Learned Fear of “Unseen” Faces After Pavlovian, Observational, and Instructed Fear

Andreas Olsson and Elizabeth A. Phelps

New York University

ABSTRACT—This study compared fear learning acquired through direct experience (Pavlovian conditioning) and fear learning acquired without direct experience via either observation or verbal instruction. We examined whether these three types of learning yielded differential responses to conditioned stimuli (CS+) that were presented unmasked (available to explicit awareness) or masked (not available to explicit awareness). In the Pavlovian group, the CS+ was paired with a mild shock, whereas the observational-learning group learned through observing the emotional expression of a confederate receiving shocks paired with the CS+. The instructed-learning group was told that the CS+ predicted a shock. The three groups demonstrated similar levels of learning as measured by the skin conductance response to unmasked stimuli. As in previous studies, participants also displayed a significant learning response to masked stimuli following Pavlovian conditioning. However, whereas the observational-learning group also showed this effect, the instructed-learning group did not.

Learning is an adaptation that enables organisms to change their behavior flexibly in a fluctuating environment. An important component of learning is the emotional reactions that guide and facilitate action when the organism encounters objects and events that should be either avoided or approached, depending on their potential impact on the organism’s survival (Rolls, 1999). Although recent investigations have examined the informative value of such emotional responses in humans during cognitive appraisal (Katkin, Wiens, & Öhman, 2001), decision making (Damasio, 1999), memory performance (Cahill & McGaugh, 1998), and action selection (Bechara, Damasio, Tranel, & Damasio, 1997), the potentially moderating role of the way events and objects acquire their values has been neglected (for recent exceptions, see Hertwig, Barron, Weber, & Erev, 2004; Phelps et al., 2001). When the method of emotional learning has been examined, it has often involved Pavlovian conditioning (e.g., LaBar, LeDoux, Spencer, & Phelps, 1995). This is noteworthy because humans may acquire the bulk of their knowledge of the emotional significance of objects and events in their surroundings through social observation and symbolic communication (Rachman, 1977).

Although no attempts have been made to systematically compare aversive learning through first-hand experiences (Pavlovian conditioning) with learning of the same causal contingencies through solely social observation (observational learning) or verbal instruction (instructed learning), available data suggest that emotional responses acquired through different kinds of learning should exhibit both similarities and differences. In addition, whereas some researchers claim that the same underlying mechanism subserves different types of learning (e.g., Lovibond & Shanks, 2002), others argue that different types of learning might draw on partially independent systems (Öhman & Mineka, 2001). In this study, we aimed to shed further light on these issues by, for the first time, systematically comparing the impact of Pavlovian, observational, and instructed learning in a fear-learning paradigm.

In the traditional Pavlovian fear-conditioning paradigm, a conditioned stimulus (CS+—e.g., a face) that has been paired with a naturally aversive unconditioned stimulus (UCS—e.g., an electric shock) elicits a greater conditioned response (CR—e.g., autonomic arousal) than a control stimulus (CS−), which has not been paired with the UCS. In this article, we use this terminology to describe stimuli and responses in all types of learning.

In an observational fear-learning protocol, a fear response is acquired without direct experience of the UCS. Instead, the representation of another individual’s emotional expression can
act as the UCS. In an early study on observational fear learning in humans (Hygge & Öhman, 1978), subjects were exposed to a confederate’s fear reactions to either fear-relevant stimuli (e.g., snakes) or fear-irrelevant stimuli (e.g., flowers). The results showed that subjects readily acquired a fear response to the stimuli paired with a fear expression in the confederate, and this response was stronger for fear-relevant stimuli. A related set of findings was reported by Mineka and her colleagues in a series of studies on vicarious fear conditioning in monkeys (Mineka & Cook, 1993; Mineka, Davidson, Cook, & Keir, 1984). In summary, these studies demonstrated rapid, strong, and persistent learning following exposure to a conspecific’s fearful reactions to a fear-relevant stimulus (i.e., a snake). A recent study, using a learning protocol similar to the one used by Mineka et al., found comparably strong and persistent fear learning in toddlers who observed their mothers’ fearful expressions in response to fear-relevant objects (Gerull & Rapee, 2002).

Another way of acquiring knowledge of the aversive qualities of a stimulus in the absence of direct experience is through language. Both clinical accounts that retrospectively identify the etiology of phobic fears to fear-relevant stimuli (e.g., King, Gullone, & Ollendick, 1998) and experimental studies involving stimuli ascribed fear-provoking qualities through storytelling (Field, Argyris, & Knowles, 2001) reveal that verbal instructions are a potent means to fear learning. A number of studies focusing on the physiological components of instructed fear learning have shown that when participants are verbally instructed to expect a shock paired with the presentation of a specific CS and then later exposed to fully visible CSs, they display an arousal response that is similar to the one demonstrated following Pavlovian fear conditioning (Grillon, Ameli, Merikangas, Woods, & Davis, 1991; Hugdahl & Öhman, 1977; Phelps et al., 2001).

AWARENESS AND EMOTIONAL LEARNING

It is unclear given current evidence whether or not the expressions of fears acquired by different kinds of learning are supported by the same underlying system (Öhman & Mineka, 2001). One proposed way to dissociate learning systems, or components of learning systems, is to establish that they produce qualitatively different outcomes if affected by a given variable, such as awareness (Merikle & Reingold, 1992). Awareness of a visual CS can be manipulated through backward masking, in which the CS (target) is presented briefly and immediately followed by another stimulus (mask) that overlaps with it spatially. The presentation of the mask interrupts the processing of early visual information related to the target and thus excludes awareness if successful (Marcel, 1983). This technique is said to short-circuit explicit knowledge of CS-UCS contingencies and to tap into information processing that is partially independent of explicit awareness (Öhman, Flykt, & Lundquist, 2000).

Although some investigators claim that emotional learning is dependent on explicit awareness of stimulus contingencies (Dawson, 1973; Lovibond, 2003; Lovibond & Shanks, 2002), a significant body of evidence indicates that these two processes are independent under certain circumstances (Bechara et al., 1995; Esteves, Dimberg, & Öhman, 1994; Mandel & Bridger, 1973; Morris, Öhman, & Dolan, 1998; Posner & Snyder, 1975). In particular, the effects of Pavlovian fear conditioning may be mediated by both explicit and implicit representations of the CS-UCS contingency; this possibility is supported, for example, by the results of studies in which subjects verbally reported both expectancies and autonomic, emotional responses (Öhman & Mineka, 2001).

In an experiment utilizing the masking paradigm subsequent to Pavlovian fear conditioning, Esteves et al. (1994) paired angry faces with electric shocks; happy faces served as the CS+. In the test phase, both unmasked and masked stimuli were presented while skin conductance response (SCR) was assessed. Subjects showed a greater response to the CS+ than to the CS− in both the unmasked and the masked conditions, results that are consistent with the findings of a number of similar experiments using fear-relevant stimuli as conditioned stimuli (e.g., Morris et al., 1998; Öhman & Soares, 1993). The evidence that emotional responses are partially independent of explicit awareness resonates well with recent findings in neuroscience indicating that the human brain comprises neural circuits that support automatic processing of emotionally relevant information (LeDoux, 1996) and that these circuits can initiate emotional responses without explicit awareness of the stimuli (e.g., Bechara et al., 1995; Critchley, Mathias, & Dolan, 2002; Morris et al., 1998; Whalen et al., 1998).

But are these emotional responses to fear-relevant stimuli dependent on the way the emotional significance is acquired? To date, no studies have investigated the role of awareness of the CS following observational and instructed learning. We aimed to investigate the degree to which a learned emotional response is modulated by (a) type of learning (Pavlovian, observational, or instructed) and (b) awareness of the CS (unmasked or masked). If fear learning through observation and fear learning through verbal instruction engage representations in the same system as Pavlovian conditioning, then all three kinds of learning may produce similar emotional responses to the CS.

In order to isolate the effects of the type of learning, we varied the learning component while keeping other factors constant. A systematic comparison of the emotional responses acquired through different types of learning can be informative about what components are necessary and sufficient in order for autonomic, emotional responses to be elicited, and what the underlying processes are. In a broader perspective, such a comparison can also help to clarify the role of perceptual and symbolic representations acquired without direct aversive experience in informing people about the emotional significance of certain situations, and is bound to have a considerable
Learned Fear of “Unseen” Faces

Impact in other fields, such as decision making, memory, and psychopathology.

In the present experiment, we examined responses of a Pavlovian-learning group in order to replicate earlier findings (e.g., Esteves et al., 1994) demonstrating that Pavlovian conditioning can produce a significant emotional response even when fear-relevant CSs are presented without the subject’s explicit awareness. Two other groups of subjects were submitted to similar test conditions following observational and verbally instructed learning, respectively.

METHOD

Participants

One hundred fifty-nine college students served in the experiment. All participants gave informed consent and were paid for their participation. Participants were randomly assigned to the Pavlovian-, observational-, or instructed-learning group. Subsequent to the experiment, explicit awareness of the masked stimuli was assessed, and 25 subjects were excluded from further analysis because they claimed to have seen the masked CSs. In the data analysis, 14 additional subjects were excluded because they displayed virtually no SCR (nonresponders). Also, 33 subjects were excluded because they showed no signs of learning in the unmasked condition (Pavlovian learning, n = 11; observational learning, n = 8; instructed learning, n = 14). The final sample consisted of 87 subjects—29 in each learning group.

Apparatus and Material

The experiment took place in a sound-attenuated room. Subjects were seated in a comfortable chair in front of a 20-in. CRT Apple monitor that projected the stimuli synchronized with a 60-Hz vertical refresh rate. The images were taken from Ekman and Friesen (1976) and consisted of three black-and-white pictures of males. Two angry faces served as CSs, and a neutral face served as a mask. Angry male faces were chosen because only conditioning to fear-relevant stimuli has been reported to survive masking (Öhman et al., 2000).

Initially, all participants were told that the purpose of the study was to measure physiological responses to pictures of human faces and that electric shocks were going to be administered during the experiment. After participants were seated, the shock and the SCR electrodes were attached.

In the Pavlovian-learning group, the amplitude of the shock was determined individually by a work-up procedure, which terminated when the shock was reported to be “uncomfortable, but not painful.” Participants were given no information about the stimuli contingency before or during the experiment. To determine the baseline SCR to both unmasked and masked CSs, we included two unreinforced presentations of each trial type (habituation phase). In the subsequent acquisition phase (six trials of each type), participants received six shocks that coterminated with the presentations of the unmasked CS+ (i.e., delayed conditioning). No shocks were given to the unmasked CS− and the masked CS+ or CS−. Our masking procedure was modeled on a well-established paradigm in which an unmasked CS+ is paired with a shock and then the CR to both unmasked and masked stimuli is assessed (e.g., Esteves et al., 1994; Morris et al., 1998; Öhman & Soares, 1993). Two unmasked CS+ trials always preceded the first masked trial, to ensure that learning occurred before the presentation of the first masked trial. During the extinction phase (five trials of each type), no more shocks were administered.

In the observational-learning group, no initial calibration procedure was performed. Participants were told that they would first watch a movie of another person participating in an experiment that was identical to the one in which they themselves would subsequently participate (except for being shorter). They were told that the experiment contained three phases and that at least one and at most three shocks would be administered during the second phase. They were also informed...
that shocks would be paired with the same stimulus during the experiment as in the movie, but that the order of the shocks within the second phase would be randomized. After the movie ended (5 min and 12 s), participants were briefly reminded of the instructions and then told that the experiment would begin. The stimuli were then presented using a procedure identical to the procedure for the Pavlovian group except that no shocks were administered.

The instructed-learning group, like the observational group, did not calibrate the shock level. Participants in this group were initially informed that they were not going to receive any shocks during the first phase of the experiment. After the end of the habituation phase, the experiment was briefly interrupted, and participants were shown the CS+ and told by the experimenter that they would receive at least one and at most three shocks paired with this face. They were then shown the CS− and the mask while being assured that they should not expect any shocks paired with these faces, nor in between trials. The test protocol that followed was identical to the acquisition phase in the Pavlovian protocol except that no shocks were administered. After the last acquisition trial, the experiment was interrupted, and subjects were assured that no shocks would be administered throughout the remaining trials (extinction phase).

When asked at the end of the experiment, all participants in the observational and instructed groups reported that they believed the instructions and thus expected at least one shock. Subsequently, in order to assess explicit awareness of the masked stimuli, we asked participants whether they noticed anything peculiar with the display of the neutral face. They were then asked whether they saw another image preceding the neutral face. Participants who reported having seen the masked faces were excluded from analysis. Finally, participants were debriefed and paid for their participation.

Scoring of Responses
SCR was measured for each trial as the peak-to-peak amplitude difference in skin conductance to the first response (in microsiemens) in the 0.5- to 4.5-s latency window following stimulus onset. The minimal response criterion was 0.02 μS. The raw SCR scores were square-root transformed to normalize the distributions.

RESULTS
All trials were used to produce four average scores per subject (CS+ and CS− for both unmasked and masked conditions), except that in the unmasked Pavlovian condition, only Trials 2 through 6 were used because the CS+ was not predictive of the UCS until after its first association with the shock. Separate analyses of variance were computed for the habituation, acquisition, and extinction phases. Data were analyzed separately for unmasked and masked trials. Learning group served as the between-subjects variable, whereas stimulus type (CS+ vs. CS−) was the within-subjects repeated measure.

Habituation
In the habituation phase, no significant differences were found.

Acquisition
Mean responses during the acquisition phase are presented in Figure 1. SCRs to CS+ trials were significantly larger than SCRs to CS− trials for both unmasked trials, $F(1, 84) = 120.18, p < .001, \eta_p^2 = .59,$ and masked trials, $F(1, 84) = 7.52, p < .01, \eta_p^2 = .08.$ Because only subjects who showed a positive differential response between the unmasked CS+ and CS− (indicating that learning was present) were selected for the statistical analysis in the first place, subsequent analyses of the acquisition data focused on the masked condition.

There was a marginally significant difference between masked CS+ trials and masked CS− trials in the Pavlovian group, $t(28) = 1.92, p = .06$ (two-tailed). An earlier study using an experimental paradigm similar to ours found a pronounced attenuation of the differential SCR over masked test trials (Esteves et al., 1994). Esteves et al. argued that joint presentation of the unreinforced CS+ with the mask may cause the mask to become an inhibitory stimulus, gradually inhibiting the response to the CS+. This reasoning was corroborated by the present results: The difference between masked CS+ and CS− trials was significant when the first five trials, rather than all six, were used to produce the average (the same number as in the unmasked Pavlovian condition), $t(28) = 2.27, p < .05,$ Cohen’s $d = 0.41.$ This result, which was predicted, indicates that explicit awareness was not necessary to elicit a differential response following Pavlovian conditioning.

The observational group also displayed a significant differential response to masked faces (CS+ vs. CS−), $t(28) = 2.22, p < .05,$ Cohen’s $d = 0.36.$ For the instructed group, however, SCRs did not differ between CS+ and CS− trials, $t(28) = 0.66, p = .52.$

Extinction
The differential responding to CS+ versus CS− in unmasked trials resisted extinction in all three learning groups—Pavlovian, $t(28) = 3.66, p = .001,$ Cohen’s $d = 0.62;$ observational, $t(28) = 4.21, p < .001,$ Cohen’s $d = 0.83;$ instructed, $t(28) = 2.09, p < .05,$ Cohen’s $d = 0.32.$ However, there was a significant interaction between stimulus type (CS+ vs. CS−) and learning group, $F(2, 84) = 4.07, p < .05, \eta_p^2 = .09.$ Post hoc comparisons (with Bonferroni correction) revealed less extinction in the observational than in the instructed group, $p < .05,$ and marginally less extinction in the Pavlovian than in the instructed group, $p = .08.$ There were no significant effects in the masked conditions. Recall that just before the start of the extinction phase, the instructed group received explicit
instructions that no more shocks would be administered, whereas the observational and Pavlovian groups had to gradually discover this on their own. This difference may have contributed to the observed group differences in resistance to extinction. It is interesting to note that the differential response to CS\textsubscript{1} versus CS\textsubscript{0}/C\textsubscript{0} remained significant in the instructed group despite their being explicitly told that the shocks would terminate.

**DISCUSSION**

Pavlovian, observational, and instructed fear learning have been examined in a variety of research traditions, using a range of different paradigms. Although several similarities among these types of learning have emerged (e.g., Mineka et al., 1984; Phelps et al., 2001), important differences are also apparent (Öhman & Mineka, 2001). However, until now, no systematic comparison of the differential impact of different types of learning on learned autonomic responses has been conducted. Our aim was to provide such a comparison by submitting the three kinds of fear learning to similar test conditions.

In the acquisition phase, the differentially greater SCR for masked CS+ versus CS− found in the Pavlovian group replicated earlier findings (e.g., Esteves et al., 1994) and gave us confidence in the effectiveness of the experimental methods we were using. The observational-learning group also showed this effect, a finding that is consistent with studies reporting behavioral and psychophysiological manifestations of fast, strong, and persistent learning in observational-learning paradigms in both human (e.g., Hygge & Öhman, 1978) and nonhuman primates (e.g., Mineka & Cook, 1993). Recent findings of overlapping activations of neural networks that support one’s own emotional expressions and the perception of emotions in others (Adolphs, 2002; Carr, Iacoboni, Dubeau, Mazziotta, & Lenzi, 2003) provide one possible explanation for the strong manifestations of fear learning following observation. Thus, in shaping an aversive response to a stimulus, the perception of fear reactions in another individual may serve as a UCS that is as powerful as the corresponding first-person experience.

The absence of the same data pattern in the instructed group shows that the expression of fear learned through verbal instructions does not survive masked presentations. Although abstract representations of associations between specific objects and their aversive implications are enough to prepare the individual for action and guide his or her behavior, the individual will express an emotional response only if these representations are accompanied by explicit awareness of the target. Recent findings on the lateralization of amygdala activity may shed some light on the reasons for the absence of a differential SCR to masked stimuli following instructed learning. Two recent studies demonstrated that stimuli that are verbally linked to an aversive outcome activate mainly the left amygdala (Funayama, Grillon, David, & Phelps, 2001; Phelps et al., 2001), whereas other studies have suggested that masked presentations of aversively conditioned stimuli engage
predominantly the right hemisphere (Morris et al., 1998; Peper & Karcher, 2001).

The current study employed a subjective measure of explicit awareness of the masked conditioned stimuli. A subjective response is problematic for several reasons (for an overview, see Holender, 1986), most notably because it leaves the definitional burden to the observer (Merikle & Reingold, 1992). However, the practice of using a more conservative (objective) threshold, such as chance performance on a forced-choice discrimination task, is based on the false presumption that there is one exhaustive measure of all aspects of consciousness (Merikle & Joordens, 1997). Rather than making any claims about the nature of conscious awareness, and in accordance with earlier studies (e.g., Esteves & Öhman, 1993; Whalen et al., 1998), we used verbal report by the subject as the indicator of conscious awareness. Because our specific interest was in the differential responses between learning groups, the potential problems raised by a subjective measure were not critical to our research hypotheses.

We have shown, for the first time, that fear learning following observation, like Pavlovian conditioning, need not be accompanied by explicit awareness of the CS for an emotional response to be expressed. In contrast, knowledge acquired through linguistic input does require explicit awareness of the CS to produce an emotional response. In other words, although some preferences need no inferences, others do. These results also lend support to the notion that there might be partially dissociable systems involved in different modes of emotional learning. Pavlovian and observational learning, which humans share with other primates, might be supported by an evolutionarily old system that predates the emergence of language.

In conclusion, our results suggest that the emotional arousal response to conditioned stimuli is mediated by both (a) how learning is acquired and (b) the form in which the conditioned stimuli are presented. Further studies are needed in order for investigators to better understand the neural mechanisms underlying these phenomena, as well as the wider implications for social learning and behavior.

Acknowledgments—We thank Kristen Stedenfeld for her help in data collection, Laura Thomas and Brett Sedgewick for their technical assistance, and Arne Öhman and two anonymous reviewers for their helpful comments on an earlier version of this article. This study was supported by National Institutes of Health Grant MH62104 to E.A.P.

REFERENCES


(RECEIVED 9/17/03; REVISION ACCEPTED 3/1/04)