Learned Helplessness at Fifty: Insights From Neuroscience

Steven F. Maier
University of Colorado at Boulder

Martin E. P. Seligman
University of Pennsylvania

Learned helplessness, the failure to escape shock induced by uncontrollable aversive events, was discovered half a century ago. Seligman and Maier (1967) theorized that animals learned that outcomes were independent of their responses—that nothing they did mattered—and that this learning undermined trying to escape. The mechanism of learned helplessness is now very well-charted biologically, and the original theory got it backward. Passivity in response to shock is not learned. It is the default, unlearned response to prolonged aversive events and it is mediated by the serotonergic activity of the dorsal raphe nucleus, which in turn inhibits escape. This passivity can be overcome by learning control, with the activity of the medial prefrontal cortex, which subserves the detection of control leading to the automatic inhibition of the dorsal raphe nucleus. So animals learn that they can control aversive events, but the passive failure to learn to escape is an unlearned reaction to prolonged aversive stimulation. In addition, alterations of the ventromedial prefrontal cortex-dorsal raphe pathway can come to subserve the expectation of control. We speculate that default passivity and the compensating detection and expectation of control may have substantial implications for how to treat depression.

Keywords: depression, dorsal raphe nucleus, hope, learned helplessness, ventromedial prefrontal cortex

In the early 1960s, Richard Solomon and his students at the University of Pennsylvania wanted to know how prior Pavlovian fear conditioning would influence later instrumental learning. To find out they restrained dogs in a hammock and the dogs got 64 mild-moderate electric shocks to their back paws, each shock heralded by a tone. Twenty-four hours later the dogs were placed in a shuttlebox and were supposed to learn to escape shock by jumping a short barrier between the two chambers. The two-factor theory of avoidance learning predicted that turning on the fear-inducing tone in the shuttlebox would generate fear and accelerate jumping. But to the experimenters’ annoyance, they could not test this because the dogs often failed to escape altogether in the shuttlebox and passively waited the shock out (Leaf, 1964; Overmier & Leaf, 1965).

We arrived at Penn as first year graduate students in 1964. We thought that a profound failure to escape was the phenomenon and we began to try to understand it. After 50 years of research we believe we finally do understand it, and this paper presents the evolution and destination of our theory.

From the beginning we thought the phenomenon looked like helplessness, as first suggested by Overmier and Seligman (Overmier & Seligman, 1967). But what, we puzzled, could helplessness consist of? How did it come about? How could we test for it?

Defining Helplessness

The intuitive notion of helplessness entails, we reasoned, the belief that nothing one does matters. This decomposes into objective and subjective helplessness. More formally, an animal is objectively helpless with respect to an important outcome (O) such as shock offset if the probability of (O) given a response (R) is not different from the probability of (O) given the absence of that response (notR). When this is true of all responses, objective helplessness exists.

But being subjectively helpless is another matter. We theorized that helplessness was cognitive and that it was learned. The animal must “detect” the lack of contingency as defined above and so must have “expected” that in the future shock would be independent of its responses. This was a radical suggestion for the 1960s. The learning theories of that era held that organisms could only learn associations or pairings, for example a response paired with shock strengthened this association (acquisition) or a response paired with no shock weakened this association (extinction). The rationale for the narrow associationistic view stemmed from behaviorists’ shunning cognitions in animals and it seemed that the integration of two conditional probabilities—the probability of
shock given a response integrated with the probability of shock in the absence of that response and then generalized across all responses—must be highly cognitive. Importantly we called the theory “learned” helplessness, rather than “conditioned” helplessness, because integrating these two conditional probabilities did not seem compatible with the associationistic premise that both Pavlovian conditioning and instrumental learning held dear. Emblematic of the tension between learning theory and cognitive theory was an encounter at the Princeton conference in which we first laid the theory out to the major learning theorists (Maier, Seligman & Solomon, 1969): Richard Herrnstein, a prominent Harvard Skinnerian, retorted, “You are proposing that animals learn that responding is ineffective. Animals learn responses; they don’t learn that anything.”

**Operationalizing Learned Helplessness**

The triadic design to be described operationalizes this definition of objective and subjective learned helplessness. To know that it is the noncontiguency between responding and shock and not the shock itself that produces later passivity, the noncontiguency has to be isolated from the shock. So three groups are needed. One group gets escapable shock (ESC) where shock offset is contingent on the animal’s response. So in the original stressor controllability experiment (Seligman & Maier, 1967), this group of dogs learned in the hammock to press a panel with their noses to turn off each shock. The second group is yoked to the ESC group. In this initial experiment, on each trial the yoked group subjects received the average duration of shock that the ESC group produced on each trial. However, in most subsequent experiments described below the yoking was done on each trial for each pair of subjects (ESC and yoked), so that the 2 groups receive exactly the same duration, intensity, and pattern of shocks, but for the yoked subjects their responses have no effect on the shock. In this inescapable shock group (INESC) shock offset and all of the animal’s response are noncontingently related. A third group (0) gets no shock.

The next day the animals are tested in a very different environment—shuttlebox escape—and the well-replicated result was that two thirds of the animals from the INESC group failed to learn to escape, whereas 90% of the animals in both the ESC and 0 groups easily learned to escape. Importantly the ESC and the 0 group learned to escape equally well. We concluded from this result that the animals in the INESC group had learned in the hammock that shock offset was independent of their responses and when shock occurred the next day in the shuttlebox they expected that its offset would once again be independent of responding. This expectation undermined their trying (“response initiation”) to escape. The fact that the escapable shock group did not do better than the zero group and that the main effect was that the inescapable group did worse than both other groups strongly influenced our belief that helplessness had been learned. It should be noted that at almost exactly the same time Weiss (1968) used a similar paradigm to study the effects of coping on ulcer formation.

In one variant Maier (1970) found that the passivity was not a superstition of acquired response. The contiguity-minded learning theorists countered the cognitive account by claiming that in the hammock, shock offset occasionally was paired with not moving in the INESC group and this “superstitiously” reinforced the association of not moving with shock offset. Hence in the shuttlebox the animals engaged in not moving and eventually shock went off—further strengthening the superstitious no-movement—shock-off association. So Maier ran a special escapable shock group in the hammock: For this group shock went off when the animal held still, explicitly reinforcing not moving—one step better than superstition. The cognitive theory predicted that these animals would not sit still in the shuttlebox since they had learned that they could control shock; whereas the associationistic theory predicted that they would show the competing response of “helplessness.” This was a crucial test of contiguity versus cognition and Maier found that this escapable shock group easily learned to escape in the shuttlebox by jumping the barrier. Hence we discarded contiguity accounts of helplessness.

This work led us to define a dimension that we called control over outcomes, with control being present whenever the probability of (O), given a response (R) is different from the probability of (O) given the absence of that response (notR). Clearly, the escapable subjects have control over an aspect of the aversive event (when each shock terminates), and the inescapable subjects do not. This was exactly why we used the triadic design and included the escapable subjects as a control group because it isolated the element of uncontrollability—if failure to escape in the shuttlebox was caused by learning uncontrollability, then this failure should not occur if uncontrollability is removed but shock stays constant. That is, we assumed that uncontrollability is the active ingredient in producing passivity, and that escapable subjects were later normal because they lacked this critical learning ingredient. The escapable group was thus really included as a control group.

We must mention that running dog experiments was a harrowing experience for both of us. We are both dog lovers and as soon as we could we stopped experimenting with dogs and used rats, mice, and people in helplessness experiments, with exactly the same pattern of results. The next section provides a brief summary of this early work.

**Animals**

Research with animals quickly switched from dogs to rodents, but using the same triadic design that compares exactly equal ESC, yoked INESC, and no shock. Behavioral work focused on two issues. The first issue was whether the later effects of IS are limited to the induction of passivity in tasks such as the shuttlebox. To summarize briefly, a wide range of other behavioral changes followed INESC including reduced aggression, reduced social dominance, reduced food and water intake, exaggerated attention to external cues, reduced preference for sweet tastes, potentiated fear conditioning, slowed fear extinction, neophobia, anxious behavior on measures such as juvenile social exploration, potentiated opioid reward, exaggerated stereotypy to stimulants, and others. Importantly, none of these occurred if the shocks were escapable (see Maier & Watkins, 2005, for a review). Clearly, some of these outcomes can be related to passivity but others cannot. Many of these other behaviors seem indicative of anxiety, by which we mean fear out of proportion with an existing threat—potentiated fear conditioning, reduced fear extinction, neophobia, reduced juvenile social exploration, avoidance of open arms on an elevated maze, and so forth. In this paper for simplicity, we will refer to this suite of changes as “passivity/anxiety,” but we emphasize that we...
are interested in all of the changes separately—but not for the purpose of the present theorizing.

The second issue concerned testing the learned helplessness hypothesis against a variety of alternative ideas that were developed to explain why the experience of INESC leads to later failure to learn to escape in a different environment and whether control/lack of control is the critical underlying dimension (summarized in Maier & Seligman, 1976). Most of these investigations were focused on why INESC produces consequences such as failure to learn to escape, not why ESC did not do so. Indeed, the ESC group was typically omitted as this was not deemed of interest. In one important exception, Volpicelli, Ulm, Altenor, and Seligman (1983) using a triadic design found that the ESC group was much superior to the zero group and the INESC group. During later inescapable shock, the ESC group which had first learned to bar press to escape shock continued to run in a shuttlebox during long duration inescapable shock. This learned “mastery” effect foreshadows the main findings of the neuroscientific work below, in which first learning about escapable shock inhibits the default response of passivity. The work of R. L. Jackson and T. Minor was also an exception. They argued that INESC produces later behavioral changes because it produces intense fear during the INESC session. With ESC, however, the organism has a “safety signal” in that proprioceptive and other feedback from the escape response is followed by a shock-free interval of time. Indeed, these stimuli are as far away from the next shock as possible, and such stimuli do become conditioned inhibitors of fear (Maier, Rapaport, & Wheatley, 1976). This was argued to greatly reduce the total fear experienced during the session, and therefore no later behavioral “symptoms” (Jackson & Minor, 1988; Minor, Trauner, Lee, & Dess, 1990; Weiss, 1971). Note that by this view there is no learning about uncontrollability or controllability, just Pavlovian processes and fear. As support, the addition of external stimuli such as a light or a tone occurring immediately after each INESC prevented the occurrence of later failure to learn to escape (Minor et al., 1990; Weiss, 1971).

The early 1970s witnessed the first research directed at understanding the neural basis of these phenomena. Of course, these investigations could use only the neuroscience tools then available, and for the most part they involved either the systemic or intracerebroventricular administration of various receptor agonists and antagonists, and a role for a number of receptors and their endogenous ligands (e.g., acetylcholine, norepinephrine, dopamine, serotonin, and adenosine) was suggested (Anisman, Remington, & Sklar, 1979; Glazer, Weiss, Pohorecky, & Miller, 1975; Petty & Sherman, 1979). Given the nature of these studies no particular circuitry or structures could be implicated. However, the work of J. Weiss was an exception. He and his colleagues showed that INESC depletes norepinephrine (NE) in the region of the locus coeruleus (LC), a brainstem cell cluster that provides NE to most of the forebrain. Locus coeruleus NE neurons express alpha-2 receptors on their soma and dendrites, and these are inhibitory autoreceptors. Thus, NE within the locus coeruleus restrains the activity of locus coeruleus neurons, and so depletion of NE within this structure actually increases the activity of NE neurons, and Weiss had shown this to be important in the development of learned helplessness, a phenomenon he called behavioral depression (see Weiss & Simson, 1988, for review). However, it was never entirely clear how or why increased activity of these neurons would produce the passivity or anxiety. In sum, by the mid 1980s there were nascent neurochemical views, but their detailed mechanism(s) of operation were necessarily murky given the tools that were then available, and their relationship to behavioral explanations unclear.

Humans

From this point on, we each went off to do other things. Seligman began to study humans exclusively. The human work went in three directions.

First, guided by the original theory, the learned helplessness procedures were replicated in apparently analogous human settings (e.g., Hiroto & Seligman, 1975). In the triadic design, for example, one group of college students received loud noise that could be escaped by button pressing, a second group was yoked, and a third group received nothing. Then they went to a human shuttlebox in which moving the hand from one side to the other turned off the noise. As with dogs and rats, most of the people from the yoked group failed to escape in the shuttlebox, whereas people from the escapable group and the zero group escaped well in the shuttlebox. Importantly the same pattern in the shuttlebox emerged when preceded by solvable and unsolvable anagrams (and no anagrams) instead of loud noise. In these studies, subjective unsystematic reports occasionally revealed that people from the inescapable group said that “nothing worked so why try?”

The second direction that Seligman took explored and manipulated the explanations people made for the causes of their failure to escape in the inescapable group (Abramson, Seligman, & Teasdale, 1978; Alloy, Peterson, Abramson, & Seligman, 1984). In the “attributional reformulation” of learned helplessness, Abramson et al. (1978) claimed that inescapability itself was not sufficient to produce anything more than momentary helplessness. Rather, the explanations that subjects made of the causes of their helplessness predicted the time course and the extent of helplessness. Subjects who attributed their helplessness to permanent causes (e.g., these problems will always be unsolvable) would show long-term helplessness in that situation. In contrast, subjects who attributed their helplessness to temporary causes (e.g., only verbal puzzles are unsolvable) would not show helplessness later in that situation. Subjects who attributed their helplessness to pervasive factors (e.g., most problems are unsolvable) would show passivity across situations, whereas subjects who attributed helplessness to local factors (e.g., this problem is unsolvable) would only show helplessness in the original situation. These predictions were tested and largely borne out (Alloy et al., 1984).

The third endeavor that Seligman pursued was the possibility that learned helplessness was a laboratory model of clinical depression (Seligman, 1974; Simson & Weiss, 1988). In the Diagnostic and Statistical Manual of Mental Disorders, third edition (DSM–III; American Psychiatric Association, 1980) and fourth edition (DSM–IV; American Psychiatric Association, 1994), major depressive disorder was diagnosed by the presence of at least 5 of the following 9 symptoms:

- Sad mood
- Loss of interest
- Weight loss
- Sleep problems
- Psychomotor problems
• Fatigue
• Worthlessness
• Indecisiveness or poor concentration
• Thoughts of suicide

Learned helplessness in the laboratory—combining the animal and human experimental results—produced eight of the nine symptoms, with the only exception being suicide and suicidal thoughts—an unlikely symptom to be produced in the laboratory by mild aversive events. Not only did inescapable shock and noise produce the symptoms of depression, but the converse occurred as well: Depressed people, who had not received inescapable events, behaved in the laboratory as if they had—showing passivity in the shuttlebox and giving up on cognitive problems (Klein, Fencil-Morse, & Seligman, 1976; Klein & Seligman, 1976; Miller & Seligman, 1976). Overall learned helplessness by mapping into the symptoms of depression seemed like a plausible laboratory model.

After 2000, Seligman went in two further new directions: First he began to work on “Positive Psychology,” the study of the causes and consequences of positive events (Seligman & Csikszentmihalyi, 2000), among them having control as opposed to being helpless. Second, he began to work on “prospection,” the study of periaqueductal gray and striatum acts at 5-HT receptors to inject neurons to numerous regions including the periaqueductal gray, striatum, and extended amygdala. 5-HT released in the neural actions, activates serotonergic (5-HT) neurons in the dorsal raphe nucleus pathway, which now operates as an EXPECT circuit. So, capable shock now activates the sensitized prelimbic-dorsal raphe nucleus pathway, which they would not otherwise do. Inescapable shock now activates the sensitized prelimbic-dorsal raphe nucleus pathway, which now operates as an EXPECT circuit. So, inescapable shock is not being detected as controllable, but it is being responded to as if it were controllable. This bias to respond as if the shock was escapable we shall call an EXPECTATION of control. However, it should be clearly understood that this EXPECTATION may not be a cognitive process or entity as psychologists tend to view them. It is a circuit that provides an expectation function, in the sense that it changes or biases how organism’s respond in the future as a consequence of the events that occur in the present.

The Neural Circuitry of Learned Helplessness

We now go through the neural circuitry dataset in detail using the PASSIVITY/ANXIETY, DETECT, ACT, and EXPECT terminology both to make the argument more easily understood and because we believe that it is a useful translation from the neural level of analysis to the psychological level of analysis. We are aware that any such translation is merely a hypothesis that can be tested and falsified.

By the mid-1990s it seemed that the neuroscience tools that had become available might allow a more detailed understanding of how the brain produces the behavioral consequences of uncontrollable aversive events. As noted above, a variety of neurotransmitters and receptors had already been implicated, but how the sequelae of inescapable shock are actually caused was obscure.

We state inclusion/exclusion criteria for any adequate neural learned helplessness study at the outset. A study must meet two criteria. First, control over the stressor must be manipulated to determine whether any neural change measured is indeed a consequence of the uncontrollability/controllability of the event. Oth
otherwise, the measured change could be a simple consequence of the stressor per se. There are numerous consequences of exposure to an aversive event that are, in fact, not modulated by control (Helmreich et al., 1999; Maier, Ryan, Barksdale, & Kalin, 1986; McDevitt et al., 2009; Woodmansee, Silbert, & Maier, 1993). Thus, it is not enough to compare only inescapable shock and nonshocked controls. In the research to be described below, rats are the subject and the response that the ESC subjects can perform to terminate each shock is the turning of a small wheel located on the front of the chamber. Of course, once having established that a particular outcome that follows a particular stressor is indeed a function of controllability, the triadic design may not then be needed in further studies designed to explore the mechanisms by which the incontrollable stressor produces behavioral outcomes. Second, the initial stressor must occur in an environment very different from the test environment since one of the hallmarks of learned helplessness is trans-situationality. When common cues are shared between the first environment and the test environment, processes such as fear conditioning could mediate the behavioral change. For example, there are a large number of reports under the label “biological mechanisms of learned helplessness” that have delivered inescapable gridshocks while the subjects are constrained to one side of a shuttlebox, and then escape learning is tested in that very same shuttlebox. Poor test shuttlebox escape learning could be mediated by fear conditioning to the shuttlebox environment, since freezing is a prominent fear response. Indeed, Greenwood, Strong, and Flesher (2010) have shown this to be the case. In their studies, manipulations that reduced fear conditioning reduce the shuttle escape deficit when the prior inescapable shocks were administered in the shuttlebox, but not when they were administered outside the shuttlebox.

Passivity/Anxiety and the Dorsal Raphe Nucleus

It is, of course, difficult to know where to start in a search for the circuitry that mediates learned helplessness. Maier and his colleagues began by reasoning backward from the behavioral sequelae of inescapable shock. As already noted, many of the behavioral consequences seemed to be captured as either inhibited fight/flight (poor escape, reduced aggression, reduced social dominance) or exaggerated fear/anxiety (decreased social investigation, potentiated fear conditioning, neophobia). By the mid-1990s there was quite a bit known about the neural circuitry that regulates fight/flight and fear/anxiety, and so this information could be used. Most behaviors and emotions are mediated not by a particular structure but rather by a circuit, so the idea was to identify structures that were the most proximal mediators of fight/flight and fear/anxiety, that is, the most efferent part of the circuit closest to the behaviors themselves. The most proximate mediator of fight/flight seemed to be the dorsal periaqueductal gray (dPAG), whereas the extended amygdala (bed nucleus of the stria terminals, BNST, together with the amygdala proper) mediated fear/anxiety.

Serotonin (5-HT) and the Dorsal Raphe Nucleus

So it seemed as if the subjects that received inescapable shock later behaved as if they had inhibited dorsal periaqueductal gray function and exaggerated amygdala/BNST function. There is a structure—the dorsal raphe nucleus—that projects to both, inhibiting one and potentiating the other when it itself is activated. Activation of this structure might then recapitulate the behavioral pattern produced by inescapable shock. The dorsal raphe nucleus sends 5-HT projections to both the dorsal periaqueductal gray and to the amygdala, with 5-HT released in the dorsal periaqueductal gray inhibiting its function and 5-HT in the amygdala potentiating its function (see Graeff, Guimarães, De Andrade, & Deakin, 1996 for review).

Clearly, then, if inescapable shock were to produce a powerful activation of the dorsal raphe nucleus 5-HT neurons and lead to the release of 5-HT in structures such as the amygdala and dorsal periaqueductal gray, then this structure would hold the potential to be a crucial node in any learned helplessness circuit. It would also have to be true that escapable shock does not activate the dorsal raphe nucleus. Of course, it was necessary to investigate whether inescapable shock does not just activate it in some nonselective way, but rather that inescapable shock activates specifically 5-HT neurons. 5-HT containing cell bodies are largely localized to the raphe nuclei, with the dorsal raphe nucleus being the largest and providing much of the 5-HT innervation of forebrain and limbic structures. However, only roughly 1/3 of dorsal raphe nucleus neurons contain 5-HT, and so simply showing generalized activation is not enough. To approach this issue Grahn et al. (1999) labeled 5-HT cells in the dorsal raphe nucleus with an antibody directed at 5-HT. Then, subjects received escapable shock (ESC), yoked inescapable shock (INESC), or no shock, and the expression of markers for neural activation was examined (e.g., the expression of the protein product of the immediate-early gene c-fos) using immunohistochemistry specifically in the cells known to be 5-HT cells. Thus, she was able to show that inescapable shock activated the neurons in the dorsal raphe nucleus that contained 5-HT, and exactly equal escapable shock did not.

The technique of in vivo microdialysis allows the measurement of the levels of 5-HT in discrete brain regions in real-time in live, awake, behaving animals. The results were dramatic. Figure 1 shows the levels of 5-HT within the dorsal raphe nucleus during

Figure 1. Levels of serotonin (5-HT) in the dorsal raphe nucleus (DRN) measured by in vivo microdialysis before, during, and after exposure to escapable (ESC) and yoked inescapable (IS) tailshocks. Level of 5-HT is expressed as a percentage of baseline values, and the Baseline, during stress, and Post-Stress is measured in 20-min intervals. INESC produced a sustained rise in levels of extracellular 5-HT, whereas levels during ESC dropped rapidly as the subjects learned the controlling response.
This process is either necessary or sufficient to produce the passivity and anxiety that follows inescapable shock. Three strategies have been adopted to determine necessity.

- Blockade of the dorsal raphe nucleus activation produced by inescapable shock. Here, activation of the dorsal raphe nucleus during inescapable shock was prevented by either lesion (Maier et al., 1993; Will et al., 2004) or microinjection of pharmacological agents that prevent dorsal raphe nucleus 5-HT activation (Maier, Grahm, & Watkins, 1995; Maier, Kalman, & Grahm, 1994). These treatments all prevented inescapable shock from producing its usual poor escape and heightened anxiety, and these subjects behaved as did nonshocked controls.

- Prevention of the desensitization of 5-HT1A receptors on dorsal raphe nucleus 5-HT neurons produced by inescapable shock. Here an antagonist to the 5-HT1A receptor was microinjected into the dorsal raphe nucleus during inescapable shock, and as above these subjects behaved later as if they had not received the inescapable shock.

- Blockade of 5-HT receptors in the dorsal raphe nucleus target regions during later testing. The argument is that failure to escape and increased anxiety occur because excessive 5-HT is released in critical target structures such as the amygdala during behavioral testing. Thus, blocking the receptors to which the 5-HT binds should eliminate the passivity and increased fear that typically occurs after inescapable shock. Indeed, microinjection of 5-HT2C antagonists directly into these structures does block the passivity and increased anxiety (Christianson et al., 2010; Strong et al., 2011).

Sufficiency of Dorsal Raphe Nucleus Activity for Passivity/Anxiety

With regard to sufficiency, simply activating the dorsal raphe nucleus by microinjecting agents into the dorsal raphe nucleus that activate 5-HT neurons should produce the same passivity and anxiety as does inescapable shock. Although there is less work directed at this issue, this appears to be the case. Direct activation of the dorsal raphe nucleus by microinjection of the GABA antagonist picROTOxin or the benzodiazepine receptor antagonist beta-carboline both produce the typical behavioral outcomes that are produced by inescapable shock (Maier, Grahm, Maswood, & Watkins, 1995; Short & Maier, 1993).

Learning: How and What Does the Dorsal Raphe Nucleus Know?

The work above indicates that dorsal raphe nucleus 5-HT neurons are selectively activated if the shock is inescapable, and that this activation is necessary and sufficient to produce passivity and anxiety. But the key question is why the dorsal raphe nucleus responds only if the shock is inescapable. The most obvious option is that the dorsal raphe nucleus DETECTS the uncontrollability of the shock. To do so the dorsal raphe nucleus would have to extract the conditional probability of the shock offset given that the wheel turn or some other escape response occurs, and the conditional probability of the shock offset occurring in the absence of those responses, and compare these two probabilities. When the probabilities are equal the shock is uncontrollable. However, to do this,
the dorsal raphe nucleus would require inputs informing it whether the motor responses have occurred and whether the shock is present or not, but the dorsal raphe nucleus does not receive these types of somatomotor and somatosensory inputs.

The next possibility is that the dorsal raphe nucleus receives greater excitatory inputs during inescapable than during escapable shock, thereby leading to more activation with inescapable shock. Indeed, a number of inputs to the dorsal raphe nucleus during stress have been discovered, but none provide more excitatory input during inescapable shock than during escapable shock. For example, recall that Weiss and his colleagues (Weiss & Simson, 1988) found that inescapable shock activates locus coeruleus noradrenergic (NE)-containing neurons. These project to the dorsal raphe nucleus, and consistent with the Weiss work, blockade of NE receptors in the dorsal raphe nucleus with a microinjected antagonist during inescapable shock eliminated the passivity and anxiety (Grahn et al., 2002). However, both escapable and inescapable shock produced exactly equal levels of locus coeruleus NE activation (McDevitt et al., 2009). That is, although locus coeruleus input to the dorsal raphe nucleus was required for learned helplessness, both inescapable and escapable shock led to equivalent inputs to the dorsal raphe nucleus. Moreover, a similar pattern was found for several other inputs to the dorsal raphe nucleus occurring during the shock. So, the conclusion is that the dorsal raphe nucleus does not receive any heightened excitation from inescapable shock relative to escapable shock.

Learning: The Ventromedial Prefrontal Cortex Does DETECT and EXPECT

In sum, a number of inputs to the dorsal raphe nucleus, using a number of different transmitters, was necessary to produce learned helplessness behaviors, but these inputs did not discriminate inescapable from escapable shock. If inescapable shock produces a much greater activation of dorsal raphe nucleus 5-HT neurons than does escapable shock, but both provide equivalent excitatory input, then there is only one obvious possibility left—the presence of control must somehow inhibit dorsal raphe nucleus 5-HT neurons that would otherwise be activated by shock per se without regard to controllability. The computational complexity of detecting the presence of control suggests a cortical process, and the dorsal raphe nucleus receives virtually all of its cortical input from the prelimbic region (PL) of the ventromedial prefrontal cortex (vmPFC; Peyron, Petit, Rampon, Jouvet, & Luppi, 1998; Vertes, 2004). Importantly, electrical stimulation of the neurons that descend from the prelimbic area to the dorsal raphe nucleus inhibits dorsal raphe nucleus neuronal activity. Although these descending neurons are glutamatergic and so excitatory, they synapse preferentially on GABAAergic interneurons in the dorsal raphe nucleus that inhibit the 5-HT cells (see Figure 2 for a cartoon). This arrangement leads to the hypothesis that escapability (control) is DETECTed by the ventromedial prefrontal cortex, and that the ventromedial prefrontal cortex then ACTs to inhibit shock-induced dorsal raphe nucleus activation. The dorsal raphe nucleus is a site of convergence that sums inputs from a number of structures themselves activated by shock (see Figure 3). One idea is that these different inputs encode different aspects of aversive events, and so the more that are activated the more serious the threat. The dorsal raphe nucleus is important because it has this integrative function, and in turn projects to structures that are the proximate mediators of passivity/anxiety, our shorthand for the various behavioral and mood changes that follow inescapable shock. Thus, the dorsal raphe nucleus plays a role with respect to passivity somewhat analogous to that of the central nucleus of the amygdala in mediating fear. However, the dorsal raphe nucleus 5-HT neurons are under the inhibitory control of the prelimbic region of the ventromedial prefrontal cortex, and the detection of escapable shock activates this top-down inhibition of the dorsal raphe nucleus. We will return to a discussion of how this detection is accomplished.

Does the ventromedial prefrontal cortex actually regulate dorsal raphe nucleus activity and passivity as specified by this model (see Figure 3)?

First, does the presence of escapable shock, but not inescapable shock, activate ventromedial prefrontal cortex neurons that project to the dorsal raphe nucleus? It would be easy to administer escapable shock, yoked inescapable shock, or no shock treatment and then determine whether the ventromedial prefrontal cortex is selectively activated by the escapable shock. However, most of the cells in the ventromedial prefrontal cortex have nothing to do with projections to the dorsal raphe nucleus, and so more is needed to indicate that the specific ventromedial prefrontal cortex pathways that project to the dorsal raphe nucleus are activated by escapable shock. To answer this question Baratta et al. (2009) microinjected a retrograde tracer into the dorsal raphe nucleus. Retrograde tracers are taken up by axon terminals within the dorsal raphe nucleus and transported back to the neuronal cell bodies. This labels all cell bodies in the ventromedial prefrontal cortex that project to the dorsal raphe nucleus. Baratta et al. (2009) then later administered escapable shock, inescapable shock, or no shock. It was then only necessary to determine whether the cells that were labeled as projecting from the ventromedial prefrontal cortex to the dorsal
The dorsal raphe nucleus were activated, which was done by examining within these labeled neurons the expression of markers of neuronal activation such as the immediate-early gene c-fos. Indeed, escapable shock but not exactly equal inescapable shock, increased c-fos protein in the labeled projecting neurons.

Second, is activation of this pathway necessary for escapable shock to reduce dorsal raphe nucleus activation and block the passivity and anxiety usually produced by inescapable shock? To answer this question Amat et al. (2005) inactivated the ventromedial prefrontal cortex-to-dorsal raphe nucleus pathway during the experience of escapable shock, inescapable shock, or no shock. This was done by microinjecting a pharmacological agent into the prelimbic area that inhibits the glutamatergic pyramidal neurons that project to the dorsal raphe nucleus (see Figure 4). The results were dramatic. Although the subjects with control learned the escape response perfectly, this learning was no longer protective—the dorsal raphe nucleus was activated as if the tailshocks were actually inescapable, and the subjects showed the passivity and heightened anxiety typical of exposure to inescapable shock. That is, inactivating the ventromedial prefrontal cortex-to-dorsal raphe nucleus pathway turned an animal with control into an animal without control.

Third, is activation of this pathway sufficient for control to reduce dorsal raphe nucleus activation and block the passivity typically produced by inescapable shock? That is, does the organism actually need to escape at all, or is the mere activation of this pathway during the shock enough? To answer this question Amat, Paul, Watkins, and Maier (2008) activated this pathway directly with a microinjected pharmacological agent during the experience of inescapable shock. Now the dorsal raphe nucleus was inhibited as if the stressor was escapable, which it was not, and passivity was prevented. That is, activating the ventromedial prefrontal

---

**Figure 3.** Schematic depiction of the proposed model. Serotonin (5-HT) neurons in the dorsal raphe nucleus (DRN) integrate stress-responsive inputs that encode different aspects of a stressor and then activate brain regions that are the proximate mediators of the behavioral effects of uncontrollable stress. Glut = glutamate; vmPFC = ventral medial prefrontal cortex; GABA = γ-aminobutyric acid; 5-HT = serotonin; DRN = dorsal raphe nucleus; habenula = habenula; LC = locus coeruleus; BNST = bed nucleus of the stria terminalis; PAG = periaqueductal gray; amygdala = amygdala; N. Acc. = nucleus accumbens.

**Figure 4.** Schematic depiction of experimental strategy to determine whether activation of the ventromedial prefrontal cortex (vmPFC) to dorsal raphe nucleus (DRN) pathway is necessary for the presence of behavioral control to be protective. Blockade of the vmPFC to DRN pathway would prevent behavioral control from activating the inhibitory GABAergic cells that control the serotonergic (5-HT) neurons. See the online article for the color version of this figure.
cortex-to-dorsal raphe nucleus pathway turned an animal without control into an animal with control.

**Detection of control.** So the presence of control activates descending inhibition of shock-induced dorsal raphe nucleus activation and thereby blocks passivity and anxiety. However, this does not mean that control is necessarily detected by the ventromedial prefrontal cortex in isolation, or by the ventromedial prefrontal cortex at all. Control could be detected by a different circuit, with this information then conveyed over to the ventromedial prefrontal cortex.

The clue to how Maier and his students proceeded with this issue came from the literature on the neural mechanisms underlying appetitive instrumental learning—for example, a rat learning to press a lever for food. The history of psychology witnessed a debate as to whether instrumental learning involves the formation of a Stimulus-Response habit or instead a Response-Reinforcer expectancy. Neuroscience research indicates that each can take place and each involves different neural systems (Balleine & O’Doherty, 2010). One system, called the “act/outcome” system, is sensitive to the contingency between response and reinforcement. Contingency is “the difference between the probability of obtaining a target reward (r) given that a specific action (a) is performed and the probability of gaining the reward in the absence of the action” (Liljeholm, Tricomi, O’Doherty, & Balleine, 2011, p. 2474). The act/outcome system leads to “flexible” learning, and it is sensitive to contingency changes in the reward. A second system, the “habit” system, encodes a mere habit that is not sensitive to contingency but only to the temporal pairing between response and reward, and it produces inflexible learning that is not sensitive to changes in the contingency of the reward (Balleine & Dickinson, 1998). Importantly, the act/outcome system involves a corticostriatal circuit consisting of the prelimbic area within the ventromedial prefrontal cortex and the posterior dorsal medial striatum (DMS), while the habit system has no prefrontal cortical involvement, but instead involves the sensorimotor cortex and the dorsal lateral striatum (DLS). Thus, lesion, NMDA receptor blockade, and inactivation of either the prelimbic area or the dorsal medial striatum prevents contingency sensitive act/outcome learning. Responses can be learned, but only the habit system is then used, and so the learning is insensitive to contingency (Shiflett & Balleine, 2011).

Interestingly, this definition of instrumental contingency is identical to the definition of control that Maier and Seligman (1976) provided, although Maier and Seligman were referring to shock rather than food. Maier and Seligman defined control as being present whenever the conditional probability of the outcome (shock termination) after some response is different from the conditional probability of the outcome in the absence of that response, an identical formalism to that provided for contingency in the appetitive instrumental learning literature 40 years later.

All this suggested that perhaps DETECTion of control over shock is done by a circuit involving the prelimbic area and the dorsal medial striatum, just as is instrumental appetitive contingency learning. As would be predicted, Amat et al. (2014) found that escapable shock, but not inescapable shock activates the contingency-sensitive dorsal medial striatum, but not the habit-oriented contingency-insensitive dorsal lateral striatum. The impact of inactivating the dorsal medial striatum during the experience of escapable and inescapable shock was even more intriguing.

First, inactivating the dorsal medial striatum did not interfere with the learning and performance of the escape response. This suggests that the rats could use the “dumber” habit system to acquire and perform the escape response. Dramatically, even though the escapable subjects performed the controlling escape response perfectly, this was not protective against passivity/anxiety later. So when the habit system and not the contingency-sensitive system is used, detecting control seems to be absent. That is, the dorsal raphe nucleus was activated as if the shocks were uncontrollable and passivity/anxiety followed. Thus, it is not turning the wheel and actually escaping the shock by wheel turning that is necessary to prevent later passivity, but rather the detection of escapability by the prefrontal-correlated dorsal striatum act/outcome system. As this conclusion also predicts, inactivating the dorsolateral striatum and the habit system did not reduce the protective effects of escapability. It might be noted that the precise mechanism by which the prelimbic-dorsomedial striatum circuit does the “detecting” is not known.

In sum, the prelimbic region of the ventromedial prefrontal cortex is involved in two separable functions—the DETECTion of control in a circuit with the dorsal medial striatum and then ACTing to inhibit the dorsal raphe nucleus (see Figure 5). Is it the same prelimbic neurons that are involved in both detection of contingency and ACTing on this information by transmitting it to the dorsal raphe nucleus and inhibiting it? To answer this question M. Barratta (unpublished) in Maier’s group microinjected one color retrograde tracer in the dorsal medial striatum and a differently colored retrograde tracer into the dorsal raphe nucleus. These tracers were transported backward along the neurons to the cell bodies of the neurons in the prelimbic area that project to these structures. If the same prelimbic neurons project to both regions then both colors should be present in the same cell bodies in the prelimbic area. There was no colocalization at all, indicating that these are separate populations of prelimbic neurons. Thus, DETECT escapability and ACT to pass this information that then inhibits the dorsal raphe nucleus are subserved by different populations of ventromedial prefrontal cortex cells and so are truly different functions.

**Immunization and EXPECTation.** Initial exposure to escapable shock prevents or reduces the passivity/anxiety induced by
later inescapable shock, a phenomenon we called immunization (Maier et al., 1969). Two features of immunization are essential. First, it is quite enduring, but perhaps not permanent. Second, it is trans situational and so inescapable shock in one situation blocks the passivity/anxiety caused by even stressors that do not involve shock in different situations. For example, Amat et al. (2010) reported that experience with escapable shock blocked the passivity in shuttle escape and also blocked the reduced social investigation (anxiety) produced by social defeat occurring seven days later. Social defeat involves placing the experimental subject together with a larger and aggressive dominant subject. The experimental subject inevitably loses and adopts defeat postures, and so there is a strong element of uncontrollability. Here there is no shock at all, no restraint, the defeat is conducted on a different floor of the building by different experimenters, and yet escapable shock at all, no restraint, the defeat is conducted on a different floor of the building by different experimenters, and yet escapable shock immunized against the effects of defeat. As would be expected, social defeat also increased dorsal raphe nucleus 5-HT activity, and this increase was prevented by the prior escapable shock (Amat et al., 2010).

Why does the inescapable shock (or defeat) fail to activate dorsal raphe nucleus 5-HT neurons and thus produce the typical passivity/anxiety if the organism has first experienced control over shock? Given the circuitry above, perhaps the uncontrollable shock (or defeat) now activates the prelimbic-to dorsal raphe nucleus inhibitory pathway, even though without the prior immunizing experience of escapable shock it would not do so. To determine whether this is the case, Baratta et al. (2009) microinjected a retrograde tracer in the dorsal raphe nucleus to label prelimbic neurons that project to the dorsal raphe nucleus. Recall that escapable shock but not inescapable shock activates these labeled cells, as assessed by examining activation markers such as c-fos in these labeled neurons. Dramatically, inescapable shock now does activate these cells as if the inescapable shock was escapable shock—if the organism had experienced immunizing control seven days earlier! Furthermore, when these prelimbic neurons were inactivated during the inescapable shock via the microinjection of inhibitory pharmacological agents, immunization no longer occurred (Amat, Paul, Zarza, Watkins, & Maier, 2006).

Thus, the experience of escapable shock (control) produced a specific and persistent change in the prelimbic-dorsal raphe nucleus circuit that led to the inhibition of the dorsal raphe nucleus and prevented passivity in response to even uncontrollable stimulus. The obvious possibility is that the activation of this pathway that occurs during escapable shock is sufficient to produce the persistent change. To test this idea Amat et al. (2006) activated this pathway directly by microinjecting a pharmacological agent with no actual escapable shock present, but this did not produce the persistent pathway change or produce immunization. They then reasoned that perhaps it is the joint activation of this pathway and the occurrence of the shock that is critical. To test this, the prelimbic-dorsal raphe nucleus pathway was activated pharmacologically during the occurrence of inescapable shock, the uncontrollable form of the stressor. Remarkably, now immunization occurred. That is, inescapable shock produces immunization as long as the prelimbic pathway is activated during shock.

How can this be understood? It is known that without prior escapable shock activation of the prelimbic-dorsal raphe nucleus ACT pathway requires the detection of control by the prelimbic-dorsal medial striatum DETECT circuit. We suspect that it also requires the presence of a potent aversive event, as there would be no reason to inhibit the dorsal raphe nucleus if there were not an aversive event present. This assertion could be tested, although the appropriate experiment has not yet been done. Plasticity, or increased connectivity between neurons at a synapse, typically occurs when both are activated together, as they say, “neurons that fire together wire together.” Thus, the joint occurrence of shock and control might induce increased connectivity so that later just the presence of the shock, without control, is sufficient to activate the pathway. Recent evidence supports this idea: (a) The development of persistent increases in connectivity requires the production of new proteins in the cells in question, and blockade of new protein synthesis in the prelimbic area after the escapable shock experience prevents immunization (Amat et al., 2006). That is, even though the subjects exert control and the prelimbic area is activated, immunization only happens if new proteins can be formed. Importantly, control still blunts the impact of the stressor being experienced, but longer-term immunization is eliminated; (b) The production of particular proteins, called plasticity proteins, is required for long-term increases in connectivity. Increases in these particular proteins (e.g., phosphorylated extracellular signal-regulated kinase) are indeed induced in the prelimbic region of the ventromedial prefrontal cortex by escapable shock (Christianson et al., 2014); (c) Inhibitors of just these plasticity proteins prevent immunization when microinjected in the prelimbic area (Christianson et al., 2014); and (d) Direct electrophysiological measurement of projecting prelimbic neurons indicates that escapable shock, but not inescapable shock increases their excitability (Varela, Wang, Christianson, Maier, & Cooper, 2012). We conclude from this that the prelimbic-dorsal raphe nucleus ACT pathway can be modified over several hours after the joint experience of control and aversive stimulation, to respond to stressors in general as if they were controllable, and this is compatible with the idea that this altered pathway subserves the EXPECTation that shock will be controllable in new aversive situations. Thus, the same prelimbic-dorsal raphe nucleus pathway that operates as ACT can later operate as EXPECT.

**Amygdala**

The ventromedial prefrontal cortex projects to many structures other than the dorsal raphe nucleus. The amygdala is especially interesting in this regard. The role of the amygdala in fear and conditioned fear is well known. Briefly, the association between a stimulus predicting shock and the shock forms in basolateral regions of the amygdala. From there the information passes to the central nucleus of the amygdala, which in turns projects to the regions that control the behaviors and physiological responses that are the symptoms of fear (Davis, Raimie, & Cassell, 1994; LeDoux, 2003; Maren & Quirk, 2004). For example, the central nucleus of the amygdala projects to regions of the periaqueductal gray that produce freezing, a behavioral component of conditioned fear. Both the prelimbic and the infralimbic region of the ventromedial prefrontal cortex project to parts of the amygdala. Of special note, the infralimbic region sends excitatory glutamatergic projections to a region of the amygdala known as the intercalated cell region. The cells in this region are GABAergic and project to and inhibit the central nucleus of the amygdala (Berretta, Pante-zopoulos, Caldera, Pantzopoulos, & Pure, 2005). Thus, stimula-
tion of the infralimbic region should inhibit fear expression, and it does (Sierra-Mercado, Padilla-Coreano, & Quirk, 2011). Because the prelimbic-dorsal medial striatum circuit DETECTS control, and because the prelimbic communicates with the infralimbic region of the ventromedial prefrontal cortex, Baratta et al. (2007) wondered whether the experience of control over an aversive event might reduce later fear in a different situation. Thus, subjects were exposed to escapable shock or yoked inescapable shock in the wheel-turn apparatus and given fear conditioning in standard conditioning chambers seven days later. Inescapable shock potentiated later fear conditioning, a well-known phenomenon (Rau, DeCola, & Fanselow, 2005). It would not have been surprising if initial control merely prevented this potentiating effect, but it did more than that: Instead prior escapable shock actually retarded fear conditioning and facilitated fear extinction. This indicated an EXPECTation of control over shock. Moreover, these effects of prior control depended on the ventromedial prefrontal cortex (Baratta, Lucero, Amat, Watkins, & Maier, 2008), showing that this structure exerts top-down inhibition of more than just the dorsal raphe nucleus, and the limits of this arrangement await further exploration.

**Neurobiology of Human Control**

Although there is a long history of research investigating the controllability dimension in humans, studies using methods that allow the measurement of neural activity are quite recent and few in number. A number of studies employing painful stimuli have found that providing control, or inducing perceived control, reduces the experienced intensity of the painful stimulus. Moreover, perceived control in these pain studies increases ventromedial prefrontal cortex activity (Salomons, Johnstone, Backonja, & Davidson, 2004). In the only relevant triadic design of which we are aware, Kerr, McLaren, Mathy, and Niitschke (2012) used exposure to snake videos to subjects with snake phobias. Each trial began with an anticipation period of variable duration in which a cue signaled that a snake video or a neutral fish video might follow. A second cue indicated whether the subject would or would not have control over whether the video would occur on that trial. After a variable period of time a target then occurred and the subject was instructed to press it as rapidly as possible. The video or a fixation point then appeared. On a controllable trial subjects were told that if they responded fast enough the fixation point rather than the video would appear, but if they were too slow they would see the video. On uncontrollable trials the subjects were told that no matter how quickly they pressed, the video and the fixation point would each occur half the time, but subjects were asked to press as fast as possible anyway. However, the speed required on controllable trials was adjusted so that the subjects succeeded about half the time in avoiding the video, and so the actual frequencies on the uncontrollable trials was equated to this frequency. Thus, the controllable and uncontrollable trials were exactly yoked, as in animal studies. As expected, perceived control over the snake presentation reduced anticipatory anxiety on snake trials. Importantly, there was one condition that selectively excited ventromedial prefrontal cortex activity—snake controllable trials. Control did not increase ventromedial prefrontal cortex activity on neutral fish trials, even though the subjects pressed. Ventromedial prefrontal cortex activity was higher on controllable snake trials than in any of the other conditions. Finally, there was a negative relationship between ventromedial prefrontal cortex and amygdala activity on snake trials. These findings provide some support for generalizing the animal data reviewed above to humans.

**Contrasting Psychological and Neural Explanations of Learned Helplessness**

We believe that the neural explanations strongly inform the psychological explanations. We suspect that the learned helplessness work now provides a good, generalizable example of the complementarity between neural and psychological explanations. In the present case, the detailed knowledge concerning neural processes enabled the testing and major revision of the original psychological theory of learned helplessness—refinements that could not have happened without knowing the neural circuitry (examples below). On the other hand, the neural work would likely never have been done without the original behavioral work and psychological theorizing. Recall that the phenomenon that began this line of work was that exposure to aversive Pavlovian fear conditioning leads to later failure to learn instrumental escape/avoidance responses (Leaf, 1964). It was behavioral work and psychological theorizing that led to the isolation of behavioral control/lack of control as being the key feature of the Pavlovian conditioning that led to the failure to learn, and without this work there would not have been neuroscientific research directed at understanding the mechanisms that underlie controllability effects. The neuroscience circuitry work then clarified numerous issues (see below), but then translation back to psychological concepts also seems useful. As will be discussed below, the translation back to the psychological level enables the neuroscience work to potentially inform clinical practice.

**Hypothesis Testing**

First, Maier’s group was able to test hypotheses that did not seem testable at the psychological level. The psychological theorizing concerning learned helplessness flowed from the triadic design that compared subjects with no control and those with control. The basic result was that the subjects without control later revealed passivity and a number of other behavioral changes, whereas those with control did not and appeared to be similar to nonshocked controls. Given this pattern we inferred that detecting and expecting a lack of control was the active ingredient. The nondifference between the zero group and the escapable shock group led us to believe that organisms expected controllability as the basic “default option.” Alternatively, it has been argued (e.g., Minor, Dess, & Overmier, 1991) that the reverse could be true, that stressors per se have deleterious effects, and that these effects could then be blocked when control was added as the active ingredient. However, it was difficult to separate these two possibilities with behavioral experiments, and the idea that uncontrollability was learned remained the dominant view.

The neural work allowed the testing of whether control is the active ingredient and lack of control is the default option, rather than the other way around as the psychological theory claimed. There are several key points. The neural evidence strongly suggests that activation and sensitization of the dorsal raphe nucleus leads to the passivity and anxiety characteristic of learned helplessness. Of course, inescapable shock produces a greater activa-
tion of the dorsal raphe nucleus than does controllable shock. But there were two obvious possibilities as to why: Is this differential activation because inescapable shock provides more excitatory input to the dorsal raphe nucleus than does escapable shock, or is it because escapable shock provides more inhibitory input to the dorsal raphe nucleus? Either would produce differential activation of the dorsal raphe nucleus by inescapable versus escapable shock and of course, both could be true. But the neural data are clear. Inescapable shock does not provide more excitatory input—both forms of shock produce equal excitation of the dorsal raphe nucleus. However, when shock is escapable this is detected by the ventromedial prefrontal cortex-dorsal medial striatum circuit, and then the ventromedial prefrontal cortex ACTs, sending inputs to the dorsal raphe nucleus that inhibit it, thereby turning off the activation produced by shock per se. That is, there is nothing in the brain that is selectively turned on by a lack of control, only something that turns things off when there is the presence of control. So, aversive events per se (either controllable or uncontrollable) excite the dorsal raphe nucleus, but control over stress actively turns this off.

The reader may wonder why then in all the initial helplessness experiments the previously nonshocked group and the previously escapably shocked group performed equally well in shuttlebox escape. Recall that passivity/anxiety is explained by 5-HT accumulation during the testing in projection regions of the dorsal raphe nucleus that mediate these behaviors. If control leads to sensitization of the prelimbic-dorsal raphe nucleus pathway (EXPECT), then 5-HT activity should be inhibited from the start during shuttlebox escape testing in this group, but not in the nonshocked controls. It is easy to see why the inescapably shocked group should perform more poorly than controls—the dorsal raphe nucleus 5-HT neurons are sensitized in these subjects at the start of testing and so the aversive stimulus in the shuttlebox test (the gridshock) would lead to large and rapid 5-HT activation and consequent passivity/anxiety. But, if control leads to EXPECT (sensitized prelimbic-dorsal raphe nucleus inhibition), why should the escapably shocked subjects not perform better than controls that had not been previously stressed and so do not have EXPECT? The answer likely lies in an accidental feature of shuttlebox escape learning—it is learned very rapidly. Indeed, rodents escape with almost asymptotically fast latencies by the second or third trial (e.g., Prahl, Watkins, & Maier, 2000). This is likely because running is elicited as a species-specific defensive response (Bolles & Fanselow, 1980) on the very first trials. It is important to understand that 5-HT in response to aversive stimulation accumulates gradually across trials, and so the nonshocked controls learn control before 5-HT levels that could induce passivity have accumulated in regions such as the dPAG and striatum. This detection of control, would, of course, inhibit the dorsal raphe nucleus. Even if there were a slight difference in 5-HT, the shuttle response is learned so rapidly that there is a ceiling effect. This argument would suggest that in tasks in which the nonshocked control is not at ceiling a difference between the previously escapably shocked and nonshocked controls might emerge, and this appears to be the case (Baratta et al., 2007).

A second theoretical advance came from the neural circuitry: We know that the part of the brain that DETECTs control is a circuit formed by the prelimbic area of the ventromedial prefrontal cortex and the dorsal medial striatum. When this system was inactivated so that control/lack of control information could not be detected and subjects were exposed to escapable shock or inescapable shock, the rats reacted to the shock as if it were inescapable—both in terms of passivity/anxiety and of neurochemistry. The data showed that all the animals, regardless of whether they had an escape response that they learned perfectly, acted later as if the stressor had been inescapable. That is, if the control detecting circuit was taken off line, all animals acted as if the shock was inescapable whether the animal actually was able to escape the shock or not. This suggests that if the DETECT control circuit is absent the animal invariably reverts to the default of helplessness following exposure to any prolonged stressor.

The neural circuitry also allowed the test of a competing theory of learned helplessness. It had been argued that the feedback from the escape response becomes a Pavlovian inhibitor of fear, a safety signal that reduces the total fear experienced and that it is this excess fear—if unreduced—that produces passivity (see above). There is no question that the presence of safety signals that predict a period of time free from shock can reduce the behavioral impact of aversive events. With knowledge of the underlying neural circuitry it became simple to ask whether safety signals blunt the impact of stressors via the same or a different mechanism. As discussed above, the escape response exerts its behavioral effects by activating ventromedial prefrontal cortex top-down inhibition of brainstem and limbic stress-responsive structures. It is straightforward to ask whether the protective effects of safety signals also require the ventromedial prefrontal cortex, and the answer is no. For example, ventromedial prefrontal cortex lesions eliminated the ability of behavioral control to blunt the passivity and fear caused by inescapable shock, but these lesions did not even reduce the passivity and fear blunting impact of safety signals (Christianson, Benison, et al., 2008). Instead, safety signals had their impact via the insular cortex and insular cortex lesions eliminated the protective effects of safety signals. However, insular lesions did not reduce the passivity blunting effects of having an escape response, thereby demonstrating a double dissociation (Christianson, Benison, et al., 2008). Thus, control cannot be reduced to safety. This does not mean that safety signals are not stress-blunting, nor that safety signals do not have clinical uses, but only that stressor control and safety signals exert their effects via different neural mechanisms.

Another theoretical advance provided by the neural circuitry concerns understanding how experiences of control alter how organisms respond to future events. If the rats first experience is with escape the organism is immunized and reacts to subsequent stressors in new situations as if they are escapable. This suggests that the rat EXPECTs that shock will be escapable in the new situation and that plasticity in the prelimbic-dorsal raphe nucleus subserves this expectation and inhibits the dorsal raphe nucleus, thus blocking learned helplessness.

In addition to theory testing, knowledge of the underlying circuitry explained a number of learned helplessness phenomena that were simply mysteries at a psychological level. Here are two examples.

**Time course of learned helplessness.** The passivity produced by inescapable shock is transient, lasting for only a few days after the inescapable shock. If the behavioral effects of inescapable shock are mediated by the learned expectation that active responding will not produce relief, the original idea, then why this time course? No satisfactory explanation could be conjured at the psychological level. The neuroscience work predicts the time course. The passivity occurs because excessive 5-HT is released in
projection regions of the dorsal raphe nucleus, and this occurs because dorsal raphe nucleus 5-HT neurons have become sensitized due to the desensitization of 5-HT1A receptors on the soma and dendrites of these cells. Thus, these behavioral changes should exist only as long as the receptors remain desensitized, which proved to be for only a few days (Rozeske et al., 2011).

The hypothalamo-pituitary-adrenal (HPA) response. The hypothalamo-pituitary-adrenal response begins with the production of corticotropin releasing hormone (CRH) in the paraventricular nucleus of the hypothalamus. Corticotropin releasing hormone travels to the anterior pituitary where it stimulates the production and release of adrenocorticotropic hormone (ACTH) into the bloodstream. Adrenocorticotropic hormone in turn stimulates the production and release of glucocorticoids (corticosterone in the rat, cortisol in humans) from the adrenal cortex into the blood. The hypothalamo-pituitary-adrenal response is often considered to be the hallmark of the bodily reaction to stressors, so it would be natural to assume that inescapable shock would produce a larger hypothalamo-pituitary-adrenal response than equated escapable shock. But, at least in the rodent, it does not (see Hess, Linwick, Patterson, Overmier, & Levine, 1983 for different results in dogs). In the rodent, the corticosterone rise is not greater or more prolonged (Helmreich et al., 2012; Maier et al., 1986), the adrenocorticotropic hormone rise is not greater or more prolonged (Maier et al., 1986), nor is the increase in corticotropin releasing hormone in the hypothalamus larger (Helmreich et al., 1999). This unexpected finding has been inexplicable by psychological theory or behavioral considerations. Why should control modulate passivity, fear, and the neurochemical impact of a stressor but not the hypothalamo-pituitary-adrenal response? It follows from the circuitry. Retrograde and anterograde tracing studies indicate that the dorsal raphe nucleus does not send a major projection to the paraventricular nucleus of the hypothalamus, and perhaps none at all (Larsen, Hay-Schmidt, Vrang, & Mikkelsen, 1996). Thus, the dorsal raphe nucleus is not a major source of the hypothalamo-pituitary-adrenal activation during stress (Herman & Cullinan, 1997). The presence of control could reduce or inhibit paraventricular activation only if the structures that DETECT control project to the paraventricular nucleus. There is a pathway from the ventromedial prefrontal cortex to the paraventricular nucleus, but it goes through a relay in the bed nucleus of the stria terminalis, rather than directly (Radley & Sawchenko, 2011). We (Baratta 2015) thus utilized retrograde tracing techniques combined with the assessment of activation markers to determine whether the projections from the ventromedial prefrontal cortex to the bed nucleus of the stria terminalis are controllability-sensitive, and they are not. Thus, we would expect that control would not modulate the HPA axis response to stress because the paraventricular nucleus is not informed about controllability by the ACT circuit.

**What Did The Original Learned Helplessness Theory Get Right and What Did It Get Wrong?**

On the positive side, we found that as the original theory claimed, organisms are sensitive to the dimension of control, and this dimension is critical. However, the part of the dimension that is detected or expected seems now to be the presence of control, not the absence of control. It also appears that the passivity and increased anxiety that follows uncontrollable stressors for several days is not produced by any expectancy at all, but rather is an unlearned reaction to prolonged aversive stimulation that sensitizes a specific set of neurons. Importantly, the presence of control aborts this process. However, expectancy does play a role, but it does so in the immunization process. Here, an expectancy of control does blunt the impact of subsequent stressors. Clearly, the neurobiological data are at odds with the theory we held 50 years ago. When we first found that dogs given inescapable shock later failed to learn to escape in a shuttlebox, but that dogs given exactly equated escapable shock later escaped normally, the ideas that we developed were shaped by thinking of what might be most adaptive for dogs, rats and people. We reasoned that active coping is generally best because this would minimize exposure, and so we assumed that organisms would initially expect control to be possible. If the stressor proved to be uncontrollable, organisms would then learn this and expect it to be true in related situations in the future, with this expectation of uncontrollability undermining trying active coping. However, the neural basis of inescapable and escapable shock effects does not support this general schema. Instead, the presence of control seems to be the active ingredient, leading to the inhibition of threat-induced changes in limbic and brainstem structures.

Perhaps this counterintuitive arrangement becomes more intelligible if one considers our phylogenetic ancestors. In primitive organisms threats engage defensive reflexes (Walters & Erickson, 1986). However, these reflexes are energy intensive, and so unsuccessful it might be adaptive to inhibit them and conserve energy for use in physiological adjustments that promote survival, such as altering the responsivity of the immune system to be better able to fight any infection or wound that might occur after an attack (Frank, Watkins, & Maier, 2013). Primitive organisms do not have the sensory apparatus to detect threats at a distance, nor do they have a complex behavioral repertoire that can be used for what we are calling behavioral control. That is, adverse events for primitive organisms are generally uncontrollable by their voluntary behavior, and such organisms do not need mechanisms to detect controllability. Thus, the “successfulness” of defensive reflexes is likely related to the duration of the existing threat—if it is prolonged then conservation/withdrawal would be adaptive with energy being shifted to physiological adjustments to threat. It is important to note that 5-HT is phylogenetically very old (Hen, 1993). Moreover, 5-HT has, from the beginning, been involved in controlling and shifting the balance and flow of energy (see Andrews, Bharwani, Lee, Fox, & Thomson, 2015, for review).

As organisms became more complex, they could detect and identify threats at a distance. And so they developed rich behavioral and cognitive skills that could be used to cope with threats. Control became possible even against threats that persist over time. Clearly, if such control will work this is the best course because it will minimize injury and harm. For complex organisms behavioral control can be possible over threats that are repeated, intermittent, and so persist across time. Thus, conservation/withdrawal and other energy adjustments set in motion by the continuation of threat should be inhibited.

Thus, when encountering a threat we envisage the following scenario. First, defensive behavior will be elicited. The aversive event would activate structures such as the dorsal raphe nucleus. This is a cumulative process with 5-HT building over time, and if the transmitter reaches some threshold in target structures, defensive behaviors become inhibited and energy flow is shifted. However, if control is possible this is detected and leads to the inhibi-
tion of this process so that active responding can continue. In addition, plasticity is induced in the prelimbic-to dorsal raphe nucleus circuit so that the system is biased to initially react to aversive events as if they are controllable, thereby prolonging the duration of active responding.

Unresolved Issues

At the level of basic neural circuits, there are several important unresolved issues. The data suggest that there are two important circuits within the ventromedial prefrontal cortex engaged by control that mediate the protective effects of control—a prelimbic-dorsomedial striatum pathway and a prelimbic-dorsal raphe nucleus pathway. The prelimbic-dorsomedial striatum circuit detects control (DETECT) when control is present, and then activates the prelimbic-dorsal raphe nucleus pathway (ACT) to then blunt the behavioral effects of stress. Because Baratta (2015) has shown that the prelimbic neurons that participate in these two circuits are quite discrete, there has to be a pathway from somewhere in the prelimbic-dorsomedial striatum circuit that projects to and activates the prelimbic neurons in the prelimbic-dorsal raphe nucleus pathway, and this pathway is as yet unknown. The “gold standard” discovery of this pathway will require measuring the activity of the prelimbic neurons in each pathway separately, with the critical result being that activity in the prelimbic-dorsomedial striatum pathway precedes activity in the prelimbic-dorsal raphe nucleus pathway. This requires a method that allows the experimenter to know that a prelimbic neuron that is being recorded is in one of these two pathways (most prelimbic neurons are in neither), and which one. This further requires the experimenter to be able to activate or inhibit each pathway selectively. The experiments described earlier in this paper that activated or inhibited prelimbic neurons did so nonselectively, as they involved microinjecting excitatory or inhibitory drugs that would act on all prelimbic neurons.

Thus, the existing neural evidence, although strong, is not conclusive. However genetic/molecular tools are now available that allow these gold standard experiments, and they are underway in the Maier laboratory.

At the psychological level, there are several other loose ends.

As a general statement, neural processes in the prefrontal cortex become narrowed by stress (Armsten, 2015). Thus, the fact that in an aversive situation the brain seems to detect control as the active ingredient rather than a lack of control, does not mean that the brain cannot detect lack of control in other types of circumstances, such as uncontrollable food or unsolvable cognitive problems, or even loud noise. That is, the findings that we have reviewed do not imply that the brain does not have circuitry to detect noncontingency between events that include actions and outcomes. Rather, it may be that this processing can occur, but is not deployed in situations that are highly aversive such as the original helplessness experiments. So it is important to distinguish between what the brain does under a given set of conditions, and what the brain is capable of under different conditions. This possibility is in need of further research.

Speculations

The neural circuitry explains and predicts phenomena that are not explained or predicted at the psychological level. There are three main takeaways from the neural circuitry that might inform thinking about therapy and psychopathology. The first is that the default response of higher organisms to prolonged bad events seems to be passivity and heightened anxiety and that this is caused by the activation of the dorsal raphe nucleus. The second is that top-down higher cortical processes from the ventromedial prefrontal cortex inhibit this default response. The passivity and heightened anxiety symptoms of learned helplessness map quite well into symptoms of depression (Seligman, 1975; Weiss, Simon, Ambrose, Webster, & Hoffman, 1985) and perhaps those of posttraumatic stress disorder (LoLordo & Overmier, 2011). The third has to do with the well-established enduring effects of cognitive interventions (Cuijpers et al., 2013). Here we have reviewed research that indicates that the experience of behavioral control over a stressor also has an enduring impact, and have suggested that this occurs because that experience induces plasticity in ventromedial prefrontal cortex neurons that project to midbrain and brainstem stress-responsive structures, leading to their later inhibition. It is tempting to suggest that cognitive therapies operate via the same mechanism. Assuming that learned helplessness models these phenomena, we now speculate on the implications of the neural circuitry particularly for the treatment of depression.

The first blush reaction is that we should measure these structures in humans, and then excite the ventromedial prefrontal cortex and inhibit the dorsal raphe nucleus, pharmacologically, electrically, trans-magnetically or psychologically in therapy. So for example one might ask whether the dorsal raphe nucleus is highly excited during deep depression and whether it becomes less excited as depression wanes either in time or in therapy. One might ask whether medial prefrontal cortical activity inhibits the dorsal raphe nucleus when therapy or medication is successful. One might even look at the effect of medications and of trans-magnetic stimulation of dorsal raphe and medial prefrontal cortical structure during the course of depression. But a set of cautions should temper these speculations. A variety of psychological processes lead to increased ventromedial prefrontal cortex activation as measured by fMRI and also blunt the impact of stressors. However, this does not mean that these processes do so by activating the crucial PL-dorsal raphe nucleus pathway. The ventromedial prefrontal cortex is a large and complex structure encompassing cell types releasing a variety of transmitters and neuropeptides. Moreover, neurons in the ventromedial prefrontal cortex participate in numerous circuits with neurons in other brain regions, and the functions served by these circuits are likely unrelated, except that they share cells in a large piece of heterogeneous geography. The prelimbic neurons that are involved in the prelimbic-dorsomedial striatum and prelimbic-dorsal raphe nucleus pathways represent an extremely small percentage of the cells in the ventromedial prefrontal cortex and would not contribute measurably to a BOLD signal in the ventromedial prefrontal cortex. Thus, measuring a global ventromedial prefrontal cortex fMRI signal will not remotely imply activation of the control-related critical prelimbic neurons. It might seem that the dorsal raphe nucleus might be usefully imaged, but the dorsal raphe nucleus is a small structure, with roughly 25,000 5-HT cells in the rat and 150,000 in humans. Furthermore, the critical dorsal raphe nucleus 5-HT neurons that are involved in mediating the effects of uncontrollable stress and that are inhibited from the prelimbic by control are restricted to the caudal dorsal raphe nucleus (Grahn et al., 1999), maybe 8,000...
neurons in the rat and 50,000 in the human. This is too small a number of cells by at least an order of magnitude to be imaged currently.

With regard to therapy, ventromedial prefrontal cortex dysregulation and impaired top-down inhibition of stress responsive limbic and brainstem structures have often been noted (e.g., DeRubeis, Siegle, & Hollon, 2008; Hartley & Phelps, 2010; Koenigs & Grafman, 2009; Rive et al., 2013; Shin & Liberman, 2010). There is a growing and complex literature concerning the impact of therapies such as cognitive–behavioral therapy (CBT) on neural function that cannot be reviewed here. However, there are reports that CBT alters ventromedial prefrontal cortex activity and reduces negative emotion relative to wait-listed controls (e.g., Goldin et al., 2013).

Reappraisal of situations seen as catastrophic is central to CBT, and there is now a large literature that explores the brain regions that might be involved. In reappraisal research subjects make some stimulus or event seem less negative or anxiety arousing (e.g., “imagine that the snake is not poisonous and cannot get at you”). Subjects that are able to successfully reduce negative reactions show reduced amygdala activity and increased activity in lateral and dorsal regions of the prefrontal cortex (Beauregard, Lévesque, & Bourgoin, 2001). Importantly Urry et al. (2006) noted that these regions of the prefrontal cortex do not project to the amygdala, but they do project to the ventromedial prefrontal cortex. They also noted that when subjects reduced negative emotional reactions successfully, there was a strong negative correlation between amygdala and ventromedial prefrontal cortex activity. This and a variety of other evidence led Ray and Zald (2012) to conclude,

> These investigators either implicitly or explicitly describe emotion regulation as the deployment of top-down “cold” cognitive control region of the prefrontal cortex to down regulate bottom-up “hot” reactive processes involving the subcortical limbic regions like the amygdala. Failures in the successful deployment of prefrontal cortex top-down cognitive control mechanisms or overactive bottom-up amygdala processes have been proposed to contribute to several forms of psychopathology. (p. 487)

Reappraisal as a tool of therapy is behavioral control. It blunts the impact of a negative event, and likely involves top-down inhibition from the ventromedial prefrontal cortex to lower structures. Whether the ventromedial prefrontal cortex activation produced by reappraisal involves the specific pathways that are critical to mediating the impact of behavioral control awaits future research. This encourages speculation that CBT engages the same top-down protective circuitry that has been isolated in the study of behavioral control. After all, CBT teaches cognitive tools that can be used to reduce destructive negative thoughts and emotions. That is, they are taught that there are things that they can do—control.

**Why Versus Whither**

Given our caution about the multifarious functions and structures of the ventromedial prefrontal cortex, it has not escaped our notice that one such function is prospection, the representation of possible futures (Seligman et al., 2013). In their review of prospection research Gilbert and Wilson (2007) conclude “An extensive body shows that prefeeling depends critically on the ventromedial prefrontal cortex and that people with damage to this area find it difficult to predict the hedonic consequences of future events” (p. 1352).

Notice that the top-down process from the prelimbic to the dorsal raphe nucleus captures the notion of EXPECTing that future bad events will be controllable. This circuit is about the future and it buffers against, but does not annihilate, the default reaction of the dorsal raphe nucleus. A default of helplessness eventually overcome by the experience of mastery over aversive events is compatible with the ontogeny of the human species: beginning life in a state of almost utter helplessness and only gradually learning to control bad events.

The possibility that the prelimbic-dorsal raphe nucleus circuitry involves prospection points to a class of psychological interventions that should be useful and to another class of psychological interventions that should be less useful. This is at the heart of our most speculative thoughts.

The default reaction to past and present bad events may be concurrent passivity and heightened anxiety. These cannot be undone directly or annihilated. They can, however, be inhibited by top-down cortical, control. Treatment can only buffer against past and present events with moves that produce control over bad events in the future: These are therapy’s end-runs around bad events. We speculate that it is expectations of a better future that most matter in treatment. Psychotherapy might usefully spend less time on what is likely default and spend more time on what are likely the “end-runs” of DETECTing and then EXPECTing control.

“Why?” is a question that psychotherapy often asks. Perhaps a better question is “whither?” An exhaustive discussion of therapeutic moves that focus on understanding and undoing past and coping with present events as opposed to building buffers for the future is beyond the scope of this paper. So we will only give a few examples. Consider the class of therapy moves in which the patient reviews a past trauma to gain insight into its causes or to have catharsis about it. The dorsal raphe nucleus default reaction to trauma suggests that this is an uphill battle that will likely fail. The circuitry suggests that there is not much that can be achieved merely by confronting, understanding and reliving the trauma.

This uphill battle about confronting the past is not the province only of psychodynamic therapy, but it is also common in CBT. Reappraisal, in general, reinterprets a past or present bad event. All of the following are other examples of a prima facie past focus (see Dobson, 2010 for details): discussion of postevent processing and attendant rumination; discussion of memory biases like selective filtering (where the patient only attended to a negative part of something that happened and ignored the positive parts); behavior chain analysis (where the patient looks at all the steps that led up to a bad outcome, such as an eating or drinking binge, and considers how those steps set him up to ‘fail’); and functional analyses (determining antecedents, behaviors, & consequences).

Similarly many moves in CBT (see Dobson, 2010 for details) focus prima facie on coping with the present. So therapists discuss “mind reading” and “catastrophization,” to facilitate reappraisal of the attributions and meaning a person is making for an ongoing event. Attentional control and mindfulness similarly emphasize the present. Exposure therapies (and emotion-focused therapies) emphasize the present emotional experience and the patient observes how emotion changes over time when he stays in the situation.
We are quick to note, of course, that in the hands of a skilled and experienced therapist these exercises about the past and the present are typically done with the purpose of changing future behavior, such as better recognizing triggers for past maladaptive responses in order to avoid those triggers in the future or gaining insight into catastrophizing to learn how to be more optimistic in the future. Nevertheless, from our circuitry speculation, it is the preparation for the future that is likely to be the most effective ingredient and so it is worthwhile to be explicit about the locus of its effectiveness.

Indeed much of cognitive–behavioral therapy is actually future-oriented, even if it is not taught in this way. Problem solving, activity scheduling, crisis response plans, role play in assertiveness training, and what doors open when one door closes, all involve simulating future situations and trying to prepare for those effectively. We note that there actually exists a variant, as yet insufficiently validated, of CBT, called “Future Directed Therapy” (Vilhauer, 2014).

Perhaps one CBT “whither” vignette may help the clinician: A young man was distressed about the upcoming 1-year anniversary of his psychiatric hospitalization. He was afraid he would be distraught on this anniversary and engage in self-harm.

The therapist could tell that he was anxious and depressed about having been hospitalized, and what this meant for his future. He was probably having some distorted thoughts about it (e.g., “this means I’ll always be a fuck-up,” etc.). The therapist considered using classic CBT moves—inquiring about these automatic thoughts and helping him to reappraise them.

Instead, she played dumb: “Wow, the one-year anniversary. How are you going to celebrate it?” He was initially confused. Celebrate? “Well,” she replied, “you’ve obviously come a really long way since then: You’re working again, you’re in a great relationship. How would you commemorate that progress, if you wanted to?”

This totally refocused the conversation on future mastery. They also planned ways to prevent self-harm, and they planned the ways that he would capitalize on his progress in the coming year and they explored how he could continue to be in a good place one year from now.

In conclusion, the neural circuitry underlying the phenomenon of learned helplessness strongly suggests that helplessness was not learned in the original experiments. Rather, passivity and heightened anxiety are the default mammalian reaction to prolonged bad events. What can be learned is cortical—that bad events will be controllable in the future. The top-down circuitry that descends from the ventromedial prefrontal cortex down to the dorsal raphe nucleus and other structures acts to inhibit this default. We are mindful that in the theory of explanatory style, “hope” consists largely in the habit of expecting that future bad events will not be permanent, global, and uncontrollable, rather they will be temporary, local, and controllable (Seligman, 1991, pp. 48–49). Such expectations are likely the best natural defense against helplessness, and we speculate that the ventromedial prefrontal cortex-dorsal raphe nucleus circuit may be usefully thought of as the “hope circuit.”

References


Baratta, M. V. (October, 2015). Prefrontal control of resilience to adverse events. Society for Neuroscience, Chicago, IL.


This document is copyrighted by the American Psychological Association or one of its allied publishers.


Received January 13, 2016
Revision received March 7, 2016
Accepted March 9, 2016

E-Mail Notification of Your Latest Issue Online!

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at http://notify.apa.org/ and you will be notified by e-mail when issues of interest to you become available!