Available online at www.sciencedirect.com

ScienceDirect

Cognitive and Behavioral Practice 26 (2019) 166-175

Cognitive and Behavioral Practice

www.elsevier.com/locate/cabp

Dropping Safety Aids and Maximizing Retrieval Cues: Two Keys to Optimizing Inhibitory Learning During Exposure Therapy

Shannon M. Blakey and Jonathan S. Abramowitz, *University of North Carolina at Chapel Hill*

The inhibitory learning model of exposure therapy posits that clinical anxiety is most effectively treated when clinicians employ strategies that maximize the (a) violation of negative expectancies and (b) generalization of nonthreat associations. Translation of basic learning research to exposure therapy via this explanatory model underscores two keys to optimizing inhibitory learning during exposure: dropping safety aids and maximizing retrieval cues. Although topographically similar, safety aids and retrieval cues are functionally distinct as well as therapeutically incompatible. In the present article, we delineate safety aids and retrieval cues in the context of exposure therapy from an inhibitory learning perspective, providing illustrative case examples of how clinicians may address the two when treating patients with clinical anxiety.

Exposure Therapy for Clinical Anxiety

Several psychological conditions recognized by the American Psychiatric Association (APA, 2013) are characterized by excessive or inappropriate anxiety that causes significant distress and functional impairment, including agoraphobia, generalized anxiety disorder (GAD), illness anxiety disorder, obsessive-compulsive disorder (OCD), panic disorder, posttraumatic stress disorder (PTSD), specific phobia, and social anxiety disorder (i.e., social phobia). Exposure-based cognitive behavioral therapy for clinical anxiety involves the guided and repeated confrontation with feared stimuli (e.g., situations, objects, thoughts/memories) while refraining from unnecessary escape, avoidance, and/or anxiety-reduction strategies (i.e., "safety behaviors"). In light of its established transdiagnostic efficacy (Olatunji, Cisler, & Deacon, 2010), exposure is considered the first-line intervention for clinical anxiety by international health care bodies (e.g., APA, 2013; NICE, 2005a, 2005b).

The Inhibitory Learning Model of Exposure

Although exposure reliably ameliorates clinical anxiety, the mechanisms through which it exerts its effects are not fully understood. One dominant explanatory model for exposure's effectiveness is emotional processing theory (EPT; Foa, Huppert, & Cahill, 2006; Foa & Kozak, 1986; Foa & McNally, 1996; Rachman, 1980). According to EPT, fear extinction results from the activation of a fear structure (a fear-based association between a stimulus and its significance; e.g., germs and the fear of becoming ill) paired with corrective information that is incompatible with the fear structure (e.g., not becoming ill after contacting a contaminated object). Critical to EPT is the concept that habituation—the natural decline in fear while in the presence of the feared stimulus—indicates revision of the fear structure. Foa and Kozak (1986) describe two types of habituation: within- and between-session habituation. Within-session habituation (WSH) represents the decline in fear a patient experiences in the presence of feared stimuli (i.e., the gradual decline of peak fear during a single exposure task) and is a prerequisite for between-session habituation (BSH; the reduction in peak fear reached between trials over the course of therapy). From an EPT perspective, WSH and BSH represent modification of the fear structure and therapeutic success.

Growing research, however, does *not* uniformly support the assumption that habituation is necessary for fear extinction (for a review, see Craske et al., 2008). To address this gap in the theory, Craske and colleagues proposed an inhibitory learning framework derived from accumulated findings from experimental animal and human laboratory studies. According to this perspective, fearful associations (e.g., germs are dangerous) are not "unlearned" or "modified" during exposure, but are instead forced to compete with newly acquired safety associations (e.g., germs do not necessarily lead to illness). Accordingly, a feared stimulus is associated with *both* its original excitatory (danger) meaning and its new inhibitory (safety) meaning following exposure. Therapeutic success is optimized by conducting exposure in ways that foster superior recall of

Keywords: anxiety; exposure therapy; inhibitory learning

the inhibitory association when encountering a feared stimulus after treatment (Laborda & Miller, 2012). Thus, the aim of exposure therapy from the inhibitory learning framework is to help patients generate and strengthen new inhibitory associations relative to older fearful associations.

Rather than emphasizing habituation, inhibitory learning theory underscores two alternate mechanisms of fear extinction: (a) violation of negative expectancies and (b) generalization of inhibitory associations. Expectancy violation refers to the discrepancy between a patient's anticipated consequence of an exposure task (e.g., being laughed at for mispronouncing a word) and the actual consequence (e.g., people don't seem to care). Therefore, inhibitory associations are generated by maximally violating a patient's fear-based predictions for feared outcomes (Rescorla & Wagner, 1972). Substantial research shows that inhibitory (safety) associations are context-specific, meaning that if safety is learned in Context A, it is not necessarily recalled in Context B (Bouton, 2002, 2004). Therefore, the inhibitory learning model of exposure also emphasizes the need to generalize safety learning by violating expectancies *repeatedly* and in a *variety* of contexts (e.g., with the therapist, at home, when alone).

Finally, although not discussed as a key mechanism during exposure, inhibitory learning theory also highlights the therapeutic advantage of developing distress tolerance, which refers to an individual's ability to withstand aversive emotional and physical states (Simons & Gaher, 2005). Distinct from distress *reduction*, distress tolerance is important within the inhibitory learning model because patients generate inhibitory associations to the extent that they persist with exposure tasks or other daily activities *despite* feeling anxious (e.g., Abramowitz & Arch, 2014; Arch & Craske, 2011). Not only does increased distress tolerance relate to increased self-efficacy for coping with exposure tasks, it also reduces the likelihood that a patient will fall back on counterproductive safety behaviors that maintain anxiety in the long term (Blakey & Abramowitz, 2016).

Clinical Implications of Inhibitory Learning Theory

Although there is some overlap between inhibitory learning and other models of exposure therapy (e.g., Salkovskis, Hackmann, Wells, Gelder, & Clark, 2006), Craske and colleagues (2014) outlined several important clinical implications of the inhibitory learning approach. Two such keys to optimizing inhibitory learning during exposure are (a) dropping safety aids and (b) maximizing retrieval cues. Safety aids are stimuli that are perceived to neutralize threat or reduce anxiety (e.g., hand sanitizer). Retrieval cues, in contrast, are stimuli associated with inhibitory associations (e.g., a chair from the therapy room in which fear extinction occurs; Craske et al., 2014). Although topographically similar, safety aids and retrieval cues are functionally distinct. Furthermore, safety aids and retrieval cues are contradictory, such that these stimuli have the potential to either impede or enhance exposure therapy, respectively. The aim of the present article is to delineate safety aids and retrieval cues in the context of exposure therapy from an inhibitory learning perspective, illustrating how clinicians may address the two when treating patients with clinical anxiety.

Dropping Safety Aids

Safety aids are stimuli that anxious patients keep available in order to prevent a feared outcome and/or alleviate associated distress (e.g., benzodiazepine medication in the case of a panic disorder patient; Westra & Stewart, 1998). In terms of basic learning theory, safety aids function as conditioned stimuli that predict the absence of the unconditioned stimulus (Craske et al., 2008, 2014; Lovibond, Davis, & O'Flaherty, 2000). Anxious patients often report keeping safety aids near at hand—even if they have never needed to use them in the past (e.g., carrying benzodiazepine medication "just in case," but never actually using it). Safety aids also can extend beyond patient resources and include the clinician and/or therapy itself. For instance, a patient's therapist might be perceived as a "safe person," as might psychoeducation materials if they are referenced for reassurance-seeking (Craske et al., 2014). Examples of common safety aids observed in various anxiety-related conditions are presented in Table 1.

Safety aids (e.g., hand sanitizer) should not be confused with safety *gear* (e.g., a prescribed rescue inhaler; Thwaites & Freeston, 2005). Whereas the former are used based on the mistaken belief that a feared catastrophe will occur if they are inaccessible (e.g., insisting that one's parent accompany them so that they will not faint during a panic attack), the latter are reasonable measures used in response to an accurate assessment of risk (e.g., someone with severe peanut allergy carrying an EpiPen). Put another way, it is largely one's beliefs about danger/safety that determine whether a particular resource functions as a safety *aid* or merely safety *gear* in a particular context. Moreover, whereas attempts to remain safe when faced with actual threat ensure survival, relying on safety aids in the absence of real threat is unnecessary and can even generate or exacerbate distress (Blakey & Deacon, 2015; Lovibond et al., 2000). As discussed next, reliance on safety aids is contraindicated by the inhibitory learning approach to exposure therapy, given that they have the potential to prevent the (a) violation of negative expectancies and (b) generalization of inhibitory associations.

Safety Aids Prevent the Violation of Negative Expectancies

Although there is evidence to suggest that anxious patients, relative to nonanxious individuals, might have

Table 1
Examples of Safety Aids

Anxiety Condition	Possible Safety Aid (and Function)
Generalized anxiety disorder	Cell phone (to check on loved ones)
	Home security camera (to check on loved ones)
	WebMD app (to check symptoms and/or seek medical reassurance)
Illness anxiety disorder	Hand sanitizer (to clean hands)
	Cleaning products (to sanitize surfaces)
	Heart rate monitor (to check cardiovascular health)
Obsessive-compulsive disorder	Hand sanitizer (to clean hands)
	Religious jewelry/icons or prayer book (to facilitate prayer rituals)
	Supplementary blind spot mirrors (to avoid hitting a pedestrian)
Panic disorder and/or agoraphobia	A "safe person" (to provide help in case of a panic attack)
	Benzodiazepine medication (to neutralize panic attacks)
	Cell phone (to call for help)
Posttraumatic stress disorder	Weapon or pepper spray (to use in self-defense)
	Alcohol or medications/drugs (to neutralize hyperarousal symptoms)
	Supplementary car mirrors (to prevent a motor vehicle accident)
Specific phobia	Muzzle (to avoid a dog bite)
	Night light (to mitigate darkness)
	Gloves (to avoid an animal bite/sting)
Social phobia	Jacket (to conceal sweating)
	Cell phone (to avoid having to talk to someone nearby)
	Pocketed clothing (to conceal shaking hands)

overall impairments in fear extinction and discriminative fear responding (e.g., Duits et al., 2015; Lissek et al., 2005), another explanation for the maintenance (or suboptimal treatment) of clinical anxiety regards safety aids. Specifically, safety aids may undermine exposure's effectiveness by attenuating the discrepancy between what a patient anticipates will happen during an exposure (i.e., a catastrophic prediction) and what actually happens (a noncatastrophic outcome). The possibility that safety aids mitigate the degree of therapeutic "surprise" following an exposure task is consistent with attribution theories of safety stimuli and exposure (e.g., Powers, Smits, Whitley, Bystritsky, & Telch, 2008; Salkovskis, 1991; Wells et al., 1995). That is, if a safety aid is perceived to mitigate the risk of a feared negative outcome, then there becomes a smaller discrepancy between what the patient anticipates will occur (i.e., a *less* catastrophic prediction) and what actually occurs (a noncatastrophic outcome). As a consequence of this failure to maximally violate a patient's negative expectancy, inhibitory learning would not be optimized (Craske et al., 2014; Lovibond et al., 2000; Rescorla & Wagner, 1972). Consider a woman with panic disorder who conducts hyperventilation exposures next to a chair (to lean or sit on) and with her benzodiazepine medication nearby. If these safety aids cause the patient to feel reassured that she will not pass out and/or artificially reduce the patient's anxious arousal during interoceptive exposure, then she may be *unsurprised* when she does not pass out, as well as misattribute the exposure's outcome to having had access to her safety aids.

Research suggests that safety aids also prevent the maximal violation of negative expectancies by making feared stimuli more noticeable and/or accessible in memory (Blakey & Abramowitz, 2016; Craske et al., 2008). Similar to threat inferences engendered by anxious arousal (Arntz, Rauner, & van den Hout, 1995) or safety behaviors (e.g, Deacon & Maack, 2008; Gangemi, Mancini, & van den Hout, 2012; Olatunji, Etzel, Tomarken, Ciesielski, & Deacon, 2011; van den Hout et al., 2014), it is possible that the mere availability of a safety aid generates anxiety among patients because the presence of a safety aid implies concurrent related threat (e.g., "Why would there be a muzzle on that dog unless the dog were aggressive?"). In this sense, the oft-used term "safety signals" to describe safety aids might be a misnomer, given that safety aids have the potential to paradoxically function as *danger* signals (Blakey & Deacon, 2015).

Safety aid–related threat inferences might also have the domino effect of activating additional cognitive-behavioral processes that maintain clinical anxiety and interfere with inhibitory learning during exposure (e.g., hypervigilance and attention biases; Blakey & Deacon, 2015; Wegner, 1997). Consider a man with religious obsessions and compulsions who fears that his unwanted blasphemous thoughts will offend God. To facilitate compulsive prayer rituals after he experiences

such obsessions, this man grips the cross he wears around his neck—his safety aid. Now, imagine that later, when washing his hands in the restroom, he leans over the sink and notices his cross in the mirror. The sight of his necklace reminds him of his blasphemous obsessions and causes him to be exquisitely sensitive to any blasphemous thoughts, which paradoxically causes him to notice *even more* unwanted obsessions (Abramowitz & Jacoby, 2014; Abramowitz, Tolin, & Street, 2001). Unable to tolerate the anxiety associated with his seemingly uncontrollable obsessions, he engages in rituals and other safety behaviors that serve to maintain his OCD symptoms in the long term (e.g., Salkovskis, 1999).

Other research shows that threat-focused attention biases associated with the presence of safety aids interfere with the processing of corrective information critical to inhibitory learning and therapeutic success. For example, Thorpe and Salkovskis (1998) offered preliminary evidence that anxious individuals simultaneously attend to both threat cues (i.e., spiders) *and* safety cues (i.e., an exit), perhaps out of a need to know when to actively seek safety from threat. In a claustrophobia study, Sloan and Telch (2002) found that relative to participants encouraged to use available safety aids during exposure (e.g., a small window that might be opened to allow fresh air), participants who were instead encouraged to focus their attention on challenging threat beliefs (i.e., negative expectancies) demonstrated superior treatment gains at follow-up. In a similar study, Sy, Dixon, Lickel, Nelson, and Deacon (2011) found that the tendency for claustrophobic individuals to infer danger from available safety aids (e.g., a two-way radio that may be used to communicate with experimenters during an exposure to an enclosed chamber) was associated with greater peak fear and poorer claustrophobia cognition outcomes posttreatment. Collectively, these studies highlight the potential for safety aids to counterintuitively *hinder,* rather than facilitate, durable inhibitory learning.

Safety Aids Prevent the Generalization of Inhibitory Associations

Safety aids also have the potential to interfere with exposure therapy by contextualizing inhibitory learning to specific contexts (Craske et al., 2008). Research indicates that whereas fear conditioning can be extremely powerful after only a few trials, fear *extinction* is relatively slower and more context-dependent (Bouton, 2002, 2004). Numerous studies of animals and humans indicate that both classical (i.e., Pavlovian) and operant (i.e., stimulus–response) conditioning are subject to the influence of "the constellation of cues that are in the background wherever learning occurs" (Bouton, 2002, p. 977). One common paradigm that reliably shows the context renewal effect (i.e., relapse, "return of fear") is an "ABA renewal" paradigm, in which fear conditioning, fear extinction, and testing occur in contexts A, B, and A, respectively.[1] Findings from these studies suggest that inhibitory associations that are created in one context *via* exposure may be overcome by older, fearful associations when tested in a different context (Bouton, 2002, 2004; Craske et al., 2014).

To use a clinical example, imagine that a woman seeks treatment for claustrophobia, which interferes with her ability to take her office building elevator (Context A). If the patient's clinician only conducts claustrophobia exposures using elevators in the treatment setting (Context B), this patient is likely to experience a surge of fear upon attempting to ride the elevator at work (Context A). Although geographic location is an obvious source of ABA context shift, a *safety aid* might similarly denote an extinction context that is dissimilar to the posttreatment testing context (i.e., the "real world"). To illustrate, if a man with social phobia conducts exposure trials with access to his handkerchief (a safety aid that he keeps in his pocket "just in case" he needs to wipe off forehead sweat while nervously conversing with others), encountering a social situation without access to his handkerchief may result in greater anxiety than it would if the man had conducted exposure trials without access to the handkerchief safety aid.

Summary

In summary, inhibitory learning theory posits that safety aids interfere with exposure by preventing the (a) violation of negative expectancies and (b) generalization of inhibitory associations to other contexts. Although research on the effects of safety behaviors on exposure outcomes is mixed (e.g., Blakey & Abramowitz, 2016), empirical research does not support the continued use of safety aids in the context of exposure for clinical anxiety. Accordingly, it is recommended that clinicians help their patients drop safety aids as quickly as patients are willing (mirroring the convention of eliminating safety behaviors during exposure therapy for clinical anxiety as fast as patients are willing). This might include instructing patients to leave safety aids (e.g., water bottles, cell phones, anxiolytic medication) at home or outside the therapy room during treatment sessions.

Maximizing Retrieval Cues

To counter the fact that inhibitory learning is relatively context-specific, some investigators have begun to consider whether stimuli associated with the fear extinction

[1] Other context renewal paradigms (e.g., AAB, ABC) show similar findings, but ABA paradigms are most often used in experimental research (Bouton, 2002, 2004).

Table 2
Examples of Retrieval Cues

External cues	Internal cues
Therapy handouts/forms	Physiological arousal
Jewelry/accessories worn during successful exposure tasks (e.g., wristband)	Psychological/affective state (e.g., anxiety, disgust)
Post-It (sticky) notes with reminder "code phrases"	Visualization/mental imagery of previous successful exposures
Background noises (e.g., city traffic) or visual stimuli (e.g., downtown buildings) present during extinction	Refrains and "code phrases" learned during therapy that facilitate mental reinstatement (e.g., *lean into the discomfort*)

context could be used to *enhance* long-term inhibitory learning. Burgeoning literature suggests that certain stimuli serve as memory retrieval cues, increasing the likelihood that patients will retrieve an inhibitory association (e.g., "trash cans are not dangerous") rather than a fear-based association in other contexts after extinction has occurred (Brooks & Bouton, 1994; Dibbets & Maes, 2011; Vansteenwegen, Vervliet, Hermans, Beckers, Baeyens & Eelen, 2006). Whereas safety aids (stimuli that a patient perceives to directly prevent a catastrophic outcome) are contraindicated during exposure from an inhibitory learning perspective, retrieval cues (stimuli that bring to mind the inhibitory association between a fear cue and noncatastrophic outcome) are hypothesized to maximize exposure by promoting the generalization of inhibitory associations to varied environmental and psychological contexts.

Retrieval Cues Help Generalize Inhibitory Associations

Retrieval cues have the potential to protect patients from relapse by making inhibitory associations more accessible in multiple, diverse settings. In this sense, strategic use of extinction-related cues could serve to enhance treatment outcome by psychologically bridging the gap between the extinction and "real-world" contexts posttreatment. One study, for example, examined the effects of extinction context stimuli on return of fear in a novel setting (Culver et al., 2011). These investigators found that when a clipboard and distinctive pen present during extinction trials (retrieval cues) were also visible during the follow-up test, there was a modest[2] protection against return of fear in a sample of 40 students with elevated public speaking fears. Thus, integrating certain external cues associated with fear extinction (e.g., a patient's watch worn during a series of successful exposure tasks, the sound of traffic outside the clinician's office window) outside of or after treatment might sustain exposure learning in the long-term by priming the retrieval of inhibitory, relative to excitatory, associations. Accordingly, clinicians might consider ways to integrate retrieval cues into treatment to ensure that patients surround their natural environments with as many reminders of the fear extinction context as possible (Bouton, 2002; Smith, 1979). Select examples of common ("everyday") stimuli that might be used as retrieval cues are shown in Table 2.

Even more effective in protecting against relapse than external retrieval cues, however, is mental representation of the fear extinction context (Craske et al., 2014). In one study, spider phobic participants who were instructed to imagine what was learned during an exposure trial (i.e., mentally reinstate inhibitory associations) during follow-up assessment in a novel context 1 week later showed less return of fear (i.e., self-reported fear ratings) when reencountering the spider than did participants who were not given the instructional retrieval cue at follow-up (Mystkowski, Craske, Echiverri, & Labus, 2006). These findings are consistent with other research on mental imagery/visualization and performance (e.g., Nordin & Cumming, 2008; Pham & Taylor, 1999). Applied to exposure therapy for clinical anxiety, these findings suggest that prompting patients to deliberately recall successful exposure outcomes in naturalistic settings prevents a partial return of fear or a full relapse. For example, a clinician could encourage patients terminating treatment to think carefully and vividly about what was learned during exposure therapy when encountering previously feared stimuli or situations in the future. Additionally, clinicians might deliberately incorporate "code phrases" or refrains during therapy (e.g., "lean into the anxiety") and help patients to mentally activate these mantras when encountering feared stimuli outside of treatment.

Importantly, retrieval cues might also enhance exposure by augmenting a patient's self-efficacy for tolerating distress. That is, if a patient sees a retrieval cue associated with the fear extinction context, he or she might recall previous successful exposure tasks and mentally reactivate associated emotions (e.g., pride, confidence in his or her ability to manage anxiety). In this way, retrieval cues have the potential to reinstate the *psychological context* associated with fear extinction, which would maximize long-term outcome. This possibility is consistent with priming research showing

[2] Although there were large effects (η^2 = .23) for physiological measures, there were no significant effects of retrieval cues on fear renewal on behavioral measures.

that activating certain cognitive schemas facilitates prime-consistent behavior (e.g., Bargh, Chen, & Burrows, 1996; Dijksterhuis & van Knippenberg, 1998).

Summary

Through the lens of inhibitory learning theory, retrieval cues carry the potential advantage of extending inhibitory associations to diverse physical and psychological contexts, which is thought to enhance exposure's long-term efficacy (Craske et al., 2014). However, the extent to which retrieval cues influence the violation of negative expectancies—another mechanism postulated to drive inhibitory learning—is unclear. Additionally, if retrieval cues are misused and deteriorate into safety aids (as discussed next), then such stimuli are likely to hinder durable fear extinction. If used in a way that is consistent with inhibitory learning theory, however, retrieval cues could serve to maximize exposure therapy by optimizing inhibitory learning.

The distinction between a retrieval cue and a safety aid is a nuanced one that is best determined *via* individualized functional assessment. That is, because the same stimulus might function differently between patients (or within an individual patient over time, as illustrated in the following case example), it is critical that the clinician carefully assess the relationship between a stimulus, the patient's beliefs about the stimulus, and the corresponding effect on affect and behavior. If the patient believes that a stimulus helps to prevent harm (or carries other superstitious significance), then the stimulus is likely a safety aid that artificially reduces anxiety elicited when confronting a feared situation. Yet if the stimulus is *not* perceived by the patient to be related to threat, but rather functions as an implicit reminder of an inhibitory association (i.e., it serves to make the inhibitory association more accessible in memory such that a patient does not experience anxiety in the first place), then the stimulus has the potential to enhance long-term outcome by functioning as a retrieval cue. As we discuss later in this article, some experts question the necessity and/or long-term utility of introducing retrieval cues into exposure therapy for clinical anxiety (Dibbets, Havermans, & Arntz, 2008; Jacoby & Abramowitz, 2016; Mystkowski & Mineka, 2007). Thus, clinicians who elect to incorporate retrieval cues with their patients are advised to do so in a cautious and time-limited manner.

Application: Delineating and Addressing Safety Aids and Retrieval Cues in Practice

Given their topographic similarity, it can be difficult to distinguish a safety aid from a retrieval cue in practice. However, theory-driven functional assessment of the stimulus in question is useful when trying to differentiate a particular cue as therapeutic versus deleterious. Whereas retrieval cues activate the inhibitory associations generated during exposure (e.g., the sight of the clinician reminds the patient of interoceptive exposure tasks that did *not* result in heart attack), safety signals are perceived by the patient to be the *reason* for the noncatastrophic outcome (e.g., the only reason he or she did not have a heart attack was that the clinician was present). One simple way to determine whether a stimulus functions as a safety aid or retrieval cue is to conduct an exposure task with and without the stimulus in question. If removal of the stimulus causes the patient anxiety and/or results in an increased prediction of exposure-related harm (e.g., "driving to the next town over without my cell phone in the car makes me fear that I could have a panic attack, run off the road, and die alone in a ditch somewhere"), it is likely that it was functioning as a maladaptive safety aid (e.g., Salkovskis, 1991). In contrast, if a stimulus serves as an *occasion setter* (Bonardi & Ward-Robinson, 2001) and increases the likelihood that safety-based associations will inhibit older, fear-based associations, then the stimulus is best conceptualized as a therapeutic retrieval cue (Craske et al., 2014). Put more simply, the question becomes: Does this object function as a "charm," or a reminder?

Next, we present a case illustration of how safety aids and retrieval cues influenced the conceptualization and treatment of a patient with social anxiety disorder.

Chris

Chris, a 34-year-old administrative assistant, sought therapy for social anxiety symptoms related to public speaking at work. Specifically, Chris reported fears that he would mispronounce his words during a conversation with his supervisor or colleagues, which he predicted would lead to him being publicly ridiculed and ultimately fired from his job. In order to minimize the likelihood of these feared outcomes, Chris went to extreme lengths to avoid social interactions at work. For example, he would pretend to be making a call on his phone whenever someone was walking in the direction of his office or else make an excuse to go to the supply room in order to avoid having to converse with his colleagues.

When assessing Chris's concerns at intake, his therapist determined that Chris's tendency to avoid or escape interpersonal interactions at work prevented him from gathering disconfirmatory information that might violate his negative expectancies that a social mishap would lead to intolerable embarrassment and loss of employment. Accordingly, the therapist began by providing psychoeducation about the "vicious cycle" of anxiety, emphasizing how although Chris's safety aid (i.e., work phone) and safety behaviors (i.e., hiding in the supply room) effectively reduced his social anxiety in the moment, they ironically contributed to the maintenance of his anxiety in the long-term by preventing him from gathering information that would dispute his fearful predictions.

After reviewing the rationale for exposure therapy, Chris agreed to begin exposure and response prevention, which involved eliminating counterproductive safety-seeking strategies (e.g., using his office phone to avoid social interactions) and *in vivo* exposures to social interactions. Exposure tasks involved initiating conversations with strangers on the medical campus where his therapist worked (e.g., the pharmacist, patrons in the cafeteria) and keeping track of the number of times he mispronounced his words or caused some other social mishap. Chris was surprised to learn that after 10 conversations with strangers, he had only made two minor speech errors. After these successful exposures, Chris's therapist encouraged him to *deliberately* mispronounce words (e.g., "it's nice *wetter* we're having today") and use nonsense words during conversation (e.g., "I was planning on stopping by the *sumerator* store on my way home") in order to test the belief that such mistakes would lead to social humiliation and rejection. Again, Chris's negative expectancies were not met; people failed to notice or care about Chris's speaking errors.

After Chris successfully completed these in-session exposures, his therapist encouraged him to continue practicing social mishap exposures for homework (e.g., the grocery store near his house, the lobby receptionist at his office complex). Chris was reluctant to do so, however, explaining that whereas people at the medical campus were "safe people" (e.g., treatment providers who may be naturally empathetic, other patients who Chris would never have to see again), he was more anxious around individuals in his professional or personal networks. Accordingly, Chris's therapist encouraged him to use the watch he always wore to treatment as a retrieval cue, instructing him to look down at his watch and use it as a reminder of all of the previous exposures he successfully tackled in treatment before any "real life" social interactions.

At the next week's session, Chris reported that looking at his watch helped him recall previous exposure successes, which made him feel more confident interacting with his colleagues. Over time, however, Chris's therapist noticed that Chris's compliance to engage in social exposures was dependent on whether he was wearing his watch. In fact, Chris *credited* the watch for his successes during naturalistic exposures. Specifically, when Chris volunteered an answer to a question posed by his supervisor during a group meeting, Chris told his therapist later that week that he "couldn't have done it without his watch." At the next session, however, Chris reported that while driving to work one morning on the day of a team-building exercise (during which he would be expected to speak in front of his colleagues), he realized that he had left his watch at home. Unwilling to encounter feared social situations without the watch, he returned home to get it, arrived to work 30 minutes late, and was reprimanded by his supervisor for missing the first half of the exercise.

Upon hearing this news in session, Chris's therapist conceptualized Chris's watch as a retrieval-cue-turned-safety-aid. That is, Chris's watch—which originally functioned as an occasion setter that enhanced the retrieval of inhibitory associations in new contexts—was now a safety aid—a conditioned stimulus that Chris associated with the nonoccurrence of social catastrophes. Chris's therapist helped Chris to see that rather than serving as a reminder of the inhibitory associations formed during in-session exposures in novel settings, the watch was currently operating as a counterproductive anxiety reduction cue. Because the watch's new function paralleled that of his office phone (i.e., his pretreatment safety aid), Chris and his therapist jointly agreed that they should fade out the use of his watch during social interactions, much like they faded out Chris's use of his office phone earlier in treatment. They agreed to letting Chris use his watch as a safety aid during a meeting later that day, but leaving it at home altogether for the 4 weeks following (and instead using mental reinstatement to help Chris retrieve inhibitory associations when interacting with others at work). After a month without wearing his watch, Chris was permitted to wear it to work so long as he took it off and left it in his office when he attended meetings or other social events. At the end of treatment, when Chris no longer endorsed any social anxiety symptoms, he was told that he could wear his watch whenever he wished.

Conclusions and Future Research Directions

The inhibitory learning model is an increasingly accepted approach to conceptualizing and delivering exposure therapy. Although this model overlaps with other explanatory models in some ways (e.g., Foa et al., 2006; Salkovskis et al., 2006), it nevertheless carries unique implications for the treatment of clinical anxiety. One issue highlighted by the inhibitory learning framework is the effect of certain contextual stimuli on exposure's long-term effectiveness. In the present article, we delineated the relative influence of safety aids and retrieval cues on inhibitory learning, applying cutting-edge theoretical developments to clinical practice with illustrative case examples. We concluded with the case of Chris, demonstrating how both safety aids and retrieval cues influenced his therapist's case conceptualization and treatment planning.

Although empirical and anecdotal evidence suggests that safety aids paradoxically exacerbate, rather than alleviate, clinical anxiety, more research is needed to better understand the mechanisms through which safety aids influence inhibitory learning. For example, what are the relative effects of safety aid *availability* versus *use* on anxious patients' perceptions of threat, hypervigilance/attentional bias, and affective reactions to fear cues in the natural environment? In what situations or contexts

would a safety aid function as a *safety* signal versus a *danger* signal? Multimethod, experimental investigations on how safety aids influence cognitive-behavioral processes thought to maintain clinical anxiety would be helpful.

Some investigators question the long-term benefit of relying on retrieval cues that may not be readily available when previously feared stimuli are encountered unexpectedly or much later in time (Mystkowski & Mineka, 2007). That is, if encountering a previously feared stimulus without access to the retrieval cue results in a sudden surge of anxiety (e.g., "I won't be able to face this without _____," as in the case of Chris), then use of a retrieval cue during exposure might paradoxically leave individuals vulnerable to relapse (Dibbets et al., 2008; Jacoby & Abramowitz, 2016). Clinicians should therefore be mindful of when and how retrieval cues are introduced during exposure and are advised to use them in a time-limited manner, especially considering that research in this area is nascent and the clinical (versus statistical) significance of incorporating retrieval cues is unclear. Although research on this topic is ongoing, some investigators recommend using retrieval cues near the end of exposure therapy to prevent relapse, rather than using these cues early in therapy (when the focus is on initial generation of nonthreat associations; e.g., Bouton, 2002; Craske et al., 2014). Additional research highlighting the optimal use of retrieval cues during the course of exposure therapy is therefore warranted.

Although the present article focused on external cues that may either impede or enhance exposure, clinicians should also consider the effect of their patient's internal state on inhibitory learning (Bouton, 2004; Craske et al., 2014). That is, clinicians should remember that a patient's affective (e.g., anxiety, fear, uncertainty) and physiological state (e.g., panicked, pharmacologically sedated) may also moderate extinction recall (e.g., Bouton & Swartzentruber, 1991; Marks, Viswanathan, Lipsedge, & Gardner, 1972). As discussed in greater detail elsewhere (e.g., Craske et al., 2008), emphasizing the decrease in subjective distress within and between exposure trials (i.e., habituation) may communicate to patients that exposure tasks are safe and tolerable *to the extent that anxiety eventually comes down*. Given that fear reduction is not a reliable predictor of long-term exposure outcome, clinicians might describe within- and between-trial fear reduction as a likely, but unnecessary, result of repeated and prolonged exposure tasks. Similarly, clinicians should encourage patients to conduct exposures (and other daily activities) in a variety of psychological contexts to better generalize inhibitory associations and develop greater distress tolerance (Craske et al., 2014; Jacoby & Abramowitz, 2016).

The translation of basic learning research to the study of exposure therapy underscores several implications for the study and treatment of clinical anxiety. Although previous models of exposure therapy have been fruitful in establishing exposure therapy as an empirically supported intervention, a number of outstanding clinical challenges remain (Craske et al., 2008). Additional research is needed to determine how to maximize exposure therapy *via* optimizing inhibitory learning in typical clinical settings in order to improve long-term treatment outcomes.

Disclosure Statement

The authors declare that they have no conflict of interest (i.e., no financial, personal, or other relationships with other people or organizations within three years of beginning the work submitted that could inappropriately influence or bias this work).

References

Abramowitz, J. S., & Arch, J. J. (2014). Strategies for improving long-term outcomes in cognitive-behavioral therapy for obsessive-compulsive disorder: Insights from learning theory. *Cognitive and Behavioral Practice, 21*, 20–31. https://doi.org/10.1016/j.cbpra.2013.06.004

Abramowitz, J. S., & Jacoby, R. J. (2014). Scrupulosity: A cognitive-behavioral analysis and implications for treatment. *Journal of Obsessive-Compulsive and Related Disorders, 3*, 140–149. https://doi.org/10.1016/j.jocrd.2013.12.007

Abramowitz, J. S., Tolin, D. F., & Street, G. P. (2001). Paradoxical effects of thought suppression: A meta-analysis of controlled studies. *Clinical Psychology Review, 21*, 683–703. https://doi.org/10.1016/S0272-7358(00)00057-X

American Psychiatric Association [APA]. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Arch, J. J., & Craske, M. G. (2011). Addressing relapse in cognitive behavioral therapy for panic disorder: Methods for optimizing long-term treatment outcomes. *Cognitive and Behavioral Practice, 18*, 306–315. https://doi.org/10.1016/j.cbpra.2010.05.006

Arntz, A., Rauner, M., & van den Hout, M. (1995). "If I feel anxious, there must be danger": Ex-consequentia reasoning in inferring danger in anxiety disorders. *Behaviour Research and Therapy, 33*, 917–925. https://doi.org/10.1016/0005-7967(95)00032-S

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype-activation on action. *Journal of Personality and Social Psychology, 71*, 230–244. https://doi.org/10.1037/0022-3514.71.2.230

Blakey, S. M., & Abramowitz, J. S. (2016). The effects of safety behaviors during exposure therapy for anxiety: Critical analysis from an inhibitory learning perspective. *Clinical Psychology Review, 49*, 1–15. https://doi.org/10.1016/j.cpr.2016.07.002

Blakey, S. M., & Deacon, B. J. (2015). "If a safety aid is present, there must be danger": The paradoxical effects of hand sanitizer during a contamination exposure task. *Journal of Experimental Psychopathology, 6*, 264–277. https://doi.org/10.5127/jep.040814

Bonardi, C., & Ward-Robinson, J. (2001). Occasion setters: Specificity to the US and CS-US association. *Learning and Motivation, 32*, 349–366. https://doi.org/10.1006/lmot.2001.1089

Bouton, M. E. (2002). Context, ambiguity, and unlearning: Sources of relapse after behavioral extinction. *Biological Psychiatry, 52*, 976–986. https://doi.org/10.1016/S0006-3223(02)01546-9

Bouton, M. E. (2004). Context and behavioral processes in extinction. *Learning & Memory, 11*, 485–494. https://doi.org/10.1101/lm.78804

Bouton, M. E., & Swartzentruber, D. (1991). Sources of relapse after extinction in Pavlovian and instrumental learning. *Clinical Psychology Review, 11*, 123–140. https://doi.org/10.1016/0272-7358(91)90091-8

Brooks, D. C., & Bouton, M. E. (1994). A retrieval cue for extinction attenuates response recovery (renewal) caused by a return to the conditioning context. *Journal of Experimental Psychology: Animal Behavior Processes, 20*, 366–379. https://doi.org/10.1037/0097-7403.20.4.366

Craske, M. G., Kircanski, K., Zeilkowsky, M., Mystkowski, J., Chowdhury, N., & Baker, A. (2008). Optimizing inhibitory learning during exposure therapy. *Behaviour Research and Therapy*, *46*, 5–27. https://doi.org/10.1016/j.brat.2007.10.003

Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T., & Vervliet, B. (2014). Maximizing exposure therapy: An inhibitory learning approach. *Behaviour Research and Therapy*, *58*, 10–23. https://doi.org/10.1016/j.brat.2014.04.006

Culver, N. C., Stoyanova, M., & Craske, M. G. (2011). Clinical relevance of retrieval cues for attenuating context renewal of fear. *Journal of Anxiety Disorders*, *25*, 284–292. https://doi.org/10.1016/j.janxdis.2010.10.002

Deacon, B. J., & Maack, D. J. (2008). The effects of safety behaviors on the fear of contamination: An experimental investigation. *Behaviour Research and Therapy*, *46*, 537–547. https://doi.org/10.1016/j.brat.2008.01.010

Dibbets, P., Havermans, R., & Arntz, A. (2008). All we need is a cue to remember: The effect of an extinction cue on renewal. *Behaviour Research and Therapy*, *46*, 1070–1077. https://doi.org/10.1016/j.brat.2008.05.007

Dibbets, P., & Maes, J. H. R. (2011). The effect of an extinction cue on ABA-renewal: Does valence matter? *Learning and Motivation*, *42*, 133–144. https://doi.org/10.1016/j.lmot.2010.12.003

Dijksterhuis, A., & van Knippenberg, A. (1998). The relation between perception and behavior, or how to win a game of Trivial Pursuit. *Journal of Personality and Social Psychology*, *74*, 865–877. https://doi.org/10.1037/0022-3514.74.4.865

Duits, P., Cath, D. C., Lissek, S., Hox, J. J., Hamm, A. O., Engelhard, I. M. … Baas, J. M. P. (2015). Updated meta-analysis of classical fear conditioning in the anxiety disorders. *Depression and Anxiety*, *32*, 239–253. https://doi.org/10.1002/da.22353

Foa, E. B., Huppert, J. D., & Cahill, S. P. (2006). Emotional Processing Theory: An update. In B. O. Rothbaum (Ed.), *Pathological anxiety: Emotional processing in etiology and treatment* (pp. 3–24). New York, NY: Guilford Press.

Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear: Exposure to corrective information. *Psychological Bulletin*, *99*, 20–35. https://doi.org/10.1037/0033-2909.99.1.20

Foa, E. B., & McNally, R. J. (1996). Mechanisms of change in exposure therapy. In M. Rapee (Ed.), *Current controversies in the anxiety disorders* (pp. 329–343). New York, NY: The Guilford Press.

Gangemi, A., Mancini, F., & van den Hout, M. (2012). Behavior as information: "If I avoid, then there must be danger." *Journal of Behavior Therapy and Experimental Psychiatry*, *43*, 1032–1038. https://doi.org/10.1016/j.jbtep.2012.04.005

Jacoby, R. J., & Abramowitz, J. S. (2016). Inhibitory learning approaches to exposure therapy: A critical review and translation to obsessive-compulsive disorder. *Clinical Psychology Review*, *15*, 41–54. https://doi.org/10.1016/j.cpr.2016.07.001

Laborda, M. A., & Miller, R. R. (2012). Reactivated memories compete for expression after Pavlovian extinction. *Behavioural Processes*, *90*, 20–27. https://doi.org/10.1016/j.beproc.2012.01.012

Lissek, S., Powers, A. S., McClure, E. B., Phelps, E. A., Woldehawariat, G., Grillon, C., & Pine, D. S. (2005). Classical fear conditioning in the anxiety disorders: A meta-analysis. *Behaviour Research and Therapy*, *43*, 1291–1424. https://doi.org/10.1016/j.brat.2004.10.007

Lovibond, P. F., Davis, N. R., & O'Flaherty, A. S. (2000). Protection from extinction in human fear conditioning. *Behaviour Research and Therapy*, *38*, 967–983. https://doi.org/10.1016/S0005-7967(99)00121-7

Marks, I. M., Viswanathan, R., Lipsedge, M. S., & Gardner, R. (1972). Enhanced relief of phobias by flooding during waning diazepam effect. *The British Journal of Psychiatry: The Journal of Mental Science*, *121*, 493–505. https://doi.org/10.1192/bjp.121.5.493

Mystkowski, J. L., & Mineka, S. (2007). Behavior therapy for specific fears and phobias: Context specificity of fear extinction. In T. A. Treat, R. R. Bootzin, & T. Baker (Eds.), *Psychological clinical science: Papers in honor of Richard M. McFall* (pp. 197–222). New York: Psychology Press.

Mystkowski, J. L., Craske, M. G., Echiverri, A. M., & Labus, J. S. (2006). Mental reinstatement of context and return of fear in spider-fearful participants. *Behavior Research and Therapy*, *37*, 49–60. https://doi.org/10.1016/j.beth.2005.04.001

NICE. (2005a). *Obsessive-compulsive disorder: Core interventions in the treatment of obsessive-compulsive disorder and body dysmorphic disorder*. Clinical guideline 31. London: National Institute for Health and Clinical Excellence. Available at www.nice.org.uk.

NICE. (2005b). *Post-traumatic stress disorder (PTSD): The management of PTSD in adults and children in primary and secondary care*. Clinical guideline 26. London: National Institute for Health and Clinical Excellence. Available at www.nice.org.uk.

Nordin, S. M., & Cumming, J. (2008). Types and functions of athletes' imagery: Testing predictions from the applied model of imagery use by examining effectiveness. *International Journal of Sport and Exercise Psychology*, *6*, 189–206. https://doi.org/10.1080/1612197X.2008.9671861

Olatunji, B. O., Cisler, J. M., & Deacon, B. J. (2010). Efficacy of cognitive behavioral therapy for anxiety disorders: A review of meta-analytic findings. *Psychiatric Clinics of North America*, *33*, 557–577. https://doi.org/10.1016/j.psc.2010.04.002

Olatunji, B. O., Etzel, E. N., Tomarken, A. J., Ciesielski, B. G., & Deacon, B. J. (2011). The effects of safety behaviors on health anxiety: An experimental investigation. *Behaviour Research and Therapy*, *49*, 719–728. https://doi.org/10.1016/j.brat.2011.07.008

Pham, L. B., & Taylor, S. E. (1999). From thought to action: Effects of process- versus outcome-based mental stimulations on performance. *Personality and Social Psychology Bulletin*, *25*, 250–260. https://doi.org/10.1177/0146167299025002010

Powers, M. B., Smits, J. A., Whitley, D., Bystritsky, A., & Telch, M. J. (2008). The effect of attributional processes concerning medication taking on return of fear. *Journal of Consulting and Clinical Psychology*, *76*, 478–490. https://doi.org/10.1037/0022-006X.76.3.478

Rachman, S. (1980). Emotional processing. *Behaviour Research and Therapy*, *18*, 51–60. https://doi.org/10.1016/0005-7967(80)90069-8

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Prokasy (Ed.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York: Appleton-Century-Croft.

Salkovskis, P. M. (1991). The importance of behaviour in the maintenance of anxiety and panic: A cognitive account. *Behavioural Psychotherapy*, *19*, 6–19. https://doi.org/10.1017/S0141347300011472

Salkovskis, P. M. (1999). Understanding and treating obsessive-compulsive disorder. *Behaviour Research and Therapy*, *37*, S29–S52. https://doi.org/10.1016/S0005-7967(99)00049-2

Salkovskis, P. M., Hackmann, A., Wells, A., Gelder, M. G., & Clark, D. M. (2006). Belief disconfirmation versus habituation approaches to situational exposure in panic disorder with agoraphobia: A pilot study. *Behaviour Research and Therapy*, *45*, 877–885. https://doi.org/10.1016/j.brat.2006.02.008

Simons, J. S., & Gaher, R. M. (2005). The Distress Tolerance Scale: Development and validation of a self-report measure. *Motivation and Emotion*, *29*, 83–102. https://doi.org/10.1007/s11031-005-7955-3

Sloan, T., & Telch, M. J. (2002). The effects of safety-seeking behavior and guided threat reappraisal on fear reduction during exposure: An experimental investigation. *Behaviour Research and Therapy*, *40*, 235–251. https://doi.org/10.1016/S0005-7967(01)00007-9

Smith, S. M. (1979). Remembering in and out of context. *Journal of Experimental Psychology: Human Learning and Memory*, *5*, 460–471. https://doi.org/10.1037/0278-7393.5.5.460

Sy, J. T., Dixon, L. J., Lickel, J. J., Nelson, E. A., & Deacon, B. J. (2011). Failure to replicate the deleterious effects of safety behaviors in exposure therapy. *Behaviour Research and Therapy*, *49*, 305–314. https://doi.org/10.1016/j.brat.2011.02.005

Thorpe, S. J., & Salkovskis, P. M. (1998). Selective attention to real phobic and safety stimuli. *Behaviour Research and Therapy*, *36*, 471–481. https://doi.org/10.1016/S0005-7967(98)00054-0

Thwaites, R., & Freeston, M. H. (2005). Safety-seeking behaviours: Fact or function? How can we clinically differentiate between safety behaviors and adaptive coping strategies across anxiety disorders? *Behavioural and Cognitive Psychotherapy*, *33*, 177–188. https://doi.org/10.1017/S1352465804001985

van den Hout, M., Gangemi, A., Mancini, F., Engelhard, I. M., Rijkeboer, M. M., van Dams, M., & Klugkist, I. (2014). Behavior

as information about threat in anxiety disorders: A comparison of patients with anxiety disorders and non-anxious controls. *Journal of Behavior Therapy and Experimental Psychiatry*, *45*, 489–495. https://doi.org/10.1016/j.jbtep.2014.07.002

Vansteenwegen, D., Vervliet, B., Hermans, D., Beckers, T., Baeyens, F., & Eelen, P. (2006). Stronger renewal in human fear conditioning when tested with an acquisition retrieval cue than with an extinction retrieval cue. *Behaviour Research and Therapy*, *44*, 1717–1725. https://doi.org/10.1016/j.brat.2005.10.014

Wegner, D. M. (1997). When the antidote is the poison: Ironic mental control processes. *Psychological Science*, *8*, 148–150. https://doi.org/10.1111/j.1467-9280.1997.tb00399.x

Wells, A., Clark, D. M., Salkovskis, P., Ludgate, J., Hackmann, A., & Gelder, M. (1995). Social phobia: The role of in-situation safety behaviors in maintaining anxiety and negative beliefs. *Behavior Therapy*, *26*, 153–161. https://doi.org/10.1016/S0005-7894(05)80088-7

Westra, H. A., & Stewart, S. H. (1998). Cognitive behavioral therapy and pharmacotherapy: Complementary or contradictory approaches to the treatment of anxiety disorders? *Clinical Psychology Review*, *18*, 307–340. https://doi.org/10.1016/S0272-7358(97)00084-6

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Address correspondence to Jonathan S. Abramowitz, Ph.D., University of North Carolina at Chapel Hill, CB # 3270 (Davie Hall), Chapel Hill, NC 27599; e-mail: jabramowitz@unc.edu.

Received: October 31, 2016
Accepted: March 8, 2018
Available online 7 April 2018