Learning fears by observing others: the neural systems of social fear transmission

Andreas Olsson,[1] Katherine I. Nearing,[2] and Elizabeth A. Phelps[3,4]

[1]Department of Psychology, Columbia University, [2]Epilepsy Center, NYU School of Medicine, New York University, [3]Department of Psychology, and [4]Center for Neural Science, New York University, NY, USA

Classical fear conditioning has been used as a model paradigm to explain fear learning across species. In this paradigm, the amygdala is known to play a critical role. However, classical fear conditioning requires first-hand experience with an aversive event, which may not be how most fears are acquired in humans. It remains to be determined whether the conditioning model can be extended to indirect forms of learning more common in humans. Here we show that fear acquired indirectly through social observation, with no personal experience of the aversive event, engages similar neural mechanisms as fear conditioning. The amygdala was recruited both when subjects observed someone else being submitted to an aversive event, knowing that the same treatment awaited themselves, and when subjects were subsequently placed in an analogous situation. These findings confirm the central role of the amygdala in the acquisition and expression of observational fear learning, and validate the extension of cross-species models of fear conditioning to learning in a human sociocultural context. Our findings also provides new insights into the relationship between learning from, and empathizing with, fearful others. This study suggests that indirectly attained fears may be as powerful as fears originating from direct experiences.

Keywords: amygdala; social learning; fear conditioning; empathy; fMRI

Learning to respond appropriately to environmental stimuli that predict potentially harmful events is an adaptive mechanism crucial to the survival of any organism. Studies exploring the neural circuitry underlying this adaptive mechanism have primarily examined classical fear conditioning paradigms. In classical fear conditioning, the organism acquires a conditioned fear response (CR) to a previously neutral stimulus (CS) through the direct experience with its pairing with a naturally aversive event, the unconditioned stimulus (US). Extensive research in non-human mammals has shown that the amygdala is a critical structure involved in the acquisition, storage and expression of conditioned fear (Kapp *et al.*, 1992; Fanselow and LeDoux, 1999; Davis and Whalen, 2001), and more recent work demonstrates that the amygdala plays a similar role in human fear conditioning (LaBar *et al.*, 1995; LaBar *et al.*, 1998; Phelps and LeDoux, 2005).

Although this paradigm provides a window into the neural workings of fear-learning in humans, fear conditioning may not be the primary means that humans learn about the potentially harmful qualities of stimuli in the environment. Our sociocultural environment provides other, indirect, means of attaining fear-relevant information, such as social observation and verbal communication, which are more efficient and associated with fewer risks than learning through direct aversive experiences (Miller and Dollard, 1941; Bandura, 1977; Rachman, 1977). However, it is unknown whether the associative learning mechanisms and their related neural processes, known to underlie classical fear conditioning, can be extended to these means of emotional learning that may be more commonly represented in everyday life. In other words, it is unclear if these different learning procedures are supported by the same or different underlying neural processes.

In the present experiment, we used functional magnetic resonance imaging (fMRI) to investigate whether the neural components known to support classical fear conditioning are similarly engaged in the acquisition and expression of fears acquired without personal experience of the US, via social observation. Previous behavioral (Mineka *et al.*, 1984; Gerull and Rapee, 2002), psychophysiological (Berger, 1962; Hygge and Öhman, 1978; Vaughan and Lanzetta, 1980; Olsson and Phelps, 2004) and imaging (Phelps *et al.*, 2001) research has highlighted both similarities and differences between directly and indirectly acquired fears, but no study has investigated the neural mechanisms supporting fears transmitted through social observation.

Although the neural correlates of observational fear-learning are unidentified, behavioral work shows that, across species, observing expressions of distress in a conspecific can provide powerful means of attaining fear relevant information about co-occurring stimuli and events.

Observational fear-learning has been documented in birds (Curio, 1988), rodents (Kavaliers *et al.*, 2001), cats (John *et al.*, 1968) and primates (Cook and Mineka, 1990; Mineka and Cook, 1993). Relating more specifically to facially expressed distress, studies in monkeys (Mineka *et al.*, 1984; Cook *et al.*, 1985; Cook and Mineka, 1990; Mineka and Cook, 1993) and humans (Berger, 1962; Vaughan and Lanzetta, 1980; Gerull and Rapee, 2002; Olsson and Phelps, 2004) have established that facial fear expressions can effectively serve as an US. In particular, work on observational fear-learning in monkeys has shown that the relationships between the magnitude of a learning model's expressed distress, the observer's immediate response to the model's distress and the resulting fear-learning in the observer are similar to those existing between an US, UR and a CR in classical fear conditioning paradigms (Cook and Mineka, 1990; Mineka and Cook, 1993). A recent study directly comparing human fear-learning through conditioning, social observation and verbal instruction supports the same conclusion (Olsson and Phelps, 2004). In this study, fear responses acquired through conditioning and observation of a distressed model were subsequently expressed to both seen and unseen (backwardly masked) presentations of CSs. In contrast, fear acquired through verbal instruction was only expressed to seen CSs, suggesting that learning attained through conditioning and observational learning procedures were represented differently from verbally transmitted fears. Taken together, these similarities between conditioned and observationally acquired fear responses suggest that the same, or similar, associative learning mechanisms are involved. Thus, conforming to a vast literature on classical fear conditioning (LaBar *et al.*, 1995; Phelps, 2006), we hypothesized that observational fear-learning would recruit a neural circuitry including the amygdala bilaterally during both acquisition and expression. The involvement of the amygdala would lend support to the extension of the conditioning model to fear learning through social observation.

An alternative hypothesis is that the indirect learning procedure of observational fear-learning is supported by a qualitatively different mechanism than fear conditioning. For example, it could be caused by inductive reasoning linking a cognitive representation of fear with the CS+, or a second-order associative process, in which the model's distress constitutes another CS by virtue of an earlier acquired association with a direct experience of an aversive event (US). These alternative descriptions of observational fear-learning would be unlikely to involve the amygdala bilaterally, especially during the acquisition stage.

METHODS

Participants

A total of 14 right-handed male subjects (mean age, 26.2 years; range, 22–36 years) were paid for their participation. Two subjects were excluded from the subsequent data analysis because they voiced suspicions about whether shocks were actually being administered during the stage of the study. One subject was excluded due to technical problems. Eleven subjects remained for the final analysis. The protocol was approved by the local Institutional Review Board (IRB) at New York University.

Stimuli

A movie was created for the observation stage of the experiment. The movie (3 min, 54 s) displayed a male participant (the learning model) taking part in a differential fear conditioning experiment (Figure 1A). Two colored squares (blue and yellow) served as CS and were presented on a computer screen in-front of the learning model. Each CS was presented for 10 s in a pseudorandomized order and interleaved with an inter-stimulus-interval (ITI) varying between 10 and 14 s. During the ITI, the word 'rest' was
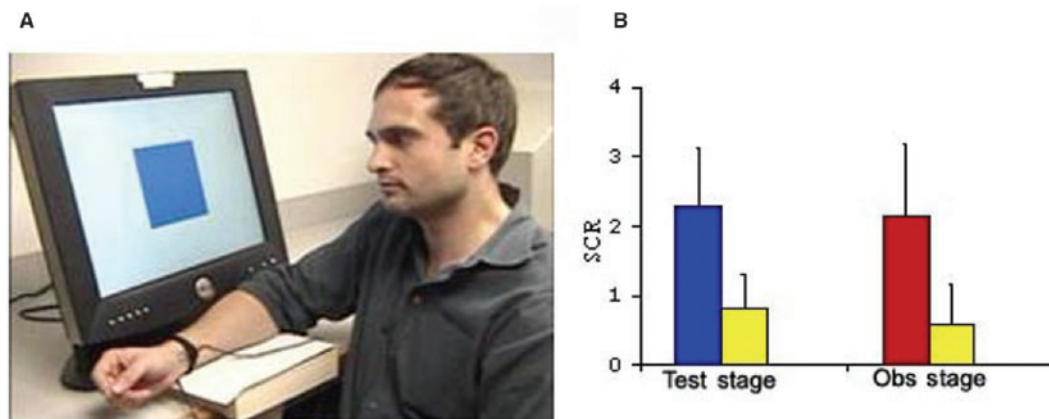
Fig. 1 (**A**) A snap-shot from the movie presented to subjects during the observation stage, depicting the learning model facing a computer screen that displayed the CS+ and CS−. To each of the three shocks paired with a CS+, the model displayed signs of distress (e.g. lowering the eyebrows, raising the cheeks and twisting the right hand indicating the administration of the shock). (**B**) Mean skin conductance response (SCR) during the test stage to the CS+ (blue bar) and the CS− (yellow bar) and during the observation stage to the model's response to the shock (red bar) and CS− (yellow bar). Error bars show standard error.

displayed on the computer screen. Each colored square was presented five times, starting with the rest period. Three presentations of the color serving as the CS+ co-terminated with the administration of an uncomfortable shock to the right wrist of the model, whereas the other (CS−) was never paired with a shock. The shock level was adjusted by the learning model to be uncomfortable but not painful prior to recording the movie. In order to counterbalance which color served as the CS+ and CS−, respectively, the original movie was video edited (Adobe Premiere). This created a second version of the movie, in which the entire visual input, apart from the colors serving as the CS+ and CS−, remained identical to the original movie.

Subjects' skin conductance response (SCR) was used to assess learning. The SCR was measured through Ag–AgCl electrodes attached to the distal phalanges of the second and third digits of the left hand. The electrode cables were grounded through an RF filter panel. The SCR signal was amplified and recorded with a BIOPAC Systems (Santa Barbara, California) skin conductance module connected to a Macintosh computer. Data were continuously recorded at a rate of 200 samples per second. An off-line analysis of the analogue SCR waveforms was conducted with AcqKnowledge software (BIOPAC Systems Inc., Goleta, California).

Procedure

Our experiment consisted of two parts; an observation and a test stage, each one corresponding to a functional scan. Before entering the scanner, subjects were told the following:

> You will now watch a movie of a person doing an experiment similar to the one you yourself are going to do afterwards. The person in the movie is going to receive shocks paired with one out of two colored squares presented to him. Please, pay attention to the movie because in the experiment that you are going to do afterwards, you are going to receive shocks to the same color as the person in the movie.

Previous research has shown that watching a movie of a learning model reacting to a CS can be as effective in transmitting fear as watching the real event (Cook and Mineka, 1990). After receiving the instructions, the subject entered the scanner. During the first functional scan (the observation stage), each subject was presented with the movie of a learning model participating in a classical conditioning experiment (see section 'Stimuli' and Figure 1A). The movie was followed by a short break of ∼30 s. The subjects were then told the following:

> You are now going to take part in an experiment similar to the one you just watched. You will be presented with the same number of colored squares as the person in the movie, but in a different order. Importantly, you will receive shocks paired with the same color as the person in the movie. However,

whereas the person in the movie received 3 shocks, you will receive between 1 and 3 shocks – at least one and at most three – paired with the same color as in the movie. No shocks will be delivered to the other color or during the rest stages.

These instructions were followed by the second functional scan (the test stage), which consisted of the presentation of the same number of colored squares, but in a different order, as was shown to the model in the video (five yellow and five blue squares). The identical stimulus material and parameters were used in the movie and test stages. Importantly, however, no shocks were administered to the subject during the test stage to ensure that learning was attained through indirect, social means only. At the end of the experiment, subjects were debriefed and asked whether they had believed the instructions.

SCR parameters and data analysis

SCR was measured for each trial as the base-to-peak amplitude difference in skin conductance to the largest response (in microSiemens, μS) in the 0.5–4.5 s latency window following stimulus onset. The minimal response criterion was 0.02 μS. Responses that did not pass this criterion were scored as '0'. The SCR data were low pass filtered and smoothed and square-root transformed to normalize the distributions. All trials were used to produce two average scores (CS+ and CS−) per subject for the test stage. For the observation stage, three averages (CS+, CS− and US) were produced that excluded the first CS+ and CS−, because the CS+ was not predictive of the US until after its first association with the shock to the learning model. Data were analyzed separately for the test stage and the observation stage.

MRI parameters and data analysis

The study was performed at the NYU Center for Brain Imaging using a 3T Siemens Allegra scanner and a Siemens head coil. Initially, MPRage anatomical scans were acquired to obtain a 3D volume for slice selection. Forty 3 mm thick slices were prescribed parallel to the AC–PC line to obtain whole brain coverage in the same plane as the functional data. Two functional scans were performed, corresponding to the observation and test stages, respectively). All functional scans lasted for 232 s, using a gradient echo sequence, TE = 30 ms, TR = 2000 ms, flip angle = 90°, FoV = 192. The inplane resolution was 3 mm × 3 mm. In-between functional runs there was a break of ∼30 s.

Imaging data were analyzed with Brain Voyager (2000, 4.9). The data were temporally and spatially smoothed (4 mm FWHM), motion corrected and transformed into Talairach space. The functional data was overlaid on one of the subject's structural scan. An overall group analysis was performed for each stage in both conditions. The data belonging to each trial type were convolved with the canonial hemodynamic response function using a general

linear model, resulting in a total of five regressors for each of the two conditions. In the observational condition, the observation stage included the regressors CS+, CS− and shock. The shock regressor comprised the 2000 ms following the administration of the shock to the model in the movie (i.e. 2 s immediately after the three CS+ trials that were paired with a shock). The test stage that followed the observational learning stage used CS+ and CS− as regressors.

RESULTS

Consistent with previous research on observational fear-learning (Hygge, Öhman, 1978; Olsson, Phelps, 2004) during the test stage subjects showed a greater SCR to the CS+ compared to CS−, $t(10) = 7.4$ $P < 0.0001$, indicating that they learned about the CS/US contingency. In addition, subjects displayed an enhanced SCR while watching the learning model being presented with the US (shock) as compared to the CS−, $t(10) = 5.0$, $P < 0.001$. The SCR means are presented in Figure 1B. This effect corroborates earlier human data reporting an increased autonomic arousal when perceiving emotional distress in others (Vaughan and Lanzetta, 1980; Levenson and Ruef, 1992). It also supports the notion that a learning model's facial expression of distress can serve as a naturally aversive US (Cook *et al.*, 1985; Cook and Mineka, 1990; Mineka and Cook, 1993).

In our examination of the neural correlates of fear-learning, the amygdala was identified as the a priori region of interest given its central role in fear conditioning. A comparison of the blood-oxygenated-level-dependent (BOLD) signal to the CS+ *vs* the CS− during the test stage revealed bilateral activation of the amygdala (x, y, $z = -20$, -2, -9 and 19, -2, -7 for the left

and right amygdala, respectively, $P < 0.005$) (Figure 2A). During the observation stage, a contrast between the CS+ and the US combined *vs* the CS− also yielded a marked activation in the amygdala bilaterally (x, y, $z = -28$, -5, -12 and 23, -2, -10 for the left and right amygdala, respectively $P < 0.005$) (Figure 2B). To directly investigate if there was an overlap in differential activation between the test and observation stages, we conducted a conjunction analysis based on the two previous contrasts. The conjunction analysis showed a significant overlap of activation in both the left and the right amygdala ($P < 0.05$; x, y, $z = -20$, -9, -13 and 19, (Phelps and LeDoux, 2005), -10 for the left and right amygdala, respectively) (Figure 2C). These findings confirm the known role of the amygdala during both acquisition and expression of learned fear (Phelps and LeDoux, 2005), and extend these functions for the first time to indirect fear-learning through social observation.

To further explore the characteristics of the amygdala response during both the acquisition and expression of observationally attained fear and their anatomical overlap, we identified regions of interest (ROIs) in the left and right amygdala from the contrast of the CS+ *vs* CS− during the test stage (see time courses presented in Figure 2D–G). For each ROI we extracted the mean beta weights for each individual predictor during both stages. Consistent with the contrast used to identify these regions of interest, there was significantly greater activation to the CS+ than the CS− for both the right ($t = 5.8$, $P < 0.001$) and left ($t = 6.7$, $P < 0.001$) amygdala during the test stage. During the observation stage, these same ROIs showed greater activation to the US compared to the CS− in both the right ($t = 4.2$, $P < 0.002$) and left ($t = 3.5$, $P < 0.006$) amygdala. These results confirm
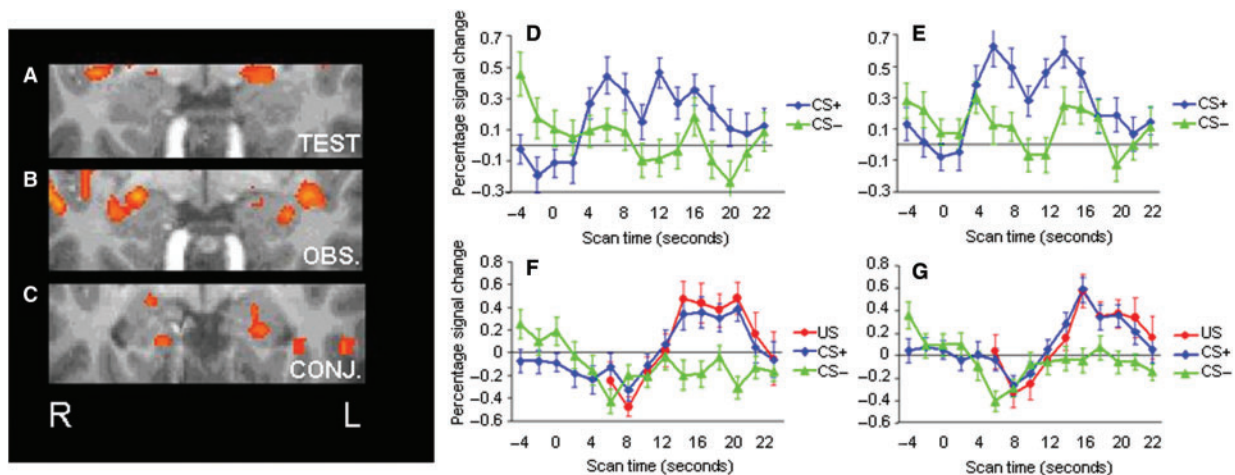

Fig. 2 Amygdala activation at $P < 0.005$, uncorrected, during (**A**) the test stage for the contrast between CS+ *vs* CS− ($x, y, z = -20, -2, -9$ and $19, -2, -7$ for the left and right amygdala, respectively) and (**B**) the observation stage for the contrast between CS+ and US combined *vs* the CS− ($x, y, z = -28, -5, -12$ and $23, -2, -10$ for the left and right amygdala, respectively). (**C**) Shows the regions commonly activated in (a) and (b) ($P < 0.05$, uncorrected; $x, y, z = -20, -9, -13$ and $19, -6, -10$ for the left and right amygdala, respectively). Region of interest (ROI) time courses for amygdala activation (percent signal change) in the (**D**) left and (**E**) right amygdala to the CS+ (blue line) and CS− (green line) during the test stage, and the (**F**) left and (**G**) right amygdala to the CS+ (blue line), US (red line) and CS− (green line) during the observation stage. ROIs were extracted from the CS+ *vs* CS− contrast during the test stage, with a P-threshold of $P < 0.005$, uncorrected. Error bars show standard error.

our results from the conjunction analysis that there was indeed an overlap in amygdala activation during the observation and test stage. Our results showed that the amygdala response was of comparable extent during the test and observation stage, although these BOLD responses peaked at non-overlapping locations. Interestingly, there was no overall response difference to the CS+ *vs* CS− during the observation stage, suggesting that the amygdala's response to observing the model's reactions to the shock paired with the CS was primarily driving the observational learning underlying the acquired anticipatory response to the CS+ as expressed during the test stage. The suggestion that the amygdala activation seen during the observation stage indeed was a signature of fear-learning is supported by the fact that it was the only source of information about the CS+/US contingency in this study.

In addition to the recruitment of the amygdala, previous studies of fear conditioning in humans have reported activation of the anterior insula (AI) and anterior cingulate cortex (ACC) (Buchel *et al.*, 1988; LaBar *et al.*, 1998; Phelps, 2006). For both these regions, robust activation was found for the contrast between the CS+ *vs* CS− in the test stage (*x*, *y*, *z* for left and right AI = −34, 18, 8; 35, 21, 5 and left and right ACC = −1, 13, 33; 5, 36, 27). Significant activation was also observed for the contrast between the CS+ and the US combined *vs* the CS− in overlapping regions of the AI and ACC during the observation stage (*x*, *y*, *z* for left and right AI = −43, 16, −3; 58, 17, 9 and right ACC = 3, 27, 32). However, unlike the amygdala activation, which was of comparable extent during the two stages, the activation in these regions was not as great during the observation stage as compared to the test stage (Tables 1 and 2).

Table 1 Group activations for the test stage

CS+ > CS− area of activation	Number of voxels	Laterality	Talairach coordinates		
			x	*y*	*z*
Superior frontal gyrus	91	R	14	52	30
Middle frontal gyrus	239	L	−25	49	6
Middle frontal gyrus	254	L	−35	46	13
Superior frontal gyrus	16	R	17	37	35
Middle frontal gyrus	30	R	29	43	0
Middle frontal gyrus	399	R	30	37	29
ACC	**782**	**R**	**5**	**36**	**27**
ACC	**340**	**L**	**−1**	**36**	**9**
ACC	**914**	**L**	**−3**	**31**	**25**
ACC	**1216**	**L**	**−1**	**13**	**33**
AI	**992**	**R**	**35**	**21**	**5**
AI	**927**	**L**	**−34**	**18**	**8**
Inferior frontal gyrus	73	R	41	20	−12
Putamen	549	R	21	9	3
Putamen	320	L	−13	9	3
AI	**962**	**L**	**−44**	**7**	**3**
AI	**660**	**R**	**55**	**7**	**6**
Caudate	312	L	−7	2	17
Caudate	179	R	11	2	17
Amygdala (*P* < 0.005, *uncorrected*)	**178**	**R**	**19**	**−2**	**−7**
Amygdala	**321**	**L**	**−20**	**−2**	**−9**
Medial frontal gyrus/motor cortex	223	R	5	−10	71
Medial frontal gyrus/motor cortex	801	L	−3	−11	62
Medial frontal gyrus	693	L	−11	−11	71
Fornix	746		0	−14	18
Thalamus	538	R	11	−14	23
Brain stem	402	R	−5	−23	−6
Posterior cingulate	653		0	−24	32
Posterior cingulate	158		0	−41	20
Inferior parietal	377	L	−64	−26	25
Brain stem	438	L	−6	−29	−8
Cerebellum medial	59	R	5	−50	−18
Cerebellum lateral	256	R	39	−55	−22
Precuneus	380	L	−10	−59	54
Cuneus	133	L	−9	−74	24
Cuneus	267	R	8	−76	35
Lingual gyrus	27	L	−18	−86	−4

P < 0.001, uncorrected, except where noted. Voxel size = 1 mm³.

Table 2 Group activations for the observation stage

CS+ & US (shock) > CS− area of activation	Number of voxels	Laterality	Talairach coordinates		
			x	y	z
Superior frontal gyrus	199	R	11	58	31
Middle frontal gyrus	220	R	37	51	16
Middle frontal gyrus	204	L	−28	52	27
Anterior-rostral medial prefrontal cortex	**302**	**R**	**1**	**46**	**24**
Superior frontal gyrus	361		0	37	43
ACC	**294**	**R**	**3**	**27**	**32**
AI	**518**	**R**	**33**	**19**	**−3**
AI	**665**	**R**	**58**	**17**	**9**
AI	**719**	**L**	**−43**	**16**	**−3**
Broca's area	684	R	47	15	1
AI	**578**	**L**	**−28**	**15**	**−4**
Inferior frontal gyrus	792	L	−43	18	−3
Inferior frontal gyrus	684	R	57	16	9
Amygdala	**203**	**R**	**23**	**−2**	**−10**
Amygdala	**103**	**L**	**−28**	**−5**	**−12**
Superior temporal gyrus	489	R	55	−2	−9
Caudate	222	R	14	−4	23
Precentral gyrus	348	R	37	−5	35
Globus pallidus	251	L	−14	−9	−7
Thalamus	156	L	−11	−11	12
Thalamus	377	R	5	−12	7
Thalamus	492	R	8	−29	0
STS	212	R	46	−33	5
Superior temporal gyrus	679	R	64	−42	21
Superior temporal gyrus	953	L	−61	−42	20
Posterior superior temporal sulcus/Extrastriate Body Area (EBA)	1032	R	51	−62	10
Posterior superior temporal sulcus/Extrastriate Body Area (EBA)	1104	L	−55	−63	9
Cuneus	451	R	10	−93	9
Cuneus	169	L	−16	−95	8

$P < 0.001$, uncorrected, except where noted. Voxel size $= 1\,mm^3$.

Given that prior research has implicated the AI and ACC in empathy (Hutchison *et al.*, 1999; Morrison, *et al.*, 2004; Singer *et al.*, 2004), we performed an exploratory analysis of the relationship between the strength in activation in these ROIs during the observation stage and the subsequent learning response during the test stage. In addition, we conducted the same exploratory analysis on a region that only showed significant activation during the observation stage; the anterior–rostral part of the MPFC which is hypothesized to be involved in mentalizing about self and others (Mitchell *et al.*, 2002; Ochsner *et al.*, 2004; Amodio and Frith, 2006). These ROIs were functionally defined (CS+ and Shock *vs* CS−) in the overall group analysis. For each subject, the beta weight was extracted from each ROI during the observation stage and correlated with the differential SCR response to the CS+ *vs* CS− in the test stage. These analyses revealed a positive correlation between the right AI during observation and the subsequent learning response ($r = 0.61$, $P < 0.05$). Although none of the other regions showed a significant correlation, if we removed one individual who was an outlier in the data set, we observed an effect in the same direction in the superior–rostral ACC ($r = 0.67$, $P < 0.05$) and a trend in the anterior–rostral

medial prefrontal cortex (arMPFC) ($r = 0.49$, $P = 0.08$, one-tailed).

DISCUSSION

Our findings indicate that fear-learning through social observation relies on an associative learning mechanism supported by neural processes similar to those underlying classical fear conditioning. Subjects in our study showed a robust fear response following observation, corroborating previous reports of comparable behavioral (Mineka *et al.*, 1984; Mineka and Cook, 1993) and psychophysiological (Olsson and Phelps, 2004) expressions of fear following observational learning and fear conditioning. Importantly, our imaging data provides the first evidence that the amygdala, which is known to be critical to the acquisition and expression of conditioned fear (Phelps and LeDoux, 2005), is similarly recruited during the acquisition and expression of fear acquired indirectly through social observation. Although there was a significant overlap in amygdala activation during the observation and test stages (Figure 2C), each stage also recruited non-overlapping regions (Figure 2A and B). These differences in activation

may be related to distinct functional processes during the observation (acquisition and expression of fear) and test stage (expression of fear).

In our study, subjects received verbal information that they were themselves going to participate in an experiment similar to the one they watched in the movie, but they were not verbally instructed about the relationship between the CS+ and the shock US. This information could only have been accessed through watching the CS+/US pairings in the movie. Nevertheless, it is possible that our instructions (see 'Methods' section) affected the learning responses seen during test. Although we cannot rule out an influence of verbally accessible knowledge in our findings, both our own data and previous research strongly suggest that it was the observed CS–US pairings that mediated the expression of learning. First, our conjunction and ROI analyses demonstrated overlapping regions of amygdala activation during the sequence of the movie when the model reacted aversively to the shock and amygdala activation during the subsequent test stage. Conforming to a vast literature on the role of the neural bases of fear conditioning, the amygdala showed bilateral activation during both the observation and test stages. In contrast, previous studies examining the expression of fear acquired solely through verbal instructions, using CSs identical to the ones in our experiment, have implicated only the left amygdala (Phelps et al., 2001; Funayama et al., 2001). Second, an earlier study from our lab used an observational learning procedure identical to the one in the current study to compare observational with instructed and conditioned fear-learning (Olsson and Phelps, 2004). These results demonstrated that both observational and conditioned fear resulted in similarly strong fear responses to unseen (backwardly masked) presentations of the CSs. In contrast, unseen CSs did not elicit a fear response in subjects that learned about the threat value of these stimuli only through verbal instructions. These data further underscore the similarities between fear responses acquired through the current observational learning paradigm and traditional fear conditioning.

Unlike the pattern of amygdala activation, which was of comparable extent during the two stages, activation in the ACC and AI was not as great during the observation stage as compared to the test stage (Tables 1 and 2). Activation of the ACC and AI has been previously linked to the anticipation and experience of pain (Davis, 2000; Peyron et al., 2000), as well as autonomic arousal (Critchley, 2005). The difference in the extent of activation in these regions during the observation and test stages may simply reflect the relative difference in the anticipation of pain or arousal the subjects experienced during these two stages.

It is also possible that the activation in the ACC and AI during the observation stage is linked to empathy. Learning about the circumstances causing distress in another person may involve taking that person's emotional perspective, triggering empathic responses in the observer. Previous research on the neural systems of empathy has demonstrated that parts of the superior–rostral zone of the ACC and the AI are recruited both when one-self is anticipating and experiencing pain and when pain is inflicted to someone else, highlighting these regions as important components in a system supporting the affective qualities of empathy (Hutchison et al., 1999; Morrison et al., 2004; Singer et al., 2004; Jackson et al., 2005). The activation seen in these areas during both the observation stage and the test stage in our study suggest that empathy may play a role in fear-learning through social observation. This assumption was further supported by the fact that activation in the AI (and to a lesser extent the ACC) predicted the subsequent learning response during the test stage.

During the observation stage, we also observed activation in the anterior–rostral medial prefrontal cortex (arMPFC) that was not apparent in the test stage. The arMPFC has been implicated in thinking about one own's and others' mental states (Mitchell et al., 2002; Ochsner et al., 2004; Amodio and Frith, 2006). The selective recruitment of this region during the observation stage provides further support for a role for empathy and, suggests that the attribution of mental states to the learning model may be involved in observational fear-learning.

Interestingly, most previous fMRI studies on empathy (Morrison et al., 2004; Singer et al., 2004; Jackson et al., 2005) have failed to report the involvement of the amygdala (but see also Botvinick et al., 2005). However, in our study this region was recruited both by the perception of pain in another and the expectancy to receive the same treatment. There are two possible reasons for this difference. First, the self-relevance of the learning context in which subjects were placed may have been critical. Our study is the first to manipulate subjects' motivation to learn from the circumstances surrounding the emotional distress in another. In our experiment, the learning model's emotional expression served as an US, and its co-occurrence with the colored squares (CS) was made directly self-relevant to the subject due to its predictability of future potentially harmful events. The motivation to learn about potentially harmful qualities in the surrounding may have triggered fear-learning mechanisms, which are known to be dependent on the amygdala (Phelps and LeDoux, 2005).

Secondly, the emotional expression displayed by the learning model may alone have served as an US without the explicit motivation to learn from the circumstances predicting the expression. In support of this, previous paradigms using images of facial (Whalen et al., 1998; Adolphs, 2003; Carr et al., 2003; Decety and Chaminade, 2003) or whole-body (de Gelder et al., 2004) emotional, in particular fearful, expressions have reported the involvement of the amygdala, without making the perceived emotional expressions explicitly self-relevant. Unlike some previous studies examining the neural systems of empathy,

which have used arbitrary cues to signal pain in another (Morrison *et al.*, 2004; Singer *et al.*, 2004) or images of single body parts being harmed (Jackson *et al.*, 2005), we presented subjects with both facial and bodily expressions of distress and pain responses displayed by the learning model. In the current study, we cannot determine whether the involvement of the amygdala is due to the self-relevance of the learning task, a response to the perceived emotional expression serving as an US, or both.

Our finding that the formation and the expression of fear through social observation relies on neural circuits that are similarly involved in fear conditioning is in accordance with the description of this form of learning as an evolutionarily old system for the transmission of emotionally relevant information as documented in a wide range of species (Öhman and Mineka, 2001; Phelps and LeDoux, 2005). Previous data have shown that the expression of observational and conditioned fear, unlike fear attained exclusively by symbolic means through verbal communication, are both partially independent on cognitive awareness (Olsson and Phelps, 2004), further supporting the conclusion that these forms of learning rely on similar neural mechanisms.

In our daily lives, we are frequently exposed to vivid images of others in emotional situations through personal social interactions as well as the media. The knowledge of somebody else's emotional state may evoke empathic responses. However, as our results reveal, when others' emotions are accompanied with vivid expressions and perceived as potentially relevant to our own future well being, we may engage additional learning mechanisms. The present results show that fears learned by observing others engage the same neural mechanisms as fear acquired through direct experience, suggesting that that social and non-social means of fear learning may be equally effective and powerful.

Conflict of Interest
None declared.

REFERENCES

Adolphs, R. (2003). Cognitive neuroscience of human social behavior. *Nature Neuroscience*, 4, 165–78.

Amodio, D.M., Frith, C.D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, 7, 268–77.

Bandura, A. (1977). *Social Learning Theory*. New York: General Learning Press.

Berger, S. (1962). Conditioning through vicarious instigation. *Psychological Review*, 69, 450–66.

Botvinick, M., Jha, A.P., Bylsma, L.M., Fabian, S.A., Solomon, P.E., Prkachin, K.M. (2005). Viewing facial expressions of pain engages cortical areas involved in the direct experience of pain. *Neuroimage*, 25, 312–19.

Buchel, C., Morris, J., Dolan, R.J., Friston, K.J. (1988). Brain systems mediating aversive conditioning: an event-related fMRI study. *Neuron*, 20, 947–57.

Carr, L., Iacoboni, M., Dubeau, M.C., Mazziotta, J.C., Lenzi, G.L. (2003). Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas. *Proceedings of the National Academy of Science USA*, 29, 5497–502.

Cook, M., Mineka, S. (1990). Selective associations in the observational conditioning of fear in rhesus monkeys. *Journal of Experimental Psychology: Animal Behavior Processes*, 16, 372–89.

Cook, M., Mineka, S., Wolkenstein, B., Laitsch, K. (1985). Observational conditioning of snake fear in unrelated rhesus monkeys. *Jornal of Abnormal Psychology*, 94, 591–610.

Critchley, H.D. (2005). Neural mechanisms of autonomic, affective, and cognitive integration. *Journal of Comparative Neurology*, 493, 154–66.

Curio, E. (1988). Cultural transmission of enemy recognition. In: Zentall, R.E., and Galef, B.G. Jr, editors. *Social Learning: Psychological and Biological Perspectives*. Hillsdale, NJ: Erlbaum, pp. 75–97.

Davis, D. (2000). The neural circuitry of pain as explored with functional MRI. *Neurological Research*, 22, 313–17.

Davis, M., Whalen, P.J. (2001). The amygdala: vigilance and emotion. *Molecular Psychiatry*, 6, 13–34.

de Gelder, B., Snyder, J, Greve, D., Gerard, G., Hadjikhani, N. (2004). Fear fosters flight: a mechanism for fear contagion when perceiving emotion expressed by a whole body. *Proceedings of the National Academy of Science USA*, 101, 16701–6.

Decety, J., Chaminade, T. (2003). Neural correlates of feeling sympathy. *Neuropsychologia*, 41, 127–38.

Fanselow, M.S., LeDoux, J.E. (1999). Why we think plasticity underlying Pavlovian fear conditioning occurs in the basolateral amygdala. *Neuron*, 23, 229–32.

Funayama, E.S., Grillon, C., Davis, M., Phelps, E.A. (2001). Activation of the left amygdala to a cognitive representation of fear. *Journal of Cognitive Neuroscience*, 13, 721–29.

Hutchison, W.D., Davis, K.D., Lozano, A.M., Tasker, R.R., Dostrovsky, J.O. (1999). Pain-related neurons in the human cingulate cortex. *Nature Neuroscience*, 2, 403–5.

Hygge, S., Öhman, A. (1978). Modeling processes in the acquisition of fears: Vicarious electrodermal conditioning to fear-relevant stimuli. *Journal of Personality and Social Psychology*, 36, 271–79.

Jackson, P.L., Meltzoff, A., Decety, J. (2005). How do we perceive the pain of others? A window into the neural processes involved in empathy. *Neuroimage*, 24, 771–79.

John, E.R., Chesler, P., Bartlett, F., Victor, I. (1968). Observation learning in cats. *Science*, 59, 1489–90.

Kapp, B.S., Whalen, P.J., Supple, W.F., Pascoe, J.P. (1992). Amygdaloid contributions to conditioned arousal and sensory information processing. In: Aggleton, J.P., editor. *The Amygdala: Neurobiological Aspects of Emotion, Memory, and Mental Dysfunction*. New York: Wiley-Liss, pp. 229–32.

Kavaliers, M., Choleris, E., Colwell, D.D. (2001). Learning from others to cope with biting flies: social learning of fear-induced conditioned analgesia and active avoidance. *Behavioral Neuroscience*, 115, 661–74.

LaBar, K.S., LeDoux, J.E., Spencer, D.D., Phelps, E.A. (1995). Impaired fear conditioning following unilateral temporal lobectomy in humans. *Journal of Neuroscience*, 15, 6846–55.

LaBar, K.S., Gatenby, J.C., Gore, J.C., LeDoux, J.E., Phelps, E.A. (1998). Human amygdala activation during conditioned fear acquisition and extinction: a mixed-trial fMRI study. *Neuron*, 20, 937–45.

Levenson, R.W., Ruef, A.M. (1992). Empathy: a physiological substrate. *Journal of Personality and Social Psychology*, 63, 234–46.

Miller, N., Dollard, J. (1941). *Social Learning and Imitation*. New Haven, NJ: Yale University Press.

Mineka, S., Cook, M. (1993). Mechanisms involved in the observational conditioning of fear. *Journal of Experimental Psychology: General*, 122, 23–38.

Mineka, S., Davidson, M., Cook, M., Keir, R. (1984). Observational conditioning of snake fear in rhesus monkey. *Journal of Abnormal Psychology*, 93, 355–72.

Mitchell, J.P., Heatherton, T.F., Macrae, C.N. (2002). Distinct neural systems subserve person and object knowledge. *Proceedings of the National Academy of Sciences*, 99, 15238–43.

Morrison, I., Lloyd, D., di Pellegrino, G., Roberts, N. (2004). Responses to pain in anterior cingulate cortex: is empathy a multisensory issue? *Cognitive, Affective and Behavioral Neuroscience*, *4*, 270–78.

Ochsner, K.N., Knierim, K., Ludlow, D.H. et al. (2004). Reflecting upon feelings: an fMRI study of neural systems supporting the attribution of emotion to self and other. *Journal of Cognitive Neuroscience*, *16*, 1746–72.

Öhman, A., Mineka, S. (2001). Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychological Review*, *108*, 483–522.

Olsson, A., Phelps, E.A. (2004). Learned fear of 'Unseen' faces after Pavlovian, observational, and instructed fear. *Psychological Science*, *12*, 822–28.

Peyron, R., Laurent, B., Garcia-Larrea, L. (2000). Functional imaging of brain responses to pain. A review and meta-analysis. *Neurophysiological Clinical*, *30*, 263–88.

Phelps, E.A. (2006). Emotion and cognition: insights from studies of the human amygdala. *Annual Review of Psychology*, *57*, 27–53.

Phelps, E.A., LeDoux, J.E. (2005). Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron*, *48*, 175–87.

Phelps, E.A., O'Connor, K.J., Gatenby, J.C., Gore, J.C., Grillon, C., Davis, M. (2001). Activation of the amygdala by cognitive representations of fear. *Nature Neuroscience*, *4*, 437–41.

Rachman, S. (1977). The conditioning theory of fear acquisition: a critical examination. *Behavioral and Reseach Therapy*, *19*, 439–47.

Vaughan, K.B., Lanzetta, J.T. (1980). Vicarious Instigation and conditioning of facial expressive and autonomic responses to a model's expressive display of pain. *Journal of Personality and Social Psychology*, *38*, 909–23.

Whalen, P.J., Rauch, S.L., Etcoff, N.L., McInerney, S.C., Lee, M.B., Jenike, M.A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *Journal of Neuroscience*, *18*, 411–18.