We describe evidence for an evolved module for fear elicitation and fear learning with four primary characteristics. First, it is preferentially activated by stimuli related to survival threats in evolutionary history. Thus, fear-relevant stimuli lead to superior conditioning of aversive associations compared with fear-irrelevant stimuli. Second, the module is automatically activated by fear-relevant stimuli, meaning that fear activation occurs before conscious cognitive analysis of the stimulus can occur. Third, the fear module is relatively impenetrable to conscious cognitive control, and fear conditioning with fear-relevant stimuli can occur even with subliminal conditioned stimuli. Fourth, the amygdala seems to be the central brain area dedicated to the fear module. Finally, we propose that there are two levels of fear conditioning, with an emotional level that is relatively independent of the cognitive contingency level, each mediated by different brain areas. Biol Psychiatry 2002;52:927–937 © 2002 Society of Biological Psychiatry

Key Words: Phobias, preparedness, fear module, selective associations, nonconscious learning, automatic fear activation

Introduction

In an evolutionary perspective, fear originates in defensive behavior systems that have helped organisms to cope with different types of survival threats (e.g., Blanchard and Blanchard 1988; Bolles 1970; Panksepp 1998). Because of their immediate survival relevance, such systems have been of central importance to mammalian evolution. Fear motivates organisms to escape and avoid sources of danger and threat with very fast, sometimes even split-second, activation of defensive behaviors (e.g., Fanselow and Lester 1988), as was necessary for survival in early evolutionary environments in which disasters could strike fast and without warning (Tooby and Cosmides 1990). Evolution has made some objects and situations innate sources of fear (e.g., Russell 1979) and has shaped some relatively hardwired and reflexive escape responses; however, all mammals are also capable of learning to fear initially neutral objects and situations that have signaled threat or danger through Pavlovian conditioning (e.g., Mineka 1979). Although there is not a complete consensus on this issue, many believe it is through such learning that many phobic fears (and other sources of anxiety seen in anxiety disorders) and avoidance behaviors develop (e.g., Mineka 1985; Mineka and Zinbarg 1996). As a product of natural selection, fear and fear learning have been shaped and constrained by evolutionary contingencies (e.g., Bolles 1970; Seligman 1970, 1971). Evidence of such shaping by evolutionary contingencies is obvious if we examine characteristics of fear and fear conditioning in humans (e.g., Öhman and Mineka 2001). For example, we are more likely to acquire fears and phobias for objects and situations that provided threats to the survival of our ancestors such as dangerous predators, heights, or wide open spaces, than for potentially dangerous objects and situations that originated with contemporary man such as guns and motorcycles, even though the latter may be equally or more likely to be associated with trauma in our everyday lives.

Selective Associations in Fear Conditioning

One highly influential (but controversial) theory about evolutionary constraints on fear conditioning was presented by Seligman (1971) who proposed a preparedness theory of phobias, which was later elaborated and expanded by Öhman and colleagues (e.g., Öhman 1979; Öhman et al 1985). The central observation of this theory was that fears and phobias do not tend to occur to an arbitrary group of objects or situations associated with trauma but rather are most likely to occur to objects and situations that were dangerous to pretechnological man. Supporting this idea were the results of three studies in which raters assessed on 5-point scales the evolutionary “preparedness” of the content and behavior of phobic patients’ fears. Objects or situations that were probably dangerous to pretechnological man under most circumstances were given a rating of 5, and those objects or situations that were unlikely to ever
have been dangerous to pretechnological man were rated as 1. With a total of more than 200 cases, each of these three studies found that raters (who were not themselves patients) placed the content of most clinical phobias in the 4–5 range (de Silva 1988; de Silva et al 1977; Zafiropoulou and McPherson 1986).1

The central component of Seligman’s proposal was that many aspects of fears and phobias can be understood in terms of selective associations in fear conditioning. Selective associations denote preferential forming of associations between some classes of stimuli (e.g., LoLordo 1979) as were first clearly demonstrated in the domain of taste aversion learning by Garcia and Koelling (1966). The most widely studied examples of selective associations for fear conditioning show that some conditioned stimuli (CSs), such as snakes or spiders, are more readily associated with certain kinds of averse unconditioned stimuli (UCSs), such as shock, than are other “unprepared” or fear-irrelevant CSs, such as flowers or mushrooms. To demonstrate that the effect specifically concerns associations, however, it is important to rule out the possibility that the fear-relevant stimulus is simply a more salient CS by demonstrating that its superior conditionability exclusively pertains to aversive UCSs. As reviewed below, there is now good evidence from various experimental paradigms that selective associations do occur in fear conditioning and that they may mediate the nonrandom distribution of fears and phobias seen clinically (see Öhman and Mineka 2001, for an extensive review).2

The Concept of the Fear Module
Although evolutionary arguments are generally somewhat controversial (for a discussion, see Öhman and Mineka 2001), we were inspired by Fodor (1983), Griffith (1997), and Tooby and Cosmides (1992) to use the concept of an evolved fear module to summarize much of what is known about fear and fear conditioning (to both “prepared” or fear-relevant stimuli and “unprepared” or fear-irrelevant stimuli; Öhman and Mineka 2001). We propose that the fear module is a relatively independent mental, behavioral, and neural system that was specifically tailored by evolutionary pressures to help solve problems of adapting to dangerous and potentially life-threatening situations frequently encountered in the ecology of our early mammalian evolutionary ancestors. The fear module, as an evolutionarily shaped behavioral system,3 is assumed to show the following characteristics: 1) selectivity with regard to input, resulting from the evolutionary history of deadly threats that have plagued mammals; 2) automaticity with regard to speed of recruitment of the behavioral and neural systems involved as a consequence of the survival premium for rapid activation; 3) encapsulation or resistance to conscious cognitive influences because the basics of fear and fear learning evolved before the emergence of conscious thought and language; and 4) specific neural circuitry (mostly in subcortical areas of the brain) that has evolved to give the module the particular adaptive characteristics that it has. We now describe each of these characteristics of the fear module more fully, providing highlights of the research supporting them. It is important to emphasize that compelling evidence for the fear module construct derives not just from strong support for one or two of these characteristics, but from evidence that the system seems to show all these features.

Selectivity of Input
Evolutionarily shaped behavioral systems are likely to be relatively selective with regard to the input to which they respond. Consistent with this premise, research shows that the fear module is especially sensitive to stimuli that seem to have been associated with recurring threatening situations in mammalian evolutionary history. Although some stimuli may innately activate the fear module, in many cases some experiential input, generally in the form of Pavlovian conditioning, is required. However, this associative apparatus for fear conditioning has been shaped by evolutionary history, and fear conditioning occurs most readily in situations that provided recurrent survival threats in mammalian evolution. Of course, fear conditioning can also occur to arbitrary stimuli that either were not dangerous in our evolutionary history, or that are currently dangerous but were not present in our evolutionary history. However, conditioning to such stimuli should (under similar conditions such as UCS intensity) be less robust in some way and show fewer of the other characteristics of the fear module noted earlier, for example, fewer signs of automaticity or encapsulation.

1 Some may wonder how phobias for flying and driving fit with the preparedness idea because airplanes and cars did not exist with pretechnological man; however, certain components of these fears clearly are prepared fears. For example, fear of flying usually stems from acrophobia (fear of heights), claustrophobia, or agoraphobia (fear of having panic attacks in enclosed or open spaces where escape might be difficult). Fear of driving is frequently secondary to agoraphobia as well; in other cases, it may stem from traumatic conditioning with a fear-irrelevant object (for example, being the victim of a car accident)—one of the relatively rare examples of unprepared phobias.

2 Although there have been many criticisms of the conditioning model of phobia acquisition, we have argued extensively elsewhere that contemporary conditioning models of phobia acquisition that take into account the complexities of the conditioning process as it is understood by learning theorists today is by far the most convincing model of phobia acquisition available (see Bouton et al 2001; Mineka 1985, Mineka and Zinbarg 1996, Mineka and Öhman 2001).
Human Experiments on Classical Conditioning of Fear

Extensive evidence from three lines of research provides good support for the concept of selective associations in fear conditioning. First, there is evidence from human experiments using classical fear conditioning paradigms; such experiments, showing similar results, have been conducted in at least six laboratories across the world, in addition to Öhman and colleagues’ laboratory in Sweden (e.g., Cook et al. 1986; Dimberg 1987; Hugdahl and Johnsen 1989; Mazurski et al. 1996; Pitman and Orr 1986; Schell et al. 1991). Briefly, as reviewed extensively by Öhman and Mineka (2001; see also Öhman 1993; Öhman et al. 2000), dozens of experiments have now demonstrated superior conditioning when slides of fear-relevant (e.g., snakes or spiders) stimuli are used as CSs and the UCS is a mild electric shock, relative to what is seen when slides of fear-irrelevant stimuli (e.g., flowers, mushrooms, geometric figures) are paired with shock. Typically in these experiments, a discriminative conditioning paradigm is used in which one fear-relevant CS is used as a CS+ and another as CS− (or one fear-irrelevant CS+ and one fear-irrelevant CS−). Such a paradigm controls for the possibility that fear-relevant CSs are simply more salient or prepotent as stimuli because superior conditioning requires the demonstration of greater responding to the fear-relevant CS+ relative to the CS−.

Even though the results from such experiments are not entirely consistent (see, e.g., McNally 1987), the failures to replicate the basic finding often involve methodologic problems (such as, for example, failing to show even basic classical conditioning effects), rendering the results inconclusive (for discussions of these failures to replicate, see Öhman 1993; Öhman and Mineka 2001). The superior conditioning is usually evidenced by enhanced resistance to extinction of the conditioned response (CR), with conditioned skin conductance responses (SCRs) being the most frequently used dependent measure. However, evidence of superior conditioning has also been shown with other psychophysiological measures such as finger pulse volume, heart rate, and cortical slow wave responses, as well as with subjective or cognitive measures of fear such as subjective ratings of acquired aversiveness and expectancy ratings of the probability of the UCS occurring (for a review, see Öhman and Mineka 2001). Moreover, further evidence of the superiority of conditioning with fear-relevant stimuli comes from an early experiment showing that conditioning can occur after only one CS–UCS pairing when fear-relevant but not fear-irrelevant CSs are used (e.g., Öhman et al. 1975).

Further evidence on differences between conditioning with fear-relevant versus fear-irrelevant stimuli shows that there is a qualitative difference in the fear responses conditioned to fear-relevant versus fear-irrelevant stimuli. For example, Cook et al. (1986) found that conditioned heart rate acceleratory responses occurred with fear-relevant CSs but that deceleratory heart rate responses occurred with fear-irrelevant stimuli. Thus, the conditioned heart rate response to fear-irrelevant stimuli conformed to the expectancy-related decelerative response of vagal origin typically seen in human conditioning (Öhman et al. 2000). The response to fear-relevant stimuli, on the other hand, was an acceleration, probably of sympathetic origin, that indexed active defense mobilization (e.g., Lang et al. 1997). That defensive responses are conditioned only to fear-relevant stimuli fits with the idea that fear-relevant stimuli have preferential access to the fear module system designed to promote defensive behavior.

This set of findings regarding superior conditioning to potentially dangerous animal stimuli such as snakes and spiders was extended to another set of fear-relevant stimuli related to social fears by Öhman and Dimberg (1978). Similar to what had been found using snakes and spiders, superior conditioning was demonstrated using angry as opposed to happy or neutral faces as CSs, which is consistent with notions of angry faces as conveying social threat (Öhman et al. 1985). Moreover, Dimberg (1987) demonstrated qualitative differences in the nature of the CRs to facial stimuli by showing that conditioning to angry faces (but not happy faces) was also evident in enhanced fear ratings, and conditioned heart rate increases, as well as in enhanced activity of the corrugator muscle controlling the frowning eyebrow response of anger. Öhman et al. (1985) presented an evolutionary framework for understanding social fears by relating them to a behavioral system controlling interactions between members of social groups. Specifically, they argued that social fears originate in dominance hierarchies that evolved to bring order into groups and minimize further aggressive encounters by having lower ranked members exhibit submissiveness (see also Mineka and Zinbarg 1995; Öhman and Wiens 2002). Thus, it was proposed that social fear and anxiety reflect exaggerated social submissiveness and fear of losing rank because of negative evaluations of other group members.

Finally, as noted earlier, to demonstrate selective associations in fear conditioning, one must demonstrate that the fear-relevant CSs that show superior conditioning when paired with aversive UCSs are not simply more salient stimuli that would also condition better with nonaversive UCSs. Accordingly, Öhman et al. (1978) examined conditioning to snakes and spiders versus flowers and mushrooms, with either an aversive shock UCS or the imperative stimulus in a reaction time task as the UCS. Several studies have demonstrated that such tasks produce conditioned SCRs (e.g., Hamm and Vaitl 1996). As expected, if
the superior conditioning with aversive UCSs reflected
selective associations, no superior conditioning to snakes
or spiders, relative to flowers or mushrooms, occurred
using the reaction time task for the conditioning paradigm.

**Monkey Experiments on Observational Conditioning of Fear**

The human classical conditioning experiments established
that there are selective associations between some classes
of fear-relevant stimuli and aversive outcomes; however,
these data have at least two important shortcomings. First,
the “fear” that is conditioned in humans is necessarily
quite mild and transient. Thus, we do not know whether
similar results would apply to the acquisition of the more
intense and persistent phobic fears that are seen clinically.
Second, it is hazardous to attribute the selective associations
documented in humans to evolutionary history such
as presumed in the fear module concept, because they
could just as well reflect cultural learning that some
creatures are dangerous. In the early 1980s, Mineka and
Cook (Mineka et al 1984; Cook et al 1985) developed an
observational fear conditioning paradigm to study the
development of snake fear in rhesus monkeys to alleviate
these shortcomings of the human database. Prior research
had demonstrated that wild-reared rhesus monkeys born in
India showed an intense phobic-like fear of snakes when
tested in the laboratory but that most lab-reared monkeys
were not afraid of snakes. The hypothesis was that the
wild-reared monkeys had developed their fear of snakes in
India through vicarious conditioning (rather than actual
traumatic conditioning with snakebites as the UCS). To
test this Mineka, Cook, and colleagues demonstrated that
lab-reared observer monkeys who simply watched wild-
reared model monkeys behaving fearfully with live and
toy snakes and nonfearfully with neutral arbitrary objects
quickly acquired an intense and persistent phobic-like fear
of snakes (Mineka 1987).

Later Cook and Mineka (1989, 1990) tested whether this
rapid acquisition of snake fear was indeed an example of
a selective association. One group of lab-reared observer
monkeys watched videotapes of model monkeys behaving
fearfully with toy snake stimuli and nonfearfully with
flower stimuli; the other group of observer monkeys
watched videotapes of model monkeys behaving fearfully
with flowers and nonfearfully with toy snakes. The fear
behavior that both groups observed was identical, which
had been accomplished through video-editing technology.
As expected, observer monkeys in the first group acquired
an intense fear of toy and real snakes, but observer
monkeys in the second group did not acquire a fear of
flowers. Parallel results were obtained in two additional
groups of observer monkeys where the videotapes they
watched showed model monkeys reacting fearfully to
either a toy crocodile (fear-relevant) or a toy rabbit
(fear-irrelevant). Furthermore, Cook and Mineka (1990)
also satisfied the second requirement for demonstration of
a selective association. Specifically, they showed that the
superior learning seen when snake stimuli were associated
with a model’s fear did not occur when snake stimuli were
used as discriminative stimuli for appetitive rewards (food
treats). Thus, selective associations have indeed been
demonstrated when strong and persistent phobic-like fears
were being vicariously conditioned in rhesus monkeys
subjects. Moreover, there is good reason to believe from
human studies using retrospective self-report that vicarious
conditioning is indeed involved in the origins of a
significant number of phobias (e.g., Ost and Hugdahl
1981; although, for a discussion of the serious limitations
of such retrospective studies, see Mineka and Öhman
2002).

**Phylogenetic versus Ontogenetic Basis for Conditioning of Fear-Related Selective Associations**

The concept of an evolved fear module rests on the assumption
that there is a phylogenetic or evolutionary basis for the
various examples of fear-relevant selective associations just
reviewed. However, the research described thus far using
human participants is not conclusive on this issue because the
fear-relevance of the stimuli used (e.g., snakes, spiders, angry
faces) could derive from culturally or ontogenetically based
influences rather than from (or in addition to) evolutionary
influences. This is because human participants entering these
studies all have prior ontogenetically based associations to
the stimuli used. In human experiments, this issue has been
addressed primarily through comparisons between effects
derived using phylogenetic fear-relevant stimuli such as
snakes or spiders and the effects derived using ontogenetic
fear-relevant stimuli such as pointed guns, knives, and
damaged electrical outlets. The fear relevance of the latter
stimuli could only stem from ontogenetic influences because
such stimuli were not present in our early evolutionary
history.

In three experiments with such comparisons using fear
conditioning paradigms, the results generally indicated
superiority of conditioning with the phylogenetically
based, fear-relevant stimuli relative to the ontogenetically-
only fear-relevant stimuli (for a review, see Öhman and
Mineka 2001). However, the strongest evidence for the
phylogenetic basis of selective associations in fear condi-
tioning comes from the studies by Cook and Mineka
fear in lab-reared monkeys. At the outset, the monkeys
used as observers in these experiments were completely
naïve with regard to the stimuli used in these experiments, having lived their entire lives in an indoor primate laboratory. Thus, the differences observed in the conditionability of fear-relevant versus fear-irrelevant stimuli must have derived from phylogenetic as opposed to ontogenetic factors. Moreover, given that the general pattern of results observed in the monkey experiments paralleled that found in the human experiments, it seems most parsimonious to argue that all the effects derived primarily from phylogenetic factors (although these sources of influence are not mutually exclusive; Öhman and Mineka 2001).

**Human Experiments Using a Covariation Bias Paradigm**

Tomarken et al (1989) developed an additional paradigm for studying fear-relevant selective associations that did not explicitly involve classical conditioning. They hypothesized that the enhanced resistance to extinction observed with fear-relevant conditioned stimuli in classical conditioning paradigms might just be one example of a more general covariation bias for fear-relevant stimuli and aversive outcomes. If this were the case, Tomarken et al predicted that selective associations should be apparent in tasks for which participants are asked explicitly to judge the association or covariation between fear-relevant stimuli and aversive outcomes. To test this hypothesis, they developed an illusory correlation paradigm containing some of the central features of Öhman and colleagues’ human classical conditioning paradigm and some features of Chapman and Chapman’s (Chapman and Chapman 1967) illusory correlation paradigm for studying clinical judgment.

Participants selected to be high or low in snake fear were exposed to a long series of slide–outcome pairs in which the slides came from one of three categories (e.g., snakes, flowers, mushrooms) and the three outcomes were aversive (mild shock) or nonaversive (tone or nothing). The three categories of slides and the three types of outcomes were, in fact, randomly related, with equal probabilities (.33) for each type of outcome to follow each category of slide. Following 72 randomly ordered trials, high and low fear participants were asked to estimate the probability that each category of slide had been followed by shock, relative to all other pertinent comparisons, which were relatively accurate; low fear participants showed a tendency in the same direction, but not all relevant comparisons were statistically significant. Moreover, participants were relatively accurate in their estimates of the base rate of occurrence of the different types of stimuli and outcomes. Thus, the overestimation of feared stimuli and aversive outcomes could be fully attributed to the overestimation of their co-occurrence, suggesting a true selective association. In a second experiment, Tomarken et al (1989) demonstrated that it was indeed the aversiveness of the shock outcome that was responsible for this effect rather than simply its high salience. These results have been replicated in several laboratories using a variety of fear-relevant stimuli ranging from blood-injury relevant stimuli (Pury and Mineka 1997) to angry facial expressions (de Jong et al 1998; Öhman and Mineka 2001).

Although Tomarken et al (1989) originally hypothesized that the covariation bias seen with fear-relevant stimuli developed over the course of the random presentations of fear-relevant stimuli and aversive outcomes, subsequent experiments showed that the covariation bias seen with fear-relevant stimuli actually seems to be present at the start of such experiments (thus an a priori expectancy bias). This a priori bias is then maintained even after the extensive exposure to random presentations of the stimuli and outcomes (an a posteriori covariation bias). By contrast, when fear-relevant stimuli of ontogenetic origins, such as weapons or electric outlets, are used, the a priori expectancy bias present at the outset disappears by the end after the extensive series of random presentations of stimuli and outcomes (e.g., Amin and Lovibond 1997; Kennedy et al 1997; McNally and Heatherton 1993). Thus, a posteriori covariation bias, reflecting an insensitivity to disconfirmation, is only present for phylogenetic fear-relevant stimuli. This is in keeping with the conclusions drawn earlier for selective associations in fear conditioning in which one of the most robust results is the pronounced resistance to extinction seen with fear-relevant stimuli paired with aversive outcomes.

**Selective Associations or Selective Sensitization?**

Although we have now summarized a large array of findings supporting the selective associability of evolutionarily fear-relevant stimuli and aversive events, other investigators have argued that associative mechanisms are not needed to account for these effects. For example, Gray (1982) and others have argued that selective sensitization, an evolutionarily shaped nonassociative process, provides a more parsimonious explanation of most of the results presented earlier (see also Lovibond et al 1993). According to this account, fear-relevant stimuli are directly encoded through genetic mechanisms to elicit fear; however, a state of arousal or anxiety is necessary before the fear actually emerges. For example, a single shock stimulus not paired with a CS is sufficient to selectively sensitise responding to a fear-relevant but not a fear-irrelevant stimulus (Öhman et al 1975). However, after reviewing the relevant studies on this topic, Öhman and
Mineka (2001) concluded that nonassociative accounts were not able to explain a number of important effects that have been reported in the literature. They noted that selective sensitization clearly exists (e.g., Öhman et al 1975) and may provide one source of input to activation of the fear module by evolutionarily fear-relevant stimuli. Nonetheless, both as traditionally conceived and as documented in laboratory experiments on this topic, it is a relatively short-lived process (e.g., a few trials) and therefore not sufficient to explain the long-lasting effects seen in prepared fear conditioning experiments nor in the acquisition of real clinical fears and phobias.

The Automaticity of Fear Activation

The second characteristic of the proposed fear module is its tendency to be preferentially activated automatically by fear-relevant stimuli. If fear evolved in animals with much more primitive brains than those of humans, it would be expected that higher order cognitive mechanisms would not be necessary for the elicitation or acquisition of fears and phobias. In fact, defensive behavior has an evolutionary origin several hundreds of millions of years ago in organisms with primitive brains incapable of advanced cognitive computations (e.g., Altman 1999). Because rapid defense recruitment continued to convey a survival advantage, the more advanced primate brain may have retained the capacity for rapid reflexive activation of defense responses independent of the more recently evolved neocortex. Thus, it would be expected that higher order cognitive mechanisms would not be necessary for the elicitation or acquisition of fears and phobias because many types of threat can be located on the basis of a rough stimulus analysis.

Over the past decade, evidence has accumulated that the fear module can be activated automatically by fear-relevant (but not fear-irrelevant) stimuli, essentially independently of any conscious analysis of the feared object or situation. This work has generally involved using subliminal or nonconsciously presented stimuli that are exposed for very brief durations (a few hundredths of a second), followed by a backward pattern mask, which precludes participants’ ability to consciously report what they saw (e.g., Esteves and Öhman 1993; Öhman and Soares 1993). In one initial study using this technique, Öhman and Soares (1994) studied the responses to subliminally presented snakes, spiders, mushrooms, and flowers in three groups of participants who differed on whether they were afraid of snakes (but not spiders), spiders (but not snakes) or neither snakes nor spiders. Fearful participants showed automatic activation of SCRs only to their own feared stimulus (but not to their nonfeared fear-relevant stimulus or the fear-irrelevant stimuli); nonfearful participants did not show SCRs to any of the subliminal stimuli.

In addition, Öhman and Soares (1993) conditioned participants to either fear-relevant or fear-irrelevant supraliminal stimuli; later during extinction, the stimuli were presented subliminally for 30 msec, followed by a backward mask. Participants conditioned with fear-relevant CSs (snakes or spiders) continued to show reliable conditioned responses to subliminal CSs in extinction, but those conditioned with fear-irrelevant stimuli (flowers or mushrooms) showed immediate extinction of their responses to subliminal CSs. Esteves et al (1994) replicated the same basic effect using angry faces as the fear-relevant stimuli. This series of experiments (for further data, see Öhman et al 2000; Öhman and Mineka 2001) clearly supports the second feature of our hypothesized fear module: automaticity of fear activation, particularly for fear-relevant stimuli.

Encapsulation of the Fear Module from Higher Cognitive Influences

Phobic fear also appears relatively independent of beliefs in the real danger of the phobic object. Thus, most people with phobias recognize that their fears are excessive and unreasonable, but they are unable to exert rational control over them. This is in keeping with our proposal that the fear module is relatively independent of higher cognitive influences such as expectancies. Although encapsulation and automaticity both bear on the role of cognition in fear responding, automaticity refers to the way in which fear of fear-relevant stimuli can be activated directly without a conscious analysis of the stimulus. Encapsulation refers more to the relative independence and resistance of the fear response, once initiated, to conscious cognitive control.

The Independence of Cognition and the Fear Module

In support of the relative independence between cognition and the fear module, there is good evidence that fear can be conditioned to masked fear-relevant stimuli, that is, in the absence of verbalized knowledge of the contingency between the CS and the UCS. Specifically, Öhman and colleagues demonstrated that acquisition of conditioned responding can occur when subliminally presented fear-relevant CSs (but not when subliminally presented fear-irrelevant CSs) are used during acquisition (e.g., for angry faces, see Esteves et al 1994; for snakes and spiders, see Öhman and Soares 1998; Katkin et al 2001). Thus, human participants can learn to associate nonrecognized fear-relevant stimuli with aversive outcomes. Fear-irrelevant stimuli, on the other hand, are clearly associable with aversive events, but only when they are presented supraliminally.
The conclusion that fear can be conditioned to nonrecognized CSs has recently been challenged by Lovibond and Shanks (2002), reflecting the most recent round of long-standing disagreements in the literature about the role of awareness in human classical conditioning. In an attack on the possibility of nonaware learning, they argued that participants in the experiments reported by Öhman and Soares (1998) actually were aware of the CS–UCS contingency because there was evidence that they showed statistically reliable (albeit very incomplete) discrimination between masked shock-associated and the masked non-shock-associated stimuli in terms of rated shock expectancy. However, this argument rests on the dubious assumption that expectancy ratings can be taken as reflecting only conscious but not any nonconscious processes (see Merikle and Reingold 1992). Indeed, Katkin et al (2001) reported data from a masked conditioning experiment that they interpreted in terms of nonconscious feedback from autonomic responses as having affected shock expectancy ratings (see Wiens and Öhman 2002, for a critical discussion of Lovibond and Shanks 2002).

An early attempt to model the resistance of phobias to rational arguments in the laboratory was reported by Hugdahl and Öhman (1977). They first conditioned participants to fear-relevant or fear-irrelevant stimuli and then told them at the outset of extinction that shocks would no longer occur. For those conditioned with fear-relevant stimuli such instructions had no effect on the conditioned responding, which persisted just as if no instructions had been given. By contrast, those conditioned with fear-irrelevant stimuli showed immediate extinction of the conditioned response with the instructed extinction procedure. Soares and Öhman (1993) replicated and extended this effect, showing that it also occurs even using subliminal fear-relevant CSs presented in extinction.

Encapsulation and Expectancy Accounts of Phobias

The encapsulation hypothesis does run counter to proposals of other contemporary researchers who posit a central role for expectancy in eliciting conditioned SCRs (e.g., Davey 1992, 1995). Indeed research directly examining people’s expectancies about the UCS before and during conditioning reveal a rather complicated pattern of results. Human participants have exaggerated expectancies that aversive events will follow fear-relevant, relative to fear-irrelevant, stimuli at the start of a fear conditioning experiment. These reportable expectancies often roughly parallel the pattern of conditioned autonomic responses that are observed in these experiments, leading to the possibility that “UCS expectancies will play a major role in determining the autonomic CRs traditionally measured” (Davey 1992, p. 37). Such correlational results led Davey (1992, 1995) to propose an expectancy-based conditioning alternative to preparedness theory that does not rely on an evolutionary framework at all. According to his model, human participants enter experiments with enhanced expectancies that aversive events will follow fear-relevant stimuli; these expectancies will diminish unless an aversive UCS is actually threatened or presented. Moreover, Davey argues that expectancies activated by the CS are not influenced by evolutionary factors, but rather by previously acquired verbal and cultural information about expected CS–UCS relationships, as well as by contingency information presented before and during the experiment. Although expectancy biases for fear-relevant stimuli are thought to dissipate if no UCSs are presented (because of a decayed UCS representation), one UCS is thought to be sufficient to reinstate the bias. Thus, for Davey (1997, p. 305), a priori expectancy biases and “situational contingency information” play a major role in mediating SCRs.

After reviewing relevant results, Öhman and Mineka (2001) concluded that at least the strong version of this theory is untenable. First, there are convincing data demonstrating dissociations between cognitions and fear such as continued SCR responding in the absence of expectancies of aversive outcomes. For example, Schell et al (2001) conditioned participants to either fear-relevant or fear-irrelevant stimuli and measured both SCRs and trial-by-trial expectancies of the UCS. In extinction they found participants conditioned with fear-relevant (but not fear-irrelevant) stimuli showed persistent SCR responding after expectancies had extinguished. Second, Davey’s appeal to “decayed UCS representations” to account for results showing that expectancies can occur without concomitant SCRs remains untestable without an independent measure of the UCS representation (other than the occurrence of the SCR). Third, an expectancy account of selective associations is silent about why there would be qualitative differences in the nature of CRs to fear-relevant versus fear-irrelevant stimuli whereas this is not at all surprising from an evolutionary standpoint. Finally, it is a mistake to pit a cognitive expectancy model against an evolutionarily shaped mechanism as Davey does. Cognition, just like fear conditioning, has been shaped and constrained by evolution. Without acknowledging an evolutionary influence on cognition, it is difficult to explain why so many nonfearful human participants have elevated UCS expectancies to fear-relevant stimuli, as well as why nonfearful observer monkeys with no prior ontogenetic experience with snake stimuli acquire snake fear so rapidly.

Thus, we argue that a priori expectancy biases for fear-relevant stimuli may supplement, but do not replace, emotional conditioning processes mediated by the fear module. Indeed, expectancy biases may help us anticipate danger by allowing responding to verbal warnings when
no explicit cues for danger are present. In this way, they may play an important role in the maintenance of phobic fears. However, expectancies per se are not sufficient to account for the range of results reviewed here, and expectancies may more often be consequences than causes of fear responding. In conjunction with evidence reviewed in the previous section on automaticity, we consider this to be strong evidence for the relative independence of fear activation and fear conditioning from cognition.

Neural Mechanisms of Fear Generation and Fear Learning

The Amygdala Fear Circuit

Following Panksepp (1998), we would typically expect an evolved behavioral module like the one for fear to be mediated by specific neural circuitry, and one would expect it to be distinct from that mediating other types of emotions or associative learning (e.g., the cerebellum is central to skeletal conditioning such as eye-blink conditioning; for review, see Medina et al 2002). Moreover, we would expect this neural circuitry to be shared by other mammals and located in evolutionarily old parts of the brain, such as in limbic structures, rather than the more recently evolved neocortex. There is now substantial evidence suggesting that the central structure involved in the control of fear and fear conditioning is the amygdala, a collection of neural nuclei in front of the hippocampus in the anterior medial temporal lobe (e.g., Fendt and Fanselow 1999; LeDoux 2000; Medina et al 2002). As reviewed by Öhman and Mineka (2001), lesions of the amygdala have long been known to produce fearlessness. Conversely, electrical stimulation of the amygdala produces behaviors in many animals, including humans (e.g., Panksepp 1998). LeDoux and colleagues (e.g., LeDoux 1996, 2000) have also clearly demonstrated in rats that there is a subcortical thalamic route to the amygdala that plays a crucial role in fear activation and fear conditioning. Thus, subcortical inputs from the thalamus to the amygdala allow sensory signals to activate it either before or simultaneously with the arrival of signals from the cortex; such subcortical inputs may potentially explain the role of unconscious processes in the activation and conditioning of fear.

Recently, evidence demonstrating the role of the amygdala in human fear conditioning has also accumulated. For example, Morris et al (1998) used positron emission tomography in normal human participants to test the hypothesis that fear conditioning to angry faces is centered on the amygdala. The most important finding of this experiment was that the human amygdala was indeed specifically activated by the conditioned fear stimuli, both when they were presented supraliminally and when they were presented subliminally. In particular, the right amygdala was activated even by fear-relevant CSs that were presented outside conscious awareness by means of backward masking. In addition, LaBar et al (1995) examined patients treated by unilateral removal of the amygdala to control epileptic seizures. They found impaired conditioning of the SCR in these patients (evidence of a conditioned emotional response), although they showed normal unconditioned responses to the shocks used as UCSs, and they could report the CS–UCS contingency. Thus, conscious awareness of the CS–UCS contingency is not sufficient for SCR (emotional) conditioning without an intact amygdala. Furthermore, Morris et al (1999, 2001) demonstrated that there is a subcortical pathway to the amygdala for fear-relevant visual stimuli.

In a related line of evidence, Bechara et al (1995) reported a double dissociation between the role of the amygdala in conditioning of the emotional response (SCR) and the role of the hippocampus in the development of CS–UCS contingency learning. Specifically, one patient with bilateral amygdala lesions did not acquire conditioned SCRs despite normal UCRs and normal acquisition of factual knowledge about the conditioning contingency. By contrast, another patient with bilateral hippocampal damage failed to learn the contingency at the cognitive level but did acquire differential conditioned SCRs. Thus the amygdala circuit seems specifically dedicated to fear learning, whereas the hippocampus is important for conscious cognitive learning.

Two Levels of Learning

Öhman and Mineka (2001) further proposed that there seem to be two levels of learning in human fear conditioning. This proposal builds on Razran’s (Razran 1971) theory about the evolution of different levels of learning through ontogeny and phylogeny, starting with basic nonassociative learning (such as habituation and sensitization), to basic associative learning (such as classical and instrumental conditioning), to higher levels of cognitive learning such as contingency and symbolic learning. We proposed that fear learning in humans occurs both at a basic associative level as evidenced by emotional responses (such as autonomic responses), mediated by the amygdala, and at a cognitive level of contingency learning, mediated by the hippocampus in at least its simple forms. When fear learning in human conditioning experiments involves fear-relevant stimuli that activate the fear module, the learning generally occurs at both levels simultaneously but independently. By contrast, when fear learning in human conditioning experiments involves fear-irrelevant stimuli, the learning would typically only engage at the cognitive level, without significant emotionality. However, fear-
Phobias and Preparedness

relevant CSs should not be understood as necessary for engaging the basic fear conditioning system. Other factors likely to have a similar effect include UCS intensity (Bridger and Mandel 1964), CS–UCS interval (Schell et al 1991), a delay rather than a trace conditioning paradigm (Manns et al 2002), and the controllability of the UCS (e.g., Mineka et al 1984). Thus, more direct and real traumatic conditioning events such as can occur outside the laboratory in the everyday lives of humans would clearly engage both emotional and cognitive levels of conditioning.

Summary and Conclusions

Our understanding of evolutionary constraints on the fear learning involved in everyday fears and phobias has dramatically expanded in the past decades. We have proposed that an evolved behavioral module—the fear module—can best explain the complex pattern of results that have emerged in the past 30 years. The primary characteristics of the fear module include selectivity with regard to input, automaticity with regard to activation, encapsulation from higher cognitive influences in learning and in the cognitive control of fear, and a dedicated neural circuitry. First, with regard to selectivity of input, we reviewed evidence that the phenomenon of selective associations in fear conditioning is robust and replicable across laboratories, having been shown with three kinds of paradigms and multiple dependent measures (e.g., autonomic responses, contingency estimates, acquired aversiveness ratings, fear and avoidance behaviors). Moreover, the nature of the CRs seen with fear-relevant CSs is often qualitatively different from those seen with fear-irrelevant CSs, as would expected if genuine defensive responses were being conditioned only with the fear-relevant stimuli. Second, with regard to automaticity, feared fear-relevant CSs (but not fear-irrelevant CSs) can activate fear responses even to subliminally presented fear-relevant CSs. Third, with regard to encapsulation, fear conditioning to fear-relevant stimuli can occur even with subliminal CSs, and such fear learning seems relatively impenetrable to conscious cognitive control. Fourth, the amygdala seems to be the central site for operations of the fear module and characteristics of fear learning in humans postulated by our proposed fear module are consistent with known operations in the amygdala regarding fear learning. Finally, we have proposed two levels of fear learning, with the emotional level being mediated by the amygdala and the cognitive contingency level being mediated by the hippocampus.

The fear module concept helps to integrate many diverse findings on fear from diverse domains such animal and human fear conditioning (including its neurobiological basis), human covariation judgments, and clinical findings on fears and phobias. Although the evidentiary basis for each of the four characteristics of the fear module may not yet be complete, it is the converging evidence from multiple lines of work that gives us confidence in the soundness of the approach. Moreover, the fear module concept clearly helps set an agenda for future research on its various characteristics that may also improve our understanding of both basic and clinical aspects of fear and anxiety in humans.

For the first author, some of the research cited was supported by grants from the National Science Foundation (Grant Nos. BNS-8507340 and BNS-8216141). For the second author, the research cited as well as the preparation of this article was made possible by grants from the Swedish Council for Research in the Humanities and Social Sciences and from the Bank of Sweden Tercentennial Foundation.

Aspects of this work were presented at the conference, “Learning and Unlearning Fears: Preparedness, Neural Pathways, and Patients,” held March 21, 2002 in Austin, TX. The conference was supported by an unrestricted educational grant to the Anxiety Disorders Association of America (ADAA) from Wyeth Pharmaceuticals, and jointly sponsored by the ADAA, the ADAA Scientific Advisory Board, and the National Institute of Mental Health.

References


