# EXPICIT AND IMPLICIT PROCESSES IN HUMAN AVERSIVE CONDITIONING

Thesis by

R. McKell Carter

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2006

*For my family*

*new or old*

*blood or bond*

## Acknowledgements

It seems unfair that I receive a degree for the work that many have done. To all of you, I can only hope I have helped you, or can help, in return. Research grinds to a halt without funding to keep it moving. I would like to thank the following for their sponsorship: Sandia National Labs, the David and Lucile Packard Foundation, the Gordon and Betty Moore Foundation, the Keck Foundation, the National Institutes of Health, the National Institutes of Mental Health, the National Science Foundation (and ERC), the Wellcome Trust and the William T. Gimbel Discovery Fund in Neuroscience at Caltech

All of my collaborators were both great scientists and enjoyable colleagues. I would like to thank all of you. Nao and Connie spent time on the early conditioning work, transforming its direction. Conferences would not have been the same without you. Ben Seymour was the king of pain and scotch eggs at the FIL. The study of trace and delay conditioning was made a great deal more entertaining with him and John O'Doherty in the control room. I also thank the rest of the FIL for their endless help and discussion. For those who have the opportunity, the FIL is a great place to live and work. Brian Cleary for his time spent thinking about functional DTI, a project that I hope will be continued. Thanks to Tony Bruguier for a great deal of time spent scratching our heads over noise in an MRI environment. Beena, Kats, and Romi all deserve thanks for introducing me to psychophysics and for a great project not included in this thesis (Khurana et al., 2006). I would like to thank the members of the Conscious Mouse Club for thought provoking discussion at meetings well after most have gone home.

Members of the Caltech Biological Imaging Center were excellent in their efficiency and assistance. The always knowledgeable Mike Tyszka and Steve Flaherty

helped make both skin conductance recording during fMRI and the rapid acquisition of DTI images possible. Nothing would be accomplished at the CBIC without Mary Munoz.

There are a number of professors who had to suffer my harassment and sometimes serve on my committees. They have all provided valuable input and I am glad to have spent time interacting with them, thank you. Ralph Adolphs, John Allman, David Anderson, Michael Fanselow and Shinsuke Shimojo all had the thankless task of serving on my committees. Ray Dolan was my host at the FIL for two projects. I rotated in Gilles Laurent's lab and still consider his lab to be one of the best at Caltech. Henry Lester was part of the Conscious Mouse Club and was a tireless advocate of clarity in presentation.

John O'Doherty deserves special thanks for introducing me to SPM and MRI as well as The Queen's Larder. I would still be lost in colored blobs had it not been for his capable direction. You have been a great friend, and I wish you even greater success here at Caltech.

I would like to thank my advisor Christof Koch. The time spent analyzing, climbing, debating, driving, running, talking and writing have all been valuable to me. You have served as a model for both academics and attitude. I thank you for your guidance and friendship. I also thank you for providing opportunities that have greatly enriched my stay at Caltech.

My interest in research was started by John Roth at the University of Utah. He has a talent for puzzles and communication that make his work very unique. He focused time and effort on each project in his lab on an almost daily basis without micromanaging, an extremely rare skill. I am very grateful for my time in your lab.

Finally, I would like to thank those who made it possible to be happy at Caltech. My family, for they are who define me and nothing is possible without them. Thanks to Klab, my doctoral family. I thank all of my friends at Caltech. I had only met a few rare individuals who were as interesting, intelligent, and educated as the whole pack I met here. Foremost among them is my fiancé Karli Watson, who continues to baffle me to this very day (and I hope she always will).

## Abstract

The ability to adapt to a changing environment is central to an organism's success. The process of associating two stimuli (as in associative conditioning) requires very little in the way of neural machinery. In fact, organisms with only a few hundred neurons show conditioning that is specific to an associated cue. This type of learning is commonly referred to as implicit learning. The learning can be performed in the absence of the subject's ability to describe it. One example of learning that is thought to be implicit is *delay conditioning***.** Delay conditioning consists of a single cue (a tone, for example) that starts before, and then overlaps with, an outcome (like a pain stimulus).

In addition to associating sensory cues, humans routinely link abstract concepts with an outcome. This more complex learning is often described as explicit since subjects are able to describe the link between the stimulus and outcome. An example of conditioning that requires this type of knowledge is *trace conditioning*. Trace conditioning includes a separation of a few seconds between the cue and outcome. Explicit learning is often proposed to involve a separate system, but the degree of separation between implicit associations and explicit learning is still debated.

We describe aversive conditioning experiments in human subjects used to study the degree of interaction that takes place between explicit and implicit systems. We do this in three ways. First, if a higher order task (in this case a working memory task) is performed during conditioning, it reduces not only explicit learning but also implicit learning. Second, we describe the area of the brain involved in explicit learning during conditioning and confirm that it is active during both trace and delay conditioning. Third, using functional magnetic resonance imaging (fMRI), we describe hemodynamic activity changes in perceptual areas of the brain that occur during delay conditioning and persist after the learned association has faded.

From these studies, we conclude that there is a strong interaction between explicit and implicit learning systems, with one often directly changing the function of the other.

Table of Contents


**1 INTRODUCTION ..... 1**

**1.1 Explicit and Implicit Aspects of Conditioning ..... 2**

**1.2 Our Approach ..... 5**

**2 WORKING MEMORY AND FEAR CONDITIONING ..... 8**

**2.1 Introduction ..... 8**

**2.2 Materials and Methods ..... 10**
2.2.1 Equipment ..... 10
2.2.2 Subjects ..... 11
2.2.3 Procedure ..... 11
2.2.4 Conditioning Stimuli (Figure 2-1A) ..... 12
2.2.5 Experimental Phases (Figure 2-1B) ..... 13
2.2.6 Distracting Tasks (Figure 2-1C) ..... 14
2.2.7 Analysis of SCR ..... 15
2.2.8 Trial Effects ..... 16

**2.3 Results ..... 17**
2.3.1 Differential Conditioning ..... 17
2.3.2 Awareness of CS/US Contingency ..... 18
2.3.3 Single Cue Conditioning ..... 18

**2.4 Discussion ..... 20**

**2.5 Figures and Legends ..... 25**

**3 THE NEURAL CORRELATES OF IMPLICIT AND EXPLICIT PROCESSES IN CONDITIONING ..... 31**

**3.1 Introduction ..... 32**

**3.2 Methods ..... 34**
3.2.1 Participants ..... 34
3.2.2 Experimental Procedure ..... 35
3.2.3 Online Subject Reports of Contingency Awareness ..... 35
3.2.4 Post-Experimental Questionnaire ..... 36
3.2.5 Unconditioned Stimuli ..... 36
3.2.6 Online Measure of SCR ..... 37
3.2.7 fMRI ..... 37
3.2.8 Learning: Contingency awareness and Conditioning ..... 39

**3.3 Results ..... 40**
3.3.1 Conditioned skin conductance responses ..... 40
3.3.2 Contingency awareness ..... 40

**3.4 Discussion ..... 41**

**3.5 Tables ..... 45**

**3.6 Figures ..... 47**

**3.7 Supplementary Methods ..... 50**

**4 PERSISTENT CHANGES IN VISUAL CORTEX DUE TO AVERSIVE CONDITIONING ..... 51**

**4.1 Introduction ..... 51**

**4.2 Methods ..... 54**
4.2.1 Participants ..... 54
4.2.2 Experimental Procedure ..... 54
4.2.3 Stimulus Presentation ..... 55
4.2.4 Skin Conductance Conditioning ..... 56
4.2.5 fMRI ..... 57

**4.3 Results ..... 58**
4.3.1 Skin Conductance Conditioning ..... 58
4.3.2 Brain Activity during Acquisition ..... 59
4.3.3 Extinction ..... 61

**4.4 Discussion ..... 62**

**4.5 Tables ..... 67**

**4.6 Figures ..... 69**

**5 DISCUSSION ..... 77**

**5.1 Explicit Influences on Implicit Processes ..... 77**

**5.2 Implicit Influences on Explicit Processes ..... 79**

**5.3 Continuum or Separate Systems ..... 81**

**5.4 Work on the Conscious Mouse ..... 82**

**5.5 Conclusion ..... 83**

**6 APPENDIX – SCR RECORDING DURING FMRI ACQUISITION ..... 85**

**6.1 Abstract ..... 85**

**6.2 Introduction ..... 86**

**6.3 Material and methods ..... 88**
6.3.1 Equipment ..... 88
6.3.2 Body simulation ..... 88
6.3.3 Leads and electrodes ..... 89
6.3.4 Heat insulation ..... 90
6.3.5 Filter design ..... 90

6.3.6 Head coils....91
6.3.7 Resonance testing....92
6.3.8 Measurement of induced voltages....92

**6.4 Results....93**
6.4.1 Resonances....93
6.4.2 Recorded waveform....93
6.4.3 Recorded voltages....94
6.4.4 GSR recording quality....95

**6.5 Discussion....96**

**6.6 Tables....97**

**7 REFERENCES....106**

Table list

TABLE 3-1.....45
TABLE 3-2.....46
TABLE 4-1.....67
TABLE 4-2.....68
TABLE 6-1.....97

# Figure List

FIGURE 2-1 ..... 25
FIGURE 2-2 ..... 26
FIGURE 2-3 ..... 27
FIGURE 2-4 ..... 28
FIGURE 2-5 ..... 29
FIGURE 2-6 ..... 30
FIGURE 3-1 ..... 47
FIGURE 3-2 ..... 48
FIGURE 3-3 ..... 49
FIGURE 4-1 ..... 69
FIGURE 4-2 ..... 70
FIGURE 4-3 ..... 71
FIGURE 4-4 ..... 72
FIGURE 4-5 ..... 73
FIGURE 4-6 ..... 74
FIGURE 4-7 ..... 75
FIGURE 4-8 ..... 76
FIGURE 6-1 ..... 99
FIGURE 6-2 ..... 100
FIGURE 6-3 ..... 101
FIGURE 6-4 ..... 102
FIGURE 6-5 ..... 103
FIGURE 6-6 ..... 104
FIGURE 6-7 ..... 105

## Abbreviations

| | |
|---|---|
| ACC | Anterior Cingulate Cortex |
| BOLD | Blood Oxygenation Level Dependent |
| CS | conditioned stimulus |
| CS+ | the conditioned stimulus sometimes followed by a US |
| CS- | the conditioned stimulus never followed by a US |
| DLPFC | Dorsal Lateral Prefrontal Cortex |
| FFA | Fusiform Face Area |
| fMRI | functional Magnetic Resonance Imaging |
| FWE | Family Wise Error |
| FWHM | Full Width at Half Maximum. |
| GP | Globus Pallidus |
| GSR | Galvanic Skin Response |
| Ins. | Insula |
| IPL | Inferior Parietal Lobule |
| IPS | Inferior Parietal Sulcus |
| MFG | Middle Frontal Gyrus |
| MRI | Magnetic Resonance Imaging |
| Operc. | Operculum |
| PFC | Prefrontal Cortex |
| ROI | Region of Interest |
| S | Siemen (as in micro Siemen or nano Siemen) |
| SC | Superior Colliculus |
| SCR | Skin Conductance Response |
| SPM | Statistical Parametric Map |
| STS | Superior Temporal Sulcus |
| SVC | Small Volume Correction |

# 1 Introduction

*Every extension of knowledge arises from making the conscious the unconscious – Nietzsche*

Organisms need to learn. Learning provides the basis for adaptation to a diverse and changing environment. Those organisms that are better at learning are more successful both in the acquisition of resources and the avoidance of potentially detrimental situations. When learning and memory are discussed in an everyday context, it is usually to retain information for a test or to avoid forgetting a person's name. This type of memory usually involves facts or concepts and is described as explicit. There are cases where individuals have shown an extraordinary associative capacity to the point where sensory associations from their explicit memories overpower current experience.

One such individual is the famous Solomon Shereshevskii, known in the literature simply as 'S'. The extent of his success (and difficulties) was described by the neuropsychologist Alexander Luria (Luria, 1968). From the Wikipedia entry for Shereshevskii, 'S' was a Russian journalist whose abilities were discovered when he was scolded for not taking notes at a speech. When questioned, 'S' was able to recite the speech verbatim. 'S' was tested for decades and his abilities were exceptional; in a matter of minutes, he was capable of memorizing text in languages he had never been exposed to. In his book, Luria describes the abilities of 'S' as being related to synaesthesia, the experience of sensation in one sensory modality when presented with a stimulus in another (for example, smelling color). 'S' formed complex representations of meaningless symbols that bridged multiple modalities. A nonsense syllable might produce a sharply shaped cloud that tasted sour, for example. Refining these associative

abilities, he became a very successful mnemonist, able to retain a great deal more than he could without the techniques. Most surprising was that 'S' scored absolutely average on intelligence tests. While his abilities sound incredibly useful to anyone who has ever stood stammering while trying to recall a name or reference, this ability had its cost. The extent of his associations eventually left him unable to interact normally; for instance he once explained that he was unable to eat strawberry ice cream because the tone of the ice cream's vendor left the taste of coal in his mouth. He spent the later portion of his life in an asylum. 'S' is an example of nearly perfect explicit association; each stimulus had such a rich sensory representation that he was able to form episodic memories that were very robust.

A type of learning that is better studied and more common is that of conditioned association, often referred to as implicit since it does not require the subject to be aware of the association. Conditioning is studied in a wide variety of organisms from mollusks to fruit flies, rodents, monkeys and humans (Baer and Fuhrer, 1982; Mackintosh, 1983; Gallistel, 1990; Thompson and Krupa, 1994; Connolly et al., 1996; Eichenbaum, 1997; Pearce et al., 1997; Tully, 1998; Squire and Kandel, 1999; Kocorowski and Helmstetter, 2001). This gives scientists a large number of tools to study how this association takes place. The most notable model used to study association is that of Pavlovian conditioning.

## *1.1 Explicit and Implicit Aspects of Conditioning*

Pavlovian conditioning involves the association of a previously neutral stimulus, such as a bell, with a meaningful stimulus, such as food. Initially, the subject of study has a

reaction, such as salivation, to only the meaningful stimulus. Over time, the subject begins to respond to the previously neutral stimulus in the same way as the meaningful one. The subject has formed an association; he or she now begins salivating to the presence of the bell alone without food. In Pavlovian conditioning terms, the initially neutral stimulus (the bell) is referred to as the Conditioned Stimulus or CS. The initially meaningful stimulus (the food) is referred to as the Unconditioned Stimulus or US.

This thesis examines the interaction between implicit and explicit learning. The simple association that takes place in most organisms is often described as *implicit*, occurring without any relationship to conscious knowledge (Manns et al., 2002). Learning a person's name or associating two abstract concepts is described as *explicit* since it occurs, essentially by definition, with conscious knowledge.

It is easy to understand how conditioning could be considered an implicit process; even the 302 neurons in the roundworm *Caenorhabditis elegans* show learning in a large number of association paradigms (Rankin et al., 1990; Wen et al., 1997; Law et al., 2004). With such a small number of neurons involved, it becomes more difficult to imagine that explicit knowledge is involved in the process.

Although the differentiation between "explicit" and "implicit" has intuitive weight, what makes a particular learned relationship explicit? The study of a patient referred to as HM served as a means to split the two learning systems. HM's case was first described by Scoville and Milner in 1957 (Scoville and Milner, 1957). HM was involved in an accident at a young age that eventually resulted in epilepsy. His condition was bad enough that the medial temporal lobes, including an area called the hippocampus, on both sides of his brain were removed. After the operation, HM was unable to form new

memories. He retained older memories, but was unable to recall what he had eaten for breakfast or what he had done yesterday. He could carry on a conversation as long as there weren't too many changes of topic, and as long as it was less than a few minutes. While HM could not form any new explicit memories, his implicit learning remained intact. Over several days he learned to trace objects in a mirror without ever being able to report that he had tried the task before. He could learn new skills, motor associations, without any explicit knowledge. This finding has been generally interpreted to mean that explicit and implicit learning depend on two separate memory systems (Squire and Kandel, 1999).

A set of experiments by Larry Squire and colleagues described a further dissociation between explicit and implicit learning using two different types of eye-blink conditioning. In eye-blink conditioning, a neutral stimulus, such as a tone (the CS), is paired with a puff of air to the eye, the US. As the experiment progresses, the subject learns to blink (non-consciously) at the appropriate time to protect their eye.

Delay conditioning is an example of Pavlovian conditioning that has been described as occurring independently of awareness (Manns et al., 2002). For the association to be described as delay conditioning, the CS must precede the US, either overlapping with it or directly before it. It is called delay conditioning because the US presented is *delayed* with respect to the CS.

Trace conditioning is an example of Pavlovian conditioning that is believed to occur only when the subject has acquired conscious knowledge (Clark and Squire, 1998). The difference between delay and trace conditioning is that in trace conditioning, the CS is separated in time from the US. It is called *trace conditioning* because the association of

the neutral (CS) and meaningful (US) stimulus requires that a *memory trace* of the CS be kept after the CS terminates in order to associate it with the US. In spite of its relative complexity, there is evidence that this type of learning occurs in the fruit fly *Drosophila melanogaster* (Tully and Quinn, 1985).

There are many examples of explicit influences in implicit processes. Studies from as early as 1937 describe conditioning physiological responses using only verbal instruction (Cook and Harris, 1937); that is, the subject begins to respond to a previously neutral stimulus because of false instructions that say the stimulus may now be paired with a shock.

## *1.2 Our Approach*

We sought to further examine the interactions between explicit and implicit processes in a classical conditioning paradigm. We chose to use a fear conditioning paradigm rather than an eye-blink conditioning paradigm because fear conditioning is well studied in a wide variety of organisms. Fear conditioning's widespread use provided us not only with better information about the paradigm being studied, but also made model systems available where electrophysiology or lesion tools could be used. It was also clear that past work with rats could be reliably and quickly reproduced in mice, a model system our lab was interested in working with for the genetic tools available. The first section of this thesis was completed in collaboration with labs at Caltech (David Anderson and Henry Lester) and UCLA (Michael Fanselow). The aim of this collaboration was to design and perform experiments with the other labs in mind, making what was learned from studies of explicit learning in humans applicable to rodent systems, and what was learned from

lesion studies in rodents applicable to behavioral results with human subjects. Details of the collaboration are contained in the discussion. This thesis is organized in three main parts.

Chapter 2 describes the effects of performing a working memory task during delay and trace conditioning. We reasoned that if trace conditioning depended on high level mental resources, such as working memory, then having subjects perform a working memory task during conditioning would eliminate trace conditioning, leaving delay unaffected. Instead, we discovered that the working memory task affected not only trace conditioning, but delay as well. These effects could be partially overcome by simplification of the protocol; for example, reducing the number of stimuli or providing the subject with information before the experiment. This study provides strong evidence of the influence that explicit processes have on implicit ones.

Chapter 3 describes areas of the brain that are important for the acquisition of both explicit and implicit information. Subjects were aversively conditioned using both trace and delay protocols during fMRI acquisition. We also recorded skin conductance responses for use as a correlate of implicit learning, and shock expectancy responses as a correlate of explicit learning. Our analysis identified portions of the brain where hemodynamic responses correlated with both of these measures. Consistent with the result of Chapter 2, the middle frontal gyrus, an area associated with working memory performance, correlated with the accuracy of shock expectancy. We also found that the amygdala, hippocampus, and areas of visual cortex correlated with the implicit measure.

The identification of visual cortex as an area correlating with implicit learning led us to specifically examine the visual areas representing the CS.

Chapter 4 describes an experiment assessing blood oxygenation changes in order to assess differences in the fusiform face area that occur as a result of delay conditioning to images of faces. These changes are consistent with an increase in the representation of the paired stimulus. They are also persistent beyond extinction of the conditioned association, indicating lasting changes in visual cortex that could continue to have effects on explicit responses well after any aversive response is removed. This study was also designed to show that the stimuli used in Chapter 2 caused similar changes in visual cortex. There was no substantial evidence indicating that this was true. Possible reasons for this are discussed in Chapter 4.

# 2 Working Memory and Fear Conditioning

This work was published under the same title in 2003 (Carter et al., 2003).

Constanze Hofstötter conducted the single cue trace uninformed 0-back experiment and all of the informed experiments. She was also involved in the analysis and write up. Her contributions in the writing process made the manuscript far better than it would have been otherwise. Naotsugu Tsuchiya conducted unpublished control experiments and was also involved in the analysis and write up. Christof Koch initiated the project and secured funding. His input in the early stages of the project (while we entered a field we had no experience in) was always very useful. He also advised on analysis procedures and made substantial contributions in the write up and review processes. Experiments and analysis were conducted at the California Institute of Technology.

Here, we investigate the extent to which human classical fear conditioning depends on working memory.

## *2.1 Introduction*

Pavlovian conditioning is widely used to study associative learning in species ranging from mollusks to flies, rodents, monkeys and humans (Baer and Fuhrer, 1982; Mackintosh, 1983; Gallistel, 1990; Thompson and Krupa, 1994; Connolly et al., 1996; Eichenbaum, 1997; Pearce et al., 1997; Tully, 1998; Squire and Kandel, 1999; Kocorowski and Helmstetter, 2001). This form of learning involves the association of an initially neutral stimulus, the conditioned stimulus (CS), with a correlated meaningful stimulus, the unconditioned stimulus (US). An unresolved question concerns the extent to

which certain forms of classical conditioning depend on higher-level cognitive processes including selective attention, working memory and awareness (Hilgard et al., 1937; Dawson and Furedy, 1976; Clark and Squire, 1998; Ohman and Soares, 1998; Carrillo et al., 2000; Knuttinen et al., 2001; Lovibond and Shanks, 2002). Eye-blink conditioning is an associative learning paradigm where the role of explicit knowledge / awareness is being investigated. The paradigm involves the association of an eye-blink (a somatic motor response) with previously meaningless stimuli (CS).

Recent data showed that, in humans, associative trace conditioning of eye-blink responses requires awareness of the contingency between the CS (a tone) and the US (a puff of air to the eye), while this is not the case for delay conditioning (Clark and Squire, 1998; Clark and Squire, 1999; Manns et al., 2000b, a). In delay conditioning, the start of the US is temporally contiguous with the CS, while in trace conditioning, an interval is interposed between the end of the CS and the start of the US. Distracting subjects by having them perform a secondary task (for example, a verbal shadowing task) during a trace procedure prevents conditioning. Furthermore, subjects' ability to report the exact nature of the CS/US relationship (e.g., "I believe the tone came before the air puff") is greatly impaired with concurrent distraction during trace conditioning. Conversely, associative delay eye-blink conditioning appears to be insensitive to distracters. Other experiments find that both trace and delay associative differential conditioning can be disrupted by tasks that demand sufficient attention, while this is not the case for single cue conditioning paradigms (Carrillo et al., 2000; Knuttinen et al., 2001). In single cue conditioning, only one CS is presented (paired with the US). In differential conditioning, two CSs are presented, one of which is correlated with US presentations (CS+), while the

other is not (CS-).

We chose fear conditioning to replicate and extend these findings with human subjects on the basis of a conditioning protocol easily extendible to mice, animals for which well established molecular tools used for manipulating genetically identifiable cell populations are available. Fear conditioning differs from eye-blink conditioning in its underlying neuronal implementation, due in part to the fact that the association involves an autonomic, rather than a somatic, motor response. Fear conditioning is easy to establish in humans and rodents, is acquired in a fraction of the trials needed for eye-blink conditioning and is tolerant to long trace periods, making it amenable to fMRI investigations (Buchel et al., 1998b; LaBar et al., 1998; Buchel et al., 1999; Knight et al., 1999). Finally, the neural circuits underlying fear conditioning, particularly the lateral nucleus of the amygdala, hippocampus and prefrontal cortex, are being vigorously explored (Fendt and Fanselow, 1999; Medina et al., 2002). We use transient elevations in skin conductance (skin conductance response or SCR) as our measure of autonomic arousal when testing responses to auditory stimuli that have been previously paired with a shock. At the same time, we distract our subjects with tasks of variable working memory load. There were parallel efforts to reproduce selected aspects of this work in mice (Han et al., 2003).

## *2.2 Materials and Methods*

### 2.2.1 Equipment

Conditioning stimuli were presented and SCRs were recorded using equipment from Contact Precision Instruments (www.psylab.com), controlled by Psylab software. Silver/

Silver Chloride electrodes filled with Med Associates paste TD-246 were used for shock presentation and recording skin conductance. CS presentations were mixed into stereo headphones. Distracting tasks were written in Matlab (Mathworks) utilizing the Psychophysics Toolbox (Brainard, 1997). Analysis was carried out using programs written in Matlab as well as SPSS 10.

### 2.2.2 Subjects

Subjects were recruited from Caltech and were paid 20 dollars for their participation, based on informed consent. Their age ranged from 18-31 with a mean of 21 years. The following differential conditioning groups consisted of six subjects each: (*i*) delay no task, (*ii*) delay 1-back, (*iii*) delay 2-back (*iv*) trace no task, (*v*) trace 1-back, (*vi*) trace 2-back. The following single cue conditioning groups consisted of four subjects each: (*i*) delay no task, (i*i*) delay 2-back, (*iii*) uninformed trace no task, (*iv*) uninformed trace 0-back, (*v*) uninformed trace 2-back, (*vi*) informed trace no task, (*vii*) informed trace 0-back, (*viii*) informed trace 2-back.

### 2.2.3 Procedure

Skin conductance electrodes were attached to the palmar surface of the first and second fingers of the non-dominant hand. Shocking electrodes were attached to the palmar surface of the third and fourth fingers of the dominant hand. Each individual's shock level was determined using a subjective rating protocol that sought a level that was "uncomfortable but not painful". This shock level was used throughout the experiment.

After determining their shock level, subjects completed task training, the third of

three sessions of approximately five minutes each to ensure the subject had reached plateau performance. Prior to conditioning, subjects were read instructions asking them to focus on either their visual task or the wall in front of them. Naïve subjects had no previous specific knowledge of the experiment except that it was a "...learning and memory experiment that involves electric shocks." Subjects in the informed groups were read instructions that explicitly stated that an "electric shock shortly follows most presentations of a tone" and that "the tone generally predicts the occurrence of the electric shock." They were asked to confirm verbally that they "understand that the tone usually predicts the occurrence of the electric shock." Subjects were given a post-experimental questionnaire to assess their knowledge of the CS/US relationship (Clark and Squire, 1998) and were debriefed. The questionnaire for differential conditioning included 17 questions to assess the subject's explicit knowledge of stimulus relationships. Subjects were not allowed to correct previous answers. The awareness index is a number between 0 and 17, corresponding to the number of correct responses. The higher the index, the more detailed the subject's ability to recall the presence or absence of a contingency relationship between stimuli.

The informed consent procedure was reviewed and approved by the Caltech committee for the protection of human subjects.

### 2.2.4 Conditioning Stimuli (Figure 2-1A)

The US used in these experiments was a 0.25 second long, constant 60 Hz AC shock, the amplitude of which was determined by each subject. During differential conditioning, the CS+ and CS- were balanced between a 2 kHz tone (83 dB) and white noise (72 dB) and

were always 1 second in length. The 2 kHz tone was always used as the CS+ during single cue conditioning. During delay conditioning, reinforced CS+ presentations coterminated with the US. Reinforced CS+ presentations during trace conditioning were followed by a shock 4 seconds after the CS+ onset, leaving a 3 second trace period.

### 2.2.5 Experimental Phases (Figure 2-1B)

The learning procedure consisted of three phases: habituation, acquisition and extinction. In the first phase, habituation, subjects received two presentations of the CS+ and two of the CS-, in that order, to familiarize them with both stimuli. During acquisition, subjects received 24 CS+ and 24 CS- presentations, a total of 48 trials. Twenty of the 24 CS+ presentations were reinforced with a US, while four were not reinforced to allow for conditioning assessment. These four stimuli were positioned by randomly removing the US following one of the six CS+ presentations in each of the four blocks of 12 trials (six CS+, six CS-) during acquisition (excluding the first two CS+/US pairings in the experiment). During the extinction phase, subjects received twelve nonreinforced CS+ and twelve CS- presentations. CS+/CS- presentations occurred in random order with the limiting factors being a) that no more than two presentations of a specific CS occurred in a row and b) six of each occurred in each block of twelve trials. Intertrial intervals were uniformly distributed from 15-25 seconds.

Single cue conditioning experiments were performed in a similar fashion using a phantom CS-, a marked period of time that had no actual stimuli instead of an explicitly unpaired stimulus. The analysis protocol for single cue conditioning was analogous to the differential protocol using these phantom CSs-. When compared to a US only control

method, our procedure has the disadvantage of not controlling for unassociated stimulus SCR; however, it also has several advantages. It allows a comparison within subjects, a more effective means of detecting conditioning. This method also avoids the pitfalls of using a US only protocol where the US/CS- relationship is randomized or explicitly unpaired. The former may be associated with elevated CS- responses due to a generally elevated anxiety level. The latter tests the subject's ability to learn the anticorrelated relationship between the CS and US to enable suppression of the aforementioned general anxiety. It should be noted that our results show that working memory tasks interfere with our single cue trace conditioning protocol, adding validity to the idea that using the phantom CS- allows for accurate and sensitive detection of conditioning.

### 2.2.6 Distracting Tasks (Figure 2-1C)

To confirm that the conditioning protocols were effective, one group of subjects was excluded from performing a task (i.e. for each procedure they simply stared at the wall). The degree to which conditioning depends on working memory was assessed by asking a group of subjects to perform an n-back memory task during a conditioning procedure. Subjects had to press a key every time a given number appeared (0-back), when the present number matched the one before it (1-back), or whenever it was identical to the one before the previous (2-back). Only single cue trace subjects were asked to perform the 0-back task. The 0-back task involves the same input and the same motor output, including frequency of response, as the 1 and 2-back tasks, but is only minimally dependent on working memory.

The numeral 1, 2, 3 or 4 appeared at a constant rate that, for a 2-back task, was

adjusted for each subject to achieve a performance of approximately 85%. The mean rate of 2-back presentation was 1 Hz for differential subjects (88% correct), 1.33 Hz for uninformed single cue subjects (84% correct), and 1.2 Hz for informed single cue subjects (85% correct). All 1-back and 0-back tasks were performed at a presentation rate of 1.33 Hz. The mean performance for subjects focusing on the 1-back task was 93.5%. The mean task performance for single cue subjects in the 0-back group was 98% for uninformed subjects and 99% for informed subjects.

### 2.2.7 Analysis of SCR

A skin conductance response was measured as the maximal amplitude difference of more than 10 nS that occurred in a 1 to 4 second window after the delay CS onset, or in a 1 to 7 second window following the trace CS onset. Valid responses were range corrected by the largest amplitude response for each subject (Lykken, 1972). When there was no response, a zero-amplitude response was included in the analysis.

Habituation analysis was performed for differential conditioning using a paired t-test and a normalized ANOVA. No significant SCR differences were observed between the CSs, with one exception. Only the differential delay group performing no task showed an SCR difference ($p<0.05$) using the normalized ANOVA. However, no difference was observed using the paired t-test. The discrepancy between these statistical tests, the robustness of the conditioning for this group and the biased presentation order of the CSs lead us to regard this difference as inconsequential.

All CS+ presentations were compared to adjacent CS- presentations. During acquisition, when there were two adjacent CS- presentations available for comparison to a CS+, one was chosen at random.

Reported ‘p’ values for conditioning were ranked F-statistics for bootstrapped ANOVAs ($10^5$ re-samples per test). Four other tests were performed for confirmation: ranked F-statistics for a permuted ANOVA ($10^5$ re-samples); a square root corrected ANOVA; a permutation test (Efron and Tibshirani, 1998) ($10^5$ re-samples); and a paired t-test (averaging each trial across subjects). These confirmation statistics yielded similar results, with the exceptions noted below. Differential awareness correlations used a least squares fit. Analysis of main factors and interactions were performed using the GLM univariate ANOVA in SPSS (v10, Macintosh). These tests utilized the mean CS+, CS- difference for each subject.

### 2.2.8 Trial Effects

Trial effects were analyzed overall for acquisition and extinction phases of the experiment to assess the possible presence of consistent trends, such as a gamblers fallacy effect. Whether or not conditioning has occurred is assessed by comparing the results of the habituation analysis to the results of the acquisition/extinction phases of the experiment. In general, no CS+ / CS- (or phantom CS-) difference is present during habituation. There is a significant difference ($p<0.05$) between CS+ and CS- (or phantom CS-) responses during acquisition and extinction when conditioning has occurred. Learning is then assessed by the presence of this difference (reported in the results below).

## *2.3 Results*

### 2.3.1 Differential Conditioning

**No task** Differential conditioning relationships were first established for trace and delay paradigms, using six subjects per group who were not asked to perform any task during conditioning. The delay group (Figure 2-2A) shows larger SCRs to the CS+ test trials than to adjacent CS- presentations ($p<0.001$). The same is true of SCRs during trace conditioning (Figure 2-2B, $p<0.001$ paired t-test $p<0.01$). No significant trial effects are present in either group. Thus, trace and delay differential protocols are suffcient to produce conditioning when performed alone, without distraction.

**Concurrent distracting task** The n-back working memory task served as a distraction from the concurrently performed conditioning protocol. When six subjects performed the 1-back working memory task during differential delay conditioning (Figure 2-2C), there is a statistically significant difference between responses to CS+ and CS- during conditioning ($p<0.01$). However, when a 1-back working memory task was performed by six subjects during differential trace conditioning, there is no significant difference between SCRs to CS+ and SCRs to CS- (Figure 2-2D). When subjects carried out the 2-back task, there is no significant difference between responses to CS+ and CS- for either delay (n=6) or trace (n=6) conditioning (Figure 2-2E, F). No significant trial effects are present.

**Differential main effects** A univariate ANOVA using the mean CS+/CS- differences for each subject showed that both the delay/trace difference and task level were significant main effects ($p < 0.05$ and $p < 0.01$, respectively). The delay/trace by task interaction was not significant, but may have been lost in the floor effect between differential trace 1-back and differential trace 2-back.

### 2.3.2 Awareness of CS/US Contingency

**Correlations between awareness and CS+/CS- amplitude differences** There is a positive correlation between the awareness index and strength of conditioning (mean [CS+ - CS-]) during extinction for the 18 subjects carrying out the differential trace learning procedure (Figure 2-3). The correlation has an adjusted $r^2$ value of 0.334 (Pearson coeff. = 0.611, $p < 0.01$). No significant correlations between contingency awareness and CS+/CS- difference are present for trace acquisition, or for either acquisition or extinction during delay conditioning.

**Differential conditioning task interference** The twelve subjects who were not performing a task during differential conditioning (six delay, six trace) have an average awareness index of 15.2 (maximum 17). Twenty-four subjects who were performing a task during differential conditioning (trace and delay, 1-back and 2-back, six subjects in each combination of conditions) have an average index of 13.4. A univariate ANOVA utilizing the awareness questionnaire score to test factors that influence awareness show significant main effects for both task ($p<0.05$) and delay/trace (delay mean = 14.8, trace mean = 13.2, $p<0.05$) with no significant interaction. In summary, both the addition of a task and the addition of a short trace interval reduce the subject's ability to report the CS/US contingency relationship in a post-experimental questionnaire.

### 2.3.3 Single Cue Conditioning

**No task** Single cue conditioning relationships were established in a group of four delay subjects and four trace subjects who did not perform any distracting task during the conditioning protocol. Both groups (Figure 2-4A and B respectively, n=4 each) show

significant differences between CS+ test trials and adjacent phantom CS- presentations ($p<0.001$). No significant trial effects are present.

**Concurrent distracting task** A group of four single cue delay subjects and a group of four single cue trace subjects were asked to focus on the 2-back working memory task during conditioning (Figure 2-4C and D respectively). The subjects that carried out the 2-back task during single cue delay conditioning show greater SCRs to CS+ test trials than to phantom CS- trials ($p<0.001$). The 4 subjects performing the same 2-back task during a trace conditioning protocol show no significant conditioning for the experiment. No significant trial effects are present. While the 2-back task interferes with single cue trace and differential delay conditioning (Figure 2-2E), there is still a significant CS difference in single cue delay conditioning during the 2-back task.

**Uninformed 0-back task** A group of four subjects had to signal whenever a particular number appeared on the screen (0-back) during the single cue trace conditioning procedure (Figure 2-5A). There is no statistically significant difference between responses to the CS+ and the phantom CS- for this group. No significant trial effects are present. Although the 0-back task is a simple signal-detection task, there is no significant CS difference during single cue trace conditioning.

**Informed subjects** For the group of four informed subjects not distracted by any additional task (Figure 2-5B), and for the four performing the 0-back task (Figure 2-5C), there are significant differences between responses to the CS+ and the phantom CS- during single cue trace conditioning ($p<0.001$). However, for the group of four informed subjects performing the 2-back task (Figure 2-5D), there are no significant differences between responses to the CS+ and the phantom CS-. No significant trial effects are

present in any group. Prior explicit knowledge of the stimulus contingency facilitates, but does not guarantee, single cue trace conditioning.

## 2.4 Discussion

It is generally held in both eye-blink and fear conditioning that acquired trace and delay CS/US associations are distinct forms of learning. While the key difference between the two is the interposition of a temporal gap between the end of the CS and the start of the US, they involve different neural circuits and obey different regularities. For instance, acquisition of trace but not delay conditioning is critically dependent on hippocampus and certain prefrontal structures (Kim and Fanselow, 1992; Phillips and LeDoux, 1992; Maren et al., 1997; Weible et al., 2000; McLaughlin et al., 2002; Quinn et al., 2002) . In addition, Clark and Squire (Clark and Squire, 1998) showed that differential trace eye blink conditioning depends on CS/US contingency awareness, while this is not the case for delay conditioning (see also (Manns et al., 2000a; Clark et al., 2001; Manns et al., 2002)). This claim has been challenged. For example, Carrillo, Gabrieli and Disterhoft (Carrillo et al., 2000) demonstrated that not only single cue delay, but also single cue trace conditioning, was unaffected by division of attention. They used a dual-task paradigm to study the ability of subjects to acquire eye blink conditioning while their attention is concurrently engaged by watching a silent movie or verbal shadowing. Differential delay conditioning is, however, affected by the division of attention. Therefore, Carrillo and colleagues argue that the additional attentional demands imposed by the need to discriminate CS+ from CS- prevent delay conditioning from occurring when subjects have to perform a second task (see also (Mayer and Ross, 1969; Knuttinen

et al., 2001), and above results).

In this paper, we present experiments on fear conditioning. Fear conditioning differs from eye-blink conditioning in that it is dependent on the amygdala for both delay and trace conditioning, while eye-blink conditioning shows a similar pattern of dependence on the cerebellum (Medina et al., 2002). Our experimental paradigm involves association between tones or noises as CSs and electric shocks as USs. As a measure of autonomic conditioning, we utilize increases in skin conductance in a comparatively young population (college students). We choose fear conditioning since it can easily be adapted to rodents, allowing the use of molecular and genetic tools to study the underlying neuronal substrates of conditioning.

The general pattern of our findings is that the extent of associative autonomic conditioning depends on the cognitive load involved. The larger the demand on the system, the less conditioning occurs. We use the mean CS+, CS- difference for each group as a measure of strength of conditioning. This measure of conditioning is plotted in Figure 2-6 for each of our experiments. Figure 2-6 A, B, and C represent the transition from uninformed differential (A) to uninformed single cue (B, removing the second anticorrelated CS) and then the addition of explicit knowledge of the CS+/US relationship in the informed single cue condition (C). Task diffculty increases from left to right on the horizontal axis. The axis into the plane of the paper separates the trace and delay groups by the stimulus onset asynchrony (SOA) between the CS+ and US (Trace SOA = 4 sec, Delay SOA = 0.75 sec). Moving in Figure 2-6 from bottom to top (panel C to A), from right to left, or out of the plane of the paper all result in an increase in overall conditioning complexity for the subject. A decrease in conditioning with any difference

from the simplest protocol supports the hypothesis that as conditioning complexity increases, the amplitude/probability of conditioning decreases. This is reflected in a univariate ANOVA where the main effects single/differential, delay/trace, task level, and informed/uninformed effects are all significant. The only significant interaction is between single/differential and delay/trace. The lack of a significant delay/trace task effect could be due to a floor effect, because the conditioning amplitude has reached zero for trace conditioning protocols in the first level, where a concurrent task has been added. We are not making any claims about the uniqueness of this representation. Others are possible and might prove advantageous.

It should be noted that Figure 2-6 is compatible with the existence of secondary tasks that do not interfere with trace conditioning in naïve subjects. A similar plot might also prove beneficial in summarizing the eye blink conditioning literature.

In Figure 2-6, there are several interesting points to note. First, similar to results shown by others in eye-blink conditioning (Mayer and Ross, 1969; Carrillo et al., 2000; Knuttinen et al., 2001), differential delay conditioning is susceptible to interference tasks. Second, it should be noted that although reduced, single cue delay conditioning still occurred during the diffcult 2-back task. Third, all of the distracting tasks tested so far interfere with the trace fear conditioning protocol in our naïve subject pool. This is even the case for the 0-back task under single cue trace conditioning, a simple signal detection task—pressing a button whenever the target appeared in a string of numbers—with minimal attentional and working memory demands (subjects only had to remember a single target number during the 20 minute conditioning procedure). Fourth, it is only when we briefed subjects ahead of time about the nature of the experiment that we could

reliably induce trace conditioning under a 0-back task. We conjecture that this focused their attention onto the CS/US relationship and boosted learning.

The evaluation of the post-experimental questionnaire showed a correlation ($r^2$=0.395) between differential trace subjects' awareness scores and conditioning during the extinction phase. We found no significant correlation in the acquisition phase, nor did we find a correlation for either phase of delay conditioning. The correlation found establishes a link between explicit knowledge of the CS/US relationship and the expression of trace fear conditioning during extinction. It is different from the explicit knowledge/conditioning correlations reported in (Clark and Squire, 1998), because our correlation occurs in fear conditioning and is true for the extinction phase as opposed to acquisition. A challenge for the future will be to develop on-line measures of CS/US contingency awareness (LaBar and Disterhoft, 1998; Lovibond and Shanks, 2002).

One might expect that subjects who are aware of the stimulus contingency would show a gambler's fallacy effect where the differential response amplitude during extinction phase increases for a number of extinction trials. Such a pattern was reported during eye blink conditioning (Clark et al., 2001). We failed to find any significant trend in response slope. In fact, it is likely that if higher awareness scores cause stronger conditioning, this may lead to more than one response strategy (for example, higher initial responses with rapid extinction or gambler's fallacy). Our results also show a reduction in awareness in those groups who were performing a task compared to the no task controls.

Two possible non-exclusive explanations for our results are the following. One, explicit knowledge of the CS/US relationship is necessary for the expression of more

complex types of conditioning. When that explicit knowledge cannot be acquired, conditioning cannot be established. This is supported by the fact that task performance reduces both the awareness index and the efficacy of differential conditioning. In addition, explicit prior knowledge of the CS/US relationship compensates for some of the interference in single cue trace conditioning caused by concurrent task performance. Two, it is possible that concurrent task performance suppresses amygdala activity and subsequently suppresses the establishment of a conditioned fear response. Medial prefrontal cortex stimulation in rodents shows suppression of the basolateral complex of the amygdala (Rosenkranz and Grace, 2001). Furthermore, the n-back task shows an increased fMRI BOLD signal in human prefrontal areas that could be linked to suppression of normal brain activity under adverse conditions (Pochon et al., 2002). Either of these observations could explain fear conditioning interference by concurrent task performance.

We find it surprising that the working memory task has such a strong effect on both delay and trace conditioning. In the next chapter, we seek to identify areas of the brain that correlate with explicit and implicit learning during conditioning using fMRI. We hypothesize that the brain areas that correlate with explicit learning will be in the same prefrontal regions that are active during working memory tasks.

## 2.5 Figures and Legends

**Figure 2-1**

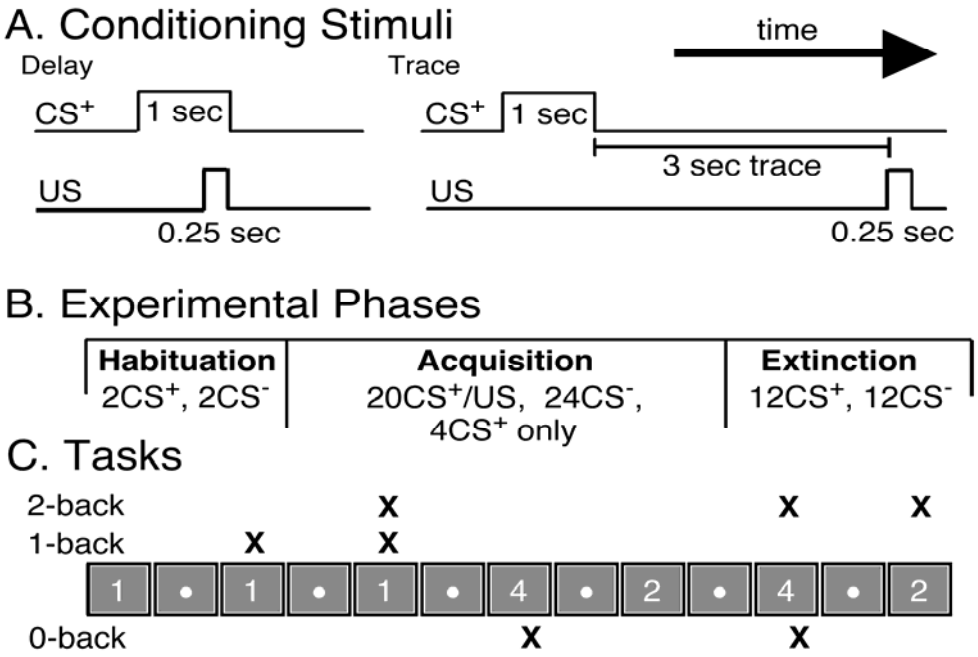



Figure 2-1 A) Delay conditioning consisted of a 0.25 second long electric shock that overlapped and co-terminated with the 1 second long CS+ (tone or noise). In trace conditioning, the CS+ was followed 3 second later by the US. B) The conditioning protocol consisted of three phases (habituation, acquisition and extinction). C) Distraction tasks and conditioning procedures were performed concurrently. During a 0-back task, the subject pressed a key (marked by an X) whenever a predetermined number appeared (4 in this case). During a 1 or 2-back task the subject pressed a key whenever the number matched the one before it or the one before the previous one, respectively.

**Figure 2-2**

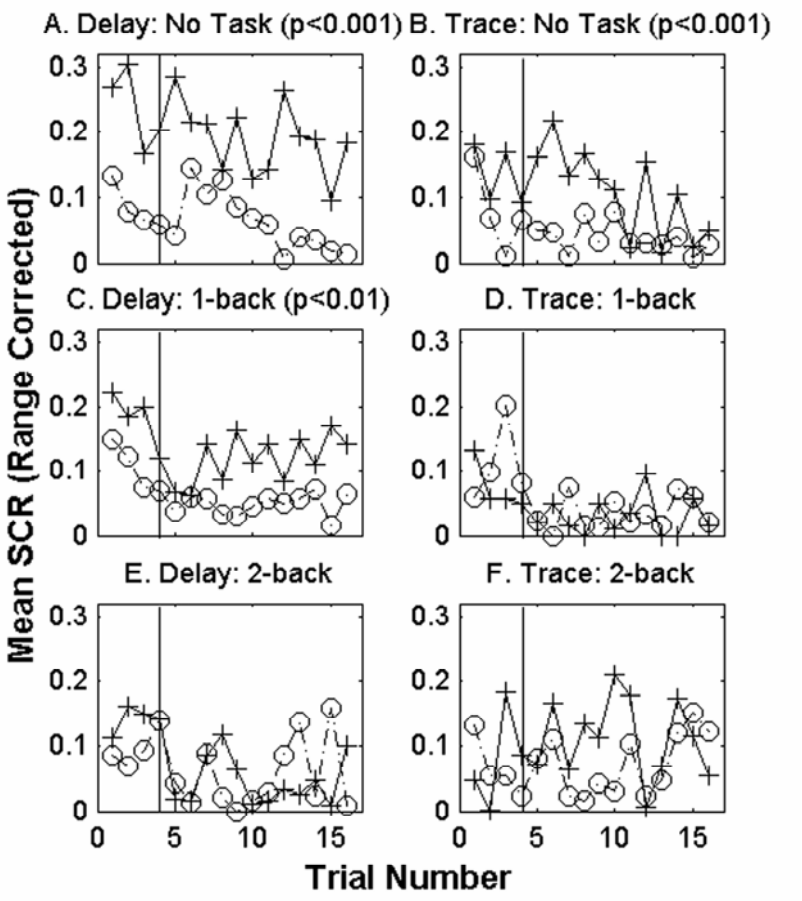



Figure 2-2 Mean range corrected SCRs to CS presentations for each trial. Thirty-six subjects (6 per group) participated in either the differential delay (A, C or E) or trace (B, D, or F) learning procedure without any task or while being distracted by a 1-back or a 2-back task. Mean range corrected SCRs to CS+ are shown in solid lines with cross markers. Mean range corrected SCRs to CS- are indicated by dashed lines with circles. Significant conditioning exists during the delay procedure with no concurrent task and while performing the 1-back task. Only under no task conditions did we find significant trace conditioning. The vertical line marks the last test trial during the acquisition phase.

**Figure 2-3**

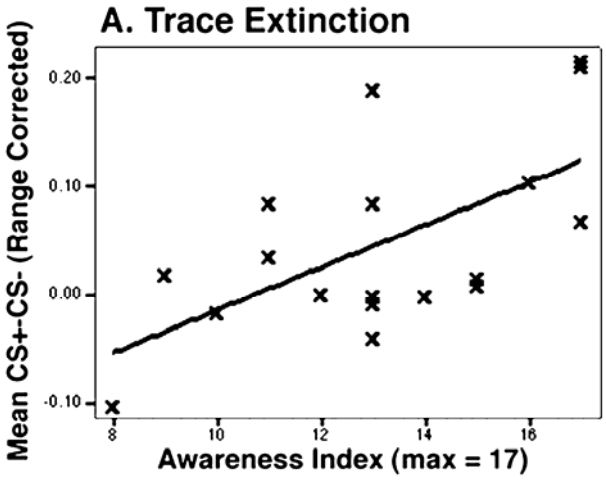



Figure 2-3 Scatter plot of mean range corrected differences between CS+ and CS- and the subject's awareness index. During differential trace extinction (Fig. 2B, D, and F; trial 5-16) subjects show a linearly increasing relationship between average amplitude of response difference and post-experimental questionnaire score (adjusted r2 = 0.334, Pearson coeff. = 0.611, $p < 0.01$, n= 18). Subjects show no significant correlation between conditioning (average range corrected CS+ -CS-) and awareness index during differential trace acquisition, differential delay acquisition or differential delay extinction.

**Figure 2-4**

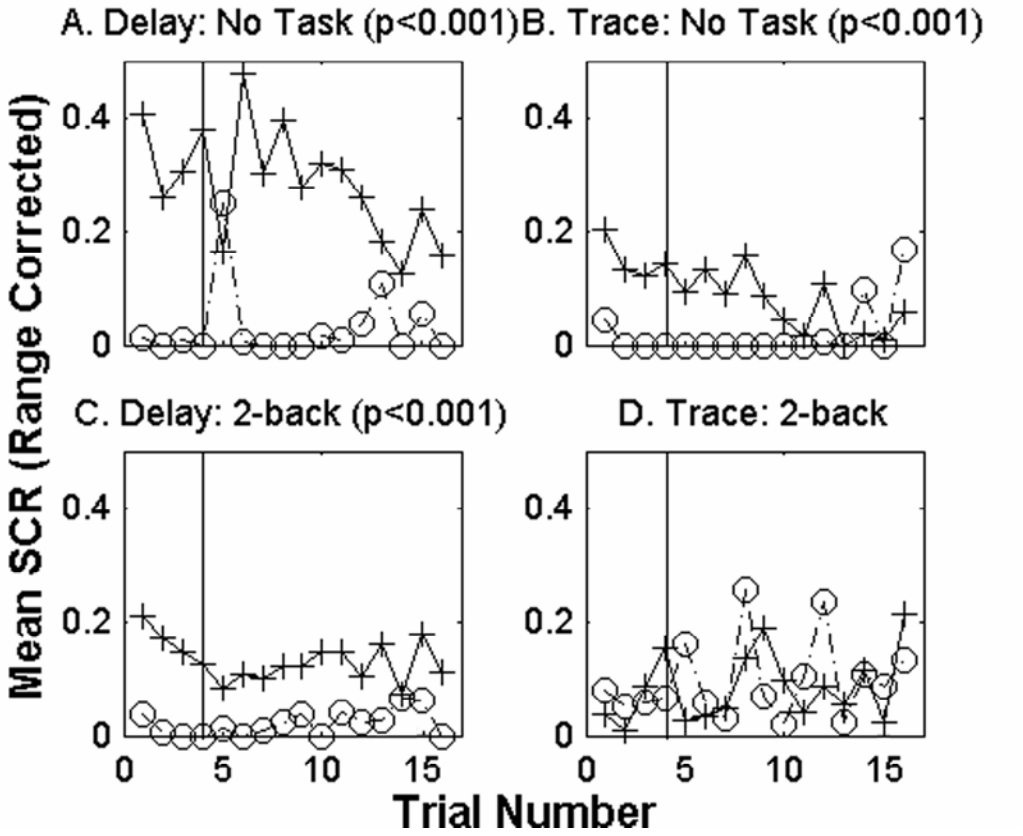



Figure 2-4 Mean range corrected SCRs to CS presentations for each trial. Sixteen subjects (4 per group) participated in either single cue delay (A or C) or trace (B or D) conditioning without any distraction or while carrying out a 2-back task. Mean range corrected SCRs to CS+ are shown in solid lines with cross markers. Mean range corrected SCRs to marked phantom CS- time points are indicated by dashed lines with circles. Significant conditioning exists for delay conditioning with no concurrent task and while performing the 2-back task. Significant trace conditioning is present only while no task was performed. The vertical line marks the last test trial presented during acquisition.

**Figure 2-5**

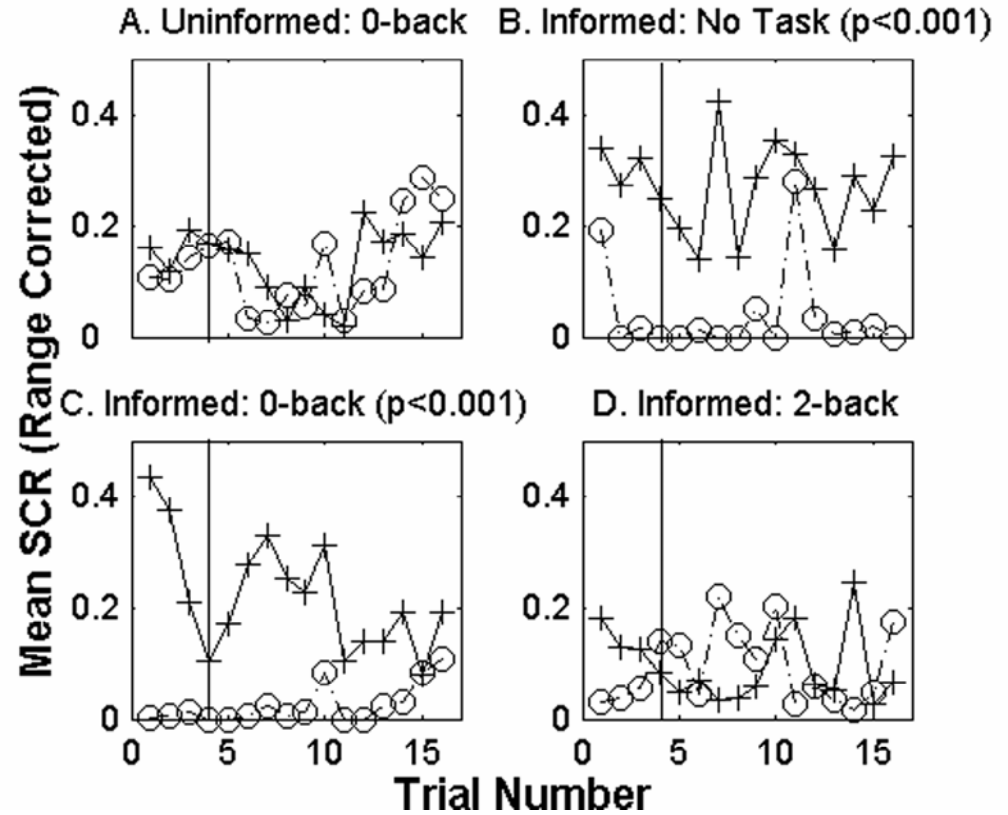



Figure 2-5 Mean range corrected SCRs to CS presentations for each trial. Sixteen subjects (4 per group) participated in either informed or uninformed single cue trace conditioning without being distracted (no task), or while carrying out a 0-back or a 2-back task. Mean range corrected SCRs to CS+ are shown in solid lines with cross markers. Mean range corrected SCRs to marked phantom CS- time points are indicated by dashed lines with circles. Significant conditioning is present for informed trace conditioning while subjects performed no task or a 0-back task. Significant uninformed trace conditioning is only present without a concurrent task (Figure 2-2B). The vertical line marks the last test trial presented during acquisition.

**Figure 2-6**

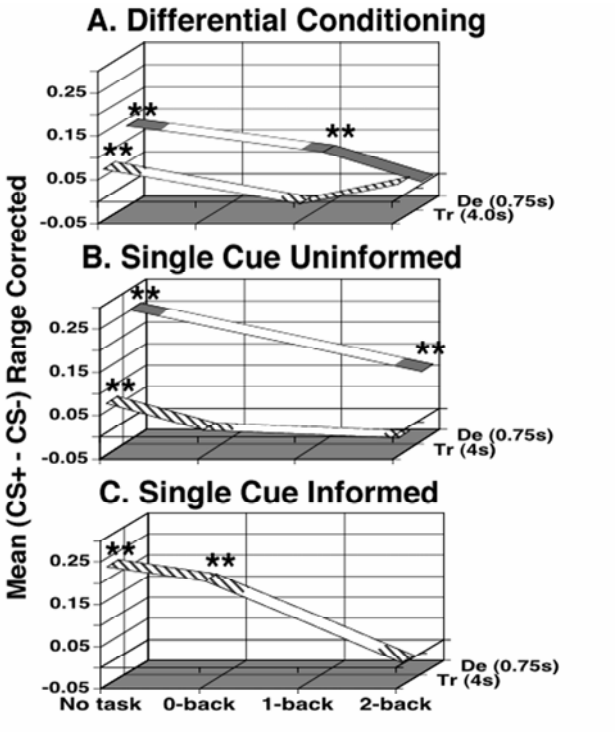



Figure 2-6 Summary of our data plotted in a 3-D space capturing the contingencies of our protocol. The vertical axis marks the group average for each subject's average range corrected and normalized CS+, CS- difference. The horizontal axis marks the task difficulty. The axis into the plane of the paper marks the group as trace or delay using the difference in CS/US onset (SOA) in seconds. In addition, the line for trace is hatched while the line for the delay group is solid. "**" indicates significant conditioning at $p<0.01$. Areas of the lines that are not filled in are meant to assist the stability of the figure, not to imply any prediction about the magnitude of conditioning in that area. A) Mean group differences for differential subjects. B) Mean group differences for uninformed single cue subjects. C) Mean group differences for single cue informed subjects. Our results indicate that the higher the cognitive load, the smaller the CS+/CS- difference.

# 3 The Neural Correlates of Implicit and Explicit Processes in Conditioning

Portions of this work were published as "Contingency Awareness in Human Aversive Conditioning Involves the Middle Frontal Gyrus" in 2006 (Carter et al., 2006). The authors were: Ronald McKell Carter, John P. O'Doherty, Ben Seymour, Christof Koch and Raymond J. Dolan. John O'Doherty served as a very knowledgeable teacher; he was involved in the experimental design, fMRI sequence choice and all data collection. He introduced me to SPM and provided base code to make the analysis easier, then continued to follow up with ideas to improve our analysis. He also helped arrange my visit to UCL and was a part of extensive manuscript review. Ben Seymour aided in the experimental design, was a great help in getting the electrical stimulation equipment, stimulation code and IRB protocols in place, helped in preparing subjects and data collection, and completed extensive manuscript reviews. He was also the artist behind the abstract images we used. Christof Koch and Raymond J. Dolan not only found funding to make the project possible and initiated the collaboration, but were also involved in the design of the experimental procedure and statistical analysis as well as extensive manuscript review. Experiments and some of the analysis were conducted at the Functional Imaging Laboratories at University College London.

We used functional magnetic resonance imaging to track the trial by trial acquisition of explicit and implicit knowledge in a concurrent trace and delay conditioning paradigm.

## 3.1 Introduction

Learning about aversive stimuli in the environment is necessary for an organism's success. One of the simplest and best studied mechanisms by which this is realized is classical conditioning, whereby a predictive association is learned between a neutral stimulus (the conditioned stimulus or CS) and a biologically meaningful signal (the unconditioned stimulus or US) (Pavlov, 1906; Cook and Harris, 1937; Wilensky et al., 1999; Buchel and Dolan, 2000; Maren, 2001; Clark et al., 2002; LeDoux, 2003; Maren and Quirk, 2004). Typically, after repeated pairings of CS and US, the CS comes to elicit a response that is appropriate to the anticipated US. In aversive conditioning, this conditioned response (CR) will often be a change in heart rate or skin conductance, and is taken as an implicit measure of successful conditioning in experimental studies. However, it is also possible to become consciously aware of the predictive contingency between CS and US, a phenomenon referred to as contingency awareness. An individual can acquire both implicit associations and contingency awareness or either may be acquired independently (Bechara et al., 1995), indicating some degree of dissociation between the two systems. Currently, a major question in conditioning (and consciousness) research is the extent, and mechanism, of contingency awareness effects in conditioning (Hilgard et al., 1937; Cole, 1939; Dawson and Furedy, 1976; Lovibond and Shanks, 2002; Wiens and Ohman, 2002; Olsson and Phelps, 2004). A better understanding of contingency awareness and how it can facilitate or inhibit implicit associations is critical for a rational treatment of phobias, placebo effects and anxiety disorders (Grillon, 2002; Quirk and Gehlert, 2003; Colloca and Benedetti, 2005).

The acquisition of contingency awareness and its interaction with conditioning differs across conditioning protocols (Clark and Squire, 1998; Ohman and Soares, 1998; Knuttinen et al., 2001; Han et al., 2003). For instance, in trace conditioning, in which there is temporal separation between CS and US, contingency awareness has been shown to positively correlate with the amplitude of conditioned responses (Clark and Squire, 1998). Those subjects in a trace conditioning experiment who do not display contingency knowledge fail to be trace conditioned. By contrast, in delay conditioning, in which there is no separation between the CS and US, no correlation between contingency awareness and successful conditioning has been observed; either can be acquired in the absence of the other. However, the simplicity of the delay protocol often results in immediate acquisition of explicit knowledge, making separation of explicit from implicit processes difficult using a delay paradigm alone.

In this study, we used functional magnetic resonance imaging (fMRI) to identify brain regions that were specifically related to the explicit acquisition of contingency awareness during both delay and trace conditioning, independent of individual protocols. We simultaneously conditioned human subjects to predict an aversive electrical stimulus (US) from arbitrary visual cues (CS) with concurrent delay and trace protocols (see Figure 3-1a/b). The use of simultaneous conditioning allowed us to identify brain responses specifically correlated with contingency awareness and distinct from responses associated with measures of implicit knowledge. To assess contingency awareness, subjects reported their shock expectancy on each trial (Figure 3-1c/d), and in addition filled out a post-experimental questionnaire. These measures were then used to identify brain responses that correlated with accurate contingency awareness. We predicted that

activity in dorsolateral prefrontal cortex and hippocampus would correlate with these measures of explicit knowledge based on evidence that these structures are involved in working memory (Leung et al., 2002), memory formation (Fanselow, 2000) and revaluation (Corlett et al., 2004), as well as from lesion studies of trace conditioning deficits (Compton et al., 1997; Clark and Squire, 1998; Kronforst-Collins and Disterhoft, 1998; McEchron et al., 1998). In addition, we identified those regions that correlated with an implicit measure of learning, differential skin conductance responses. Consistent with previous work, we found that the amygdala correlated with implicit learning as measured by skin conductance changes. Surprisingly, changes in skin conductance were also a good predictor of activity in visual cortex and the hippocampus.

## *3.2 Methods*

### 3.2.1 Participants

We recruited sixteen healthy right-handed subjects. Two were excluded: one because of excessive movement-related artifact precluding image analysis, and another subject who did not have at least one significant skin conductance response (SCR) for each trial type, precluding study of the time course of learning. The remaining subjects are reported in the analysis: 9 male and 5 female, age range 19-31 (mean 24.7). All subjects gave prior informed consent. This study was approved by the Joint Ethics Committee of the National Hospital for Neurology and Neurosurgery, UK (UCLH NHS Trust) and the Human Subject Committee at the California Institute of Technology, USA.

### 3.2.2 Experimental Procedure

We performed concurrent trace and delay Pavlovian conditioning. The CSs were abstract colored images (see Figure 3-1a) presented for 2 seconds and the US was a 1second electrical stimulus (see below). The study comprised 160 individual trials involving four separate CSs (each presented 40 times). One of the images acted as the trace conditioning cue (trace CS), which was followed on 50% of occasions by the US after a 3 second trace interval. Another image acted as the delay conditioning cue (delay CS), followed on 50% of occasions by the US, with a 0.5 second overlap between the end of the CS and the start of the US. The remaining two images acted as neutral cues (CS-), never followed by the US. Images were counter-balanced across conditions between subjects. Presentations of the CSs were arranged randomly, such that two of each CS type appeared in a block of eight. The delay and trace CS were each reinforced once in every block of eight trials. Trials were triggered on the nearest slice using a pseudo-randomized inter-trial onset asynchrony of 8, 9.25, 10.5, or 12 seconds. Presentation of stimuli and timing were controlled using Cogent 2000 (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK).

### 3.2.3 Online Subject Reports of Contingency Awareness

Subjects reported US (electric shock) expectancy for each trial by pressing one of three keys with their right hand. This was performed as quickly as possible following the presentation of each CS. One key indicated that a shock was expected, the second key indicated that the subject didn't know whether or not a shock was expected, and the third indicated that no shock was expected. Prior to conditioning, subjects practiced the procedure on a set of abstract images not used in the experiment. At no point before the

experiment were participants explicitly informed about any relationship between the images (CSs) and shock (US). Failures to respond or responses where latencies exceeded 1.5 seconds were scored as “don’t know”.

### 3.2.4 Post-Experimental Questionnaire

Following scanning, subjects were given a post-experimental questionnaire similar to that used by Clark and Squire (Clark and Squire, 1998). The questionnaire assessed their knowledge of the CS/US contingency relationships for both delay and trace protocols (see supplementary material). Subjects rated each statement on a 7 point scale ranging from “not true” through “don’t know” to “true”, capturing their degree of confidence. A response that was both accurate and very confident received a score of +3 and a response that was inaccurate and very confident received a score of -3, with all other responses falling on a scale between these limits. Scores for each subject were then totaled, giving a potential range of -48 to +48 for each protocol. More positive scores reflect greater contingency awareness.

### 3.2.5 Unconditioned Stimuli

The pain specific shock was delivered to the top of the right foot using a 100Hz train of square-waveform electrical pulses for 1 second, via a bipolar concentric surface electrode (stimulation area 20 $mm^2$), which selectively depolarizes A delta fibers (Kaube et al., 2000). This custom built concentric electrode was designed to limit activation to fast acting fibers and reduce any possibility of muscle stimulation. The electrical stimulus was delivered via an optically isolated unit with a range of 0-12mA. Current levels were

chosen for each subject before the experiment, starting at a low level and using an ascending rating method where the current amplitude was raised until the subject gave a rating of 9 on a scale of 1-10, where 1 indicated the subject could barely feel the shock and 10 indicated the shock was too uncomfortable to be used in the experiment.

### 3.2.6 Online Measure of SCR

Skin conductance data was collected at a minimum of 100Hz, and was aligned to the first slice pulse where scanning had started. Data collected at a rate higher than 100Hz was first down-sampled to that frequency. Before analysis, all skin conductance data was median filtered to reduce noise. Skin conductance responses (SCRs) were defined as the maximum amplitude response initiated no earlier than 1second with a peak no later than 5seconds after the CS onset. SCR amplitudes were range corrected by the maximum response for that subject (Lykken, 1972). A two-tailed, single sample t-test across subjects (n=14) showed a significant difference between the mean nonreinforced CS+ response and the mean CS- response for both delay ($P<0.01$) and trace ($P<0.01$) conditioning.

### 3.2.7 fMRI

Forty-four slice whole brain tilted axial BOLD images were acquired in a 3 Tesla Siemens Allegra scanner using a gradient-echo EPI sequence (Deichmann et al., 2003), at a within plane resolution of 3mm (TR = 2.86secs). A total of 565 images were acquired, including five saturation scans which were later discarded. After the completion of the experiment, a T1 weighted anatomical scan was obtained for each subject.

Functional image analysis was performed using SPM2 (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK). Prior to analysis, all functional images were realigned, slice time acquisition corrected, normalized to the MNI EPI template and smoothed using a Gaussian kernel (8mm FWHM including contrast image smoothing). Individual subject models were then constructed and random effects analysis conducted as noted below. Each trial was modeled in two segments: first, an initial event-related response to each CS presentation, and second, a 3second period following image termination. This was done for both trace and delay conditioning to ensure that correlations between the stimulus event and trace period models were treated similarly. Trace period regressors were orthogonalized with respect to their event-related CS onset equivalents to minimize any contamination of the trace period response by the CS onset response. Only results for the CS onset event responses are reported here.

The acquisitions of both conditioning and contingency awareness were modeled as a parametric modulation (Buchel et al., 1998a) of responses to a CS presentation (indexed by skin conductance and US expectancy respectively). In testing for brain activity that correlated with learning, we examined BOLD responses specific to the CS by performing a conjunction analysis (Friston et al., 1999). This technique identifies only those regions that show significant activation across the included conditions. Statistically, a conjunction analysis is identical to performing an F test with the constraint that the individual effects are positive. For all reported results, we identified regions whose activity reflected learning for both delay *and* trace conditioning during all CS+ trials (reinforced and nonreinforced). This provides a general measure of learned response differences specific to a CS presentation, leaving aside effects due to the presence or

absence of a shock and those due to peculiarities of the specific conditioning protocol. A region is reported as active if it violates the null hypothesis that on average members of the conjunction showed no effect (global null). An identified region's consistent activation for all members of the conjunction is confirmed by looking at the least significant P value for any member of that conjunction. We include plots of correlations in Figure 3-3 to demonstrate the consistency of a given effect across conditions included in the conjunction. This procedure lessens response ambiguity due to either the presence or absence of a US or any differences in protocol. Using a conjunction across these conditions allows us to infer a network that relates to the overall learned differences between CS+ (delay and trace) and CS-(neutral) representations. In areas where there was a prior hypothesis, results were family wise error (FWE) corrected for multiple comparisons using small volume correction (20mm diameter sphere centered at the peak of activation). FWE correction for multiple comparisons for the whole brain is applied for brain regions where there was no prior hypothesis.

### 3.2.8 Learning: Contingency awareness and Conditioning

*Explicit learning accuracy* was defined by the interaction between CS type (CS+ or CS-) and the reported US expectancy for each trial. Brain activity that correlates with US expectancy alone corresponds to those areas relevant for explicit fear. A shock expectancy by CS type (+ or -) interaction tests for brain activity that correlates with the accuracy of shock expectancy on each trial. The *magnitude of explicit learning* defined by the subject's score on the post-experimental questionnaire was used as a subject covariate in a second level random effects analysis (Figure 3-3 and Table 3-1b).

*Implicit learning accuracy* was defined by the interaction between CS type (+ or -) and the normalized amplitude of the skin conductance response for each trial. The *magnitude of implicit learning* was defined as the average difference between CS+ and CS- skin conductance responses and was used as a subject covariate in a second level random effects analysis (Table 3-1a).

## *3.3 Results*

### 3.3.1 Conditioned skin conductance responses

We recorded skin conductance responses associated with cues to provide an implicit measure of conditioning. Activity in the left amygdala (-27,-3,-12) correlated with the trial-by-trial time course of conditioning, indexed by the level of discriminatory skin conductance responses ($P<0.01$ corrected, see methods and Table 3-1a). This result confirms previous findings (Buchel et al., 1998b; Buchel et al., 1999; Knight et al., 2004) and in addition demonstrates that amygdala activity correlates with the specific time course of learning. It also indicates that activity in the amygdala correlates with the relative success of conditioning in different subjects.

### 3.3.2 Contingency awareness

Our principle goal was to identify neural responses correlating with contingency awareness, an example of explicit or declarative learning, measured by both online reports and post-experimental questionnaire. Online reports assessed US expectancy on a trial-by-trial basis: expectancy was accurate when it was high for a CS+ (predicts the shock) presentation and low for a CS- (neutral) presentation. This relationship is

described as the *accuracy of contingency awareness* (Figure 3-1d). The *magnitude* of a subject's contingency awareness was defined by a post-experimental questionnaire score that assessed the individual's overall contingency knowledge via a series of true/false statements about the CS/US relationship (see supplementary methods – section 3.7 ).

Brain regions that correlated with contingency awareness had greater activity during trials where a subject accurately expected a shock. We accounted for inter-subject differences by testing for regions that showed greater activity in those subjects who scored higher on the post-experimental questionnaire. This revealed responses correlated with contingency awareness in bilateral middle frontal gyri (MFG) (Figure 3-3, Table 3-1b, left -36,51,30; right 36,51,36), significant after correction for multiple comparisons. We also noted correlated activity in the para-hippocampal gyrus (-15,-15,-24, P=0.055 corrected for multiple comparisons, see methods).

## *3.4 Discussion*

Our data indicate a clear role for the middle frontal gyrus in contingency awareness during conditioning, correlated specifically with the acquisition of awareness on a trial-by-trial basis. To our knowledge, this is the first time such a trial by trial link has been demonstrated during conditioning. The role of the middle frontal gyrus in contingency awareness is contrasted with involvement of the amygdala, which we show reflects the acquisition of implicit knowledge, as indexed by autonomic activity, consistent with previous research (Buchel et al., 1998b; Buchel et al., 1999; LeDoux, 2003). These results clearly dissociate the distinct roles of the middle frontal gyrus and amygdala during classical conditioning.

In delay eye-blink conditioning, the magnitude of conditioning is independent of explicit knowledge (Manns et al., 2002). However, explicit learning is likely to be expressed in delay protocols, even if it is not correlated with the degree of conditioning. Since the degree of implicit knowledge is not correlated with the degree of explicit knowledge in delay conditioning, the substrates mediating both forms of learning can be separated. We confirmed that those same neural substrates were active during trace conditioning by testing for areas whose activation was consistent across both delay and trace conditioning. The fact that the middle frontal gyrus is active in both trace and delay conditioning, even though trace conditioning is correlated with contingency knowledge and delay is not (Clark and Squire, 1998), has implications for the mechanism by which contingency knowledge facilitates conditioning. The middle frontal gyrus is unlikely to directly facilitate conditioned associations, since doing so would require a second inhibitory mechanism active during delay conditioning. It is therefore more likely that the middle frontal gyrus facilitates conditioning by means of another brain area, such as the hippocampal complex (see below). While it is unlikely that prefrontal areas directly facilitate conditioning, it is important to keep in mind that there is evidence that areas of prefrontal cortex inhibit activity in the amygdala (Rosenkranz and Grace, 2001; Quirk et al., 2003).

An area homologous to the middle frontal gyrus, the medial prefrontal cortex in the rabbit, is necessary for trace eye-blink conditioning (Kronforst-Collins and Disterhoft, 1998). This region is also strongly implicated in tasks requiring maintenance and manipulation of information within working memory in humans (D'Esposito et al., 1998; Leung et al., 2002; Pessoa et al., 2002), and in animal models of working memory

(Goldman-Rakic, 1987; Petrides, 2000; Castner et al., 2004). In the previous chapter, we showed that working memory distraction during fear conditioning reduces explicit knowledge of the CS/US contingency (Carter et al., 2003). This reduction of explicit knowledge is consistent with the middle frontal gyri's involvement in both working memory and contingency awareness.

Brain areas central to the expression of explicit knowledge, as required in reporting contingencies, may play a role in abstract, symbolic manipulation. In line with this notion, neurons in the middle frontal gyrus of behaving macaque monkeys respond to specific rules (Wallis et al., 2001) or limit responses to a given stimulus to only those times when a specific practiced task is being performed (Asaad et al., 2000). Thus, our finding that activity in this region correlates with contingency awareness is consistent with a putative role in the representation of abstract concepts.

We also found activity in the left para-hippocampal gyrus correlated with contingency awareness during conditioning. These results point to a role for the hippocampal complex in mediating the integration of explicit knowledge of contingencies (Eichenbaum et al., 1996; Clark and Squire, 1998). It is interesting that while we found significant contingency related activation in the para-hippocampal gyrus, we did not find such effects in the hippocampus proper. By contrast, we observed a significant correlation between activity in the hippocampus proper and our implicit measure of conditioning. These results do suggest the intriguing possibility that different sub-regions of hippocampal complex have dissociable roles in associative learning, in line with evidence that the hippocampus is involved in the integration of cues and not simply

related to explicit knowledge (Chun and Phelps, 1999; Schendan et al., 2003; Degonda et al., 2005). Future studies will need to address the mechanism of integration.

In conclusion, our study provides new evidence that the trial by trial accuracy of contingency knowledge during conditioning involves the middle frontal gyri and a sub-region of the hippocampal complex across both delay and trace conditioning. These findings give insight not only into the neural substrates of classical trace protocols, where explicit knowledge correlates with conditioning, but also suggest a substrate for how explicit knowledge is coded in the human brain.

## *3.5 Tables*

**Table 3-1**

| Cluster Region | MNI Coord. | Voxels | P (L.S.) | P (G.N.) |
|---|---|---|---|---|
| **(a) Implicit learning** | | | | |
| Left Hippocampus / Subiculum | (-12,-30,-6) | 364 | 0.05 | P<0.01* |
| Right Hippocampus / Subiculum | (21,-27,-12) | | 0.04 | P<0.01** |
| Occipital Cortex (posterior pole) | (9,-102,-9) | 34 | 0.02 | P<0.01** |
| Left Amygdala | (-27,-3,-12) | 11 | 0.06 | P<0.01* |
| **(b) Explicit Learning** | | | | |
| Left Middle Frontal Gyrus | (-36,51,30) | 35 | 0.02 | P<0.001* |
| Right Middle Frontal Gyrus | (36,51,36) | 22 | 0.04 | P<0.01* |
| Left Parahippocampal Gyrus | (-15,-15,-24) | 3 | 0.096 | P=0.55 |

* FWE corrected for small volume (20mm diameter sphere)

** FWE corrected for whole brain

Table 3-1 Brain regions whose BOLD responses correlate with implicit and explicit measures of learning are shown in (**a**) and (**b**). This table specifies the anatomical labels for responsive clusters with the location of peak significance (mm in MNI space), the number of voxels included in the cluster (threshold = P<0.001) and the global null (effects of interest) P value for the peak voxel in each cluster (G.N.). To ensure the consistency of correlation across conditions, the least significant individual P value (L.S.) is also reported.

**Table 3-2**

| Cluster Region | MNI Coord. | Voxels | P (L.S.) | P (G.N.) |
|---|---|---|---|---|
| **(a) CS onset event responses (CS+ > CS-)** | | | | |
| Left Frontal Operculum<br>Left Anterior Insula<br>Left Inf. Frontal Gyrus | (-48,12,3)<br>(-33,21,6)<br>(-57,9,12) | 241 | 0.03<br>0.003<br>0.02 | P<0.05**<br>P<0.0001**<br>P<0.05** |
| Right Frontal Operculum<br>Right Anterior Insula<br>Right Inf. Frontal Gyrus | (48,21,6)<br>(42,21,-3)<br>(51,18,18) | 323 | 0.002<br>0.002<br>0.002 | P<0.001**<br>P<0.05**<br>P<0.01* |
| Left Caudate<br>Right Caudate | (-12,9,3)<br>(9,6,3) | 152 | 0.01<br>0.02 | P<0.01**<br>P<0.01** |
| Supplementary Motor Area | (3,21,60) | 145 | 0.01 | P<0.05** |
| Left Middle Frontal Gyrus (DLPFC)<br>Right Middle Frontal Gyrus (DLPFC) | (-33,45,18)<br>(33,48,15) | 50<br>49 | 0.03<br>0.06 | P<0.001*<br>P<0.01* |
| Left Anterior Cingulate<br>Right Anterior Cingulate | (-9,42,30)<br>(6,30,24) | 9<br>8 | 0.08<br>0.007 | P<0.01*<br>P<0.01* |
| Left Supramarg. Gyr. / IPL<br>Right Supramarginal Gyrus | (-60,-48,24)<br>(66,-36,33) | 209<br>37 | 0.04<br>0.06 | P<0.05**<br>P<0.01* |
| **(b) Trace Period Responses (CS+ > CS-)** | | | | |
| Left Frontal Operculum<br>Left Frontal Insula<br>Left Inf. Frontal Gyr | (-42,15,6)<br>(-36,21,27)<br>(-60,18,3) | 337 | 0.01<br>0.001<br>0.008 | P<0.01*<br>P<0.05**<br>P<0.01* |
| Right Frontal Operculum<br>Right Frontal Insula<br>Right Inf. Frontal Gyr | (51,18,6)<br>(36,15,-9)<br>(51,18,18) | 266 | 0.001<br>0.003<br>0.02 | P<0.05**<br>P<0.001*<br>P<0.05* |
| Supplementary Motor Area | (-12,-12,69) | 15 | 0.005 | P<0.01* |
| Left Middle Frontal Gyrus (DLPFC)<br>Right Middle Frontal Gyrus (DLPFC) | (-45,54,6)<br>(33,60,21) | 6<br>34 | 0.01<br>0.003 | P<0.01*<br>P<0.001* |
| Right Anterior Cingulate | (12,24,36) | 34 | 0.006 | P<0.01* |
| Left Supramarg. Gyr. /IPL<br>Right Supramarg. Gyr. / IPL | (-48,-48,33)<br>(54,-42,42) | 90<br>564 | 0.005<br>0.00005 | P<0.01*<br>P<0.0001** |

* FWE corrected for small volume (20mm diameter sphere)
** FWE corrected for whole brain

Table 3-2- Brain regions whose mean BOLD responses are greater for the CS+ than CS- during (**a**) the transient stimulus onset (across delay and trace conditioning and reinforced and nonreinforced trials), and (**b**) the 3 second long trace period (see Figure 3-1) for trace conditioning trials (reinforced and nonreinforced). This table specifies the anatomical labels for responsive clusters with the location of peak significance (mm in MNI space), the number of voxels included in the cluster (threshold = P<0.001), the least significant individual P value in the conjunction (L.S.) and the global null P value for the peak voxel in each cluster (G.N.).

## 3.6 Figures

**Figure 3-1**

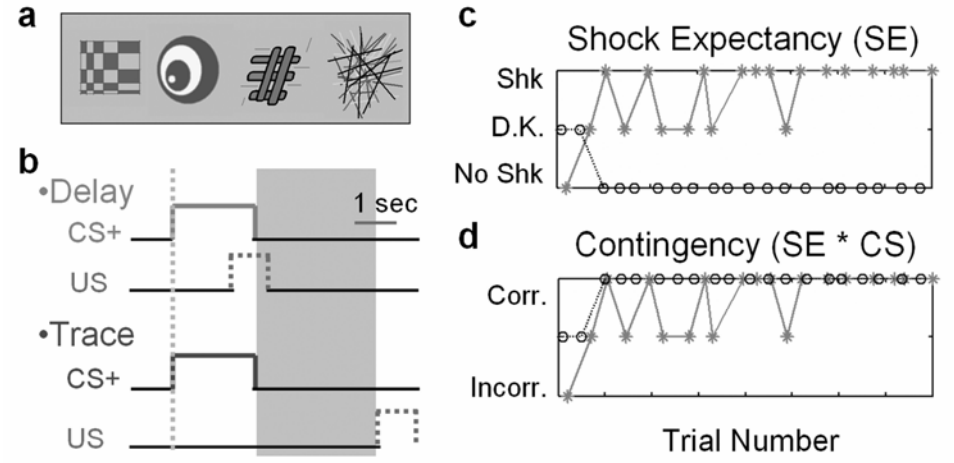



Figure 3-1 Experimental design – 14 subjects reported shock expectancy during concurrent delay and trace aversive conditioning while functional brain images were acquired. Each CS presentation was 2 seconds long. Reinforced CS+ trials were followed by an electric shock lasting 1 second. For each subject, the four images in (**a,** presented in color) acted as either the delay CS+ (image presentation overlaps with the shock, see **b**), trace CS+ (image presentation ends 3 s before the onset of the shock, see **b**), or as one of two CS- baselines. The point used for modeling event related analysis for each trial is marked with a dotted line in **b**. The memory trace period, 3 seconds marked by a shaded box (also in **b**), was analyzed separately and is not discussed here. As soon as each image appeared, subjects had to judge the likelihood of being shocked (shock expectancy; **c**) using one of three keys indicating "shock likely," "don't know," or "no shock likely" (* = CS+; o = CS-). **d**. The accuracy of contingency awareness is the interaction between shock expectancy (SE) and CS+/- (* = CS+; o = CS-). This measure reflects when a subject provided an accurate report of the CS/US contingency over the course of the experiment, and defines the measure of *accuracy of explicit knowledge*.

**Figure 3-2**

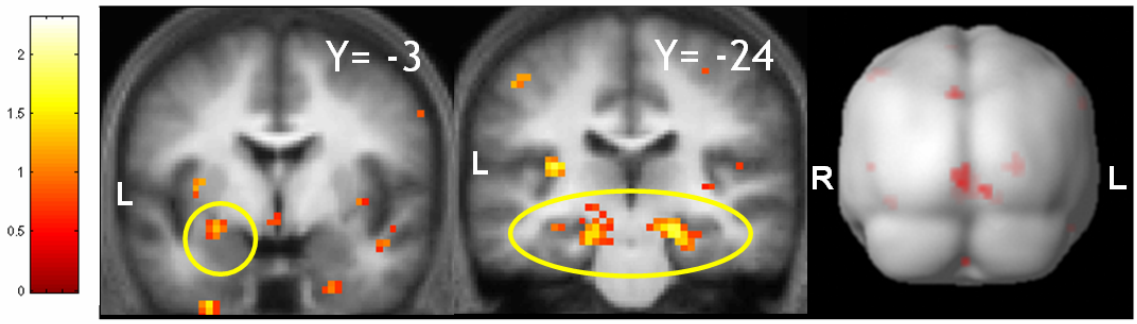



Figure 3-2 – Implicit measures of learning correlate with activity in the left amygdala (left), hippocampus (center), and visual cortex(right). See Table 3-1 for details. Images are shown at a threshold of $p<0.01$ uncorrected in order to visualize the extent of activation.

**Figure 3-3**

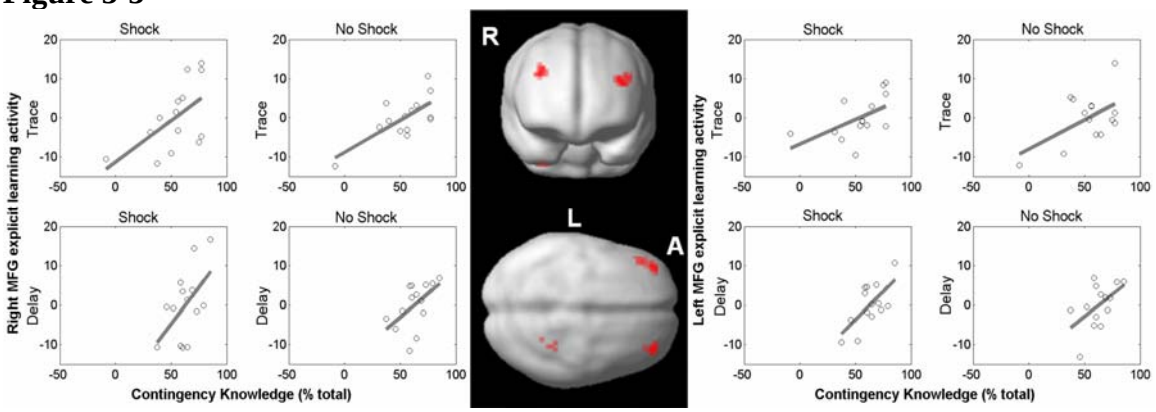



Figure 3-3 Activity in the middle frontal gyrus correlates with explicit learning. The center image shows a surface rendering of bilateral middle frontal gyri activity that correlates with explicit learning. Middle frontal gyrus activity is consistently correlated with explicit learning for all conditions, reflecting its general role in the acquisition of accurate explicit knowledge. Regression plots of brain activity (the explicit learning time course parameter estimate) vs. contingency knowledge (post experimental questionnaire score) are shown for the right and left middle frontal gyrus for trace and delay protocols during shock reinforced CS+, and no shock CS+ trials. A contingency knowledge score of 100% indicates that the subject was able to answer every question about the exact temporal relationship between CS and US accurately with high confidence. 0% is chance performance.

## 3.7 Supplementary Methods

Post Experimental Questionnaire

-Answered from -3 (not true), 0 (don't know), 3 (true)
-Presented in pseudo random order within block
-24 total questions for overall assessment
-16 questions included for delay assessment (*)
-16 questions included for trace assessment (#)

+General Contingency Questions (8)
[*#] I believe the colored sticks were not followed by the shock.:
[*#] I believe the blue ball was not followed by the shock.:
[*#] I believe the red and green hash was not followed by the shock.:
[*#] I believe the blue and orange check was not followed by the shock.:
[*#] I believe the colored sticks were often followed by the shock.:
[*#] I believe the blue ball was often followed by the shock.:
[*#] I believe the red and green hash was often followed by the shock.:
[*#] I believe the blue and orange check was often followed by the shock.:

+General Timing Questions (8)
[*] I believe the colored sticks were often immediately followed by a shock.:
[*] I believe the blue ball was often immediately followed by a shock.:
[*] I believe the red and green hash was often immediately followed by a shock.:
[*] I believe the blue and orange check was often immediately followed by a shock.:
[#] I believe the colored sticks were often followed a few seconds later by a shock.:
[#] I believe the blue ball was often followed a few seconds later by a shock.:
[#] I believe the red and green hash was often followed a few seconds later by a shock.:
[#] I believe the blue and orange check was often followed a few seconds later by a shock.:

+Specific Delay/Trace Contingency Questions (8)
[*] I believe the colored sticks were the best predictor of an immediate shock.:
[*] I believe the blue ball was the best predictor of an immediate shock.:
[*] I believe the red and green hash was the best predictor of an immediate shock.:
[*] I believe the blue and orange check was the best predictor of an immediate shock.:
[#] I believe the colored sticks were the best predictor of a shock that follows a few seconds later.:
[#] I believe the blue ball was the best predictor of a shock that follows a few seconds later.:
[#] I believe the red and green hash was the best predictor of a shock that follows a few seconds later.:
[#] I believe the blue and orange check was the best predictor of a shock that follows a few seconds later.:

# 4 Persistent Changes in Visual Cortex Due to Aversive Conditioning

This work was done with John P. O'Doherty and Christof Koch. John O'Doherty was, as always, indispensable in the analysis design and many discussions of data interpretation. Christof Koch worked on the goal and initial design of this experiment as well as review of different analysis phases. His talents in analysis were greatly appreciated during modeling and analysis of skin conductance responses. He also maintained the thankless task of insuring that funding was in place. Imaging acquisition was completed at the Caltech Biological Imaging Center. Skin conductance recording during fMRI acquisition was possible because of extensive testing and design done in collaboration with Antoine Bruguier (see Chapter 6).

We examined changes in the representation of two face images over the course of an aversive conditioning experiment and subsequent extinction.

## *4.1 Introduction*

Conditioning is commonly used to study learning and memory in a wide range of organisms (Baer and Fuhrer, 1982; Mackintosh, 1983; Gallistel, 1990; Thompson and Krupa, 1994; Connolly et al., 1996; Eichenbaum, 1997; Pearce et al., 1997; Tully, 1998; Squire and Kandel, 1999; Kocorowski and Helmstetter, 2001). Milner's study of HM (Scoville and Milner, 1957) provided evidence for the separation of explicit and implicit learning, since HM was capable of non-conscious, implicit learning but not of retaining explicit memories. Most discussions of implicit learning in conditioning center on

modifications in the amygdala and thalamic structures (Morris et al., 1999). However, the related field of perceptual learning demonstrates that modifications to the cortical sensory (in this case visual) pathway can and do take place (Poggio et al., 1992; Yang et al., 1994; Goldstone, 1998; Watanabe et al., 2002). The same is also true of the auditory pathway (Fritz et al., 2003). Perceptual learning is a change in the way a subject perceives a stimulus that is believed to reflect changes in sensory cortex. Subjects in perceptual learning studies show phenomena ranging from hyperacuity (Poggio et al., 1992), an increase in effective visual resolution, to the representation of previously unknown stimulus sets (Gauthier and Tarr, 1997; Gauthier et al., 2000). While these changes may occur with explicit influence, the resulting differences are of the automatic, implicit, variety. Modification of auditory cortex due to conditioning is well described in both electrophysiology and fMRI studies (Shamma, 2004; Weinberger, 2004). In these studies, a particular auditory stimulus which is of neutral value, the conditioned stimulus (CS), is reinforced with a reward or punishment, the unconditioned stimulus (US). The representation of this particular CS shows an increase in representation in auditory cortex as a result of conditioning (Kisley and Gerstein, 2001) . Correspondingly, individual cells show a shift in their response curve toward the frequency of the auditory stimulus used (Edeline et al., 1993). It seems likely that processes similar to those that take place during auditory conditioning may also cause modification of the visual pathway.

There are some fMRI studies that report differences in visual cortex during conditioning (Knight et al., 1999; Carter et al., 2006), but these studies were not directed at examining stimulus representation differences and also don’t examine the persistence of such changes. Recently, a study described changes in rat visual cortex due to

conditioning (Shuler and Bear, 2006). In particular, they report changes in the timing of neurons in visual cortex that reflected reward timing. Indicating that modifications to at least the timing of neural responses in visual areas can and do take place as a result of conditioning.

These studies describe changes that result from the acquisition of conditioning. If the changes in cortical response profiles described above are lasting and not due to continued modulation by an associative area of the brain, these changes should persist after the extinction of conditioned responses.

Here we describe a study of changes in a specific cortical visual area as a result of aversive conditioning in humans. We used a simple delay protocol to condition subjects to faces and abstract images during fMRI acquisition. We posited that those stimuli that were paired with a shock would elicit hemodynamic activity that increased over time when compared to a similar unpaired stimulus. We also hypothesized that this differential response in cortical visual areas would persist beyond extinction of the conditioned response.

BOLD activity in response to a reinforced face stimulus in the fusiform face area increases over time when compared to activity elicited by an unpaired face. We find that activity in this area remains consistently elevated in spite of extinction of the conditioned response, indicating that potential changes to perception of the previously reinforced stimulus are likely to be long lasting.

## *4.2 Methods*

### 4.2.1 Participants

From an initial pool of seventeen subject who participated in the experiment, our analysis includes twelve subjects who showed some indication of conditioning (greater SCRs for the CS+ than CS- at $p<0.2$ for both faces and abstract images, see below). Five of the original seventeen were not included in the final analysis because they either did not condition (2) or because of technical failures with the SCR recording (3). Subjects included in the analysis: six male and six female, age range 19-31 (mean 24.25). All subjects gave prior informed consent. This study was approved by the Human Subject Committee at the California Institute of Technology.

### 4.2.2 Experimental Procedure

Prior to entering the magnet suite, subjects were told the experiment was a learning and memory experiment. They were told that they would be presented with images and some of the images would be paired with a shock while others would not. They were asked to keep their eyes on a fixation mark and pay attention to the images presented. They were not asked to perform any other task. After confirming that they understood the instructions, the skin conductance and shock electrodes were attached and their shock level was determined. The subject was then positioned in the magnet. An anatomical scan was acquired first, followed by a 5 minute retinotopic scan using a rotating wedge stimulus. At this point, instructions for the main experiment were confirmed with the subject, and the acquisition portion of the experiment was performed.

Acquisition consisted of 60 presentations of each of the four different images (240 in total) with a 75% reinforcement rate for the CS+, and no shock pairings for the CS-. Intertrial intervals ranged from 7-9 seconds. Stimuli were presented so that no more than two of identical images occurred in a row and that nonreinforced test trials were spread evenly throughout acquisition. The stimulus that served as the CS+ was chosen at random. For subjects included in the analysis, each abstract image was reinforced for six of the twelve participants. For face stimuli, the dark haired individual was reinforced for seven of the twelve subjects and the image of the individual with lighter hair for the remaining five subjects.

Post acquisition, subjects were told they would be presented with the same images but would no longer receive any shocks (that the shock electrodes had been unplugged). They were reminded to keep their eyes on the fixation mark and to pay attention to each image presented. Each CS was then presented 20 times for 1second with an intertrial interval of 6 to 8 seconds as part of an extinction phase. Stimuli were presented so that no more than two of a particular stimulus occurred in a row.

### 4.2.3 Stimulus Presentation

Face images (Figure 4-1 top left) were obtained from the Ekman collection. Abstract images (Figure 4-1 top right) were adapted from a previous study (Carter et al., 2006). All images were presented at a size of approximately four degrees wide by six degrees tall for a period of one second. For reinforced stimuli, a shock was presented for one second overlapping with the last half second of the image presentation (see Figure 4-1 bottom). Images were presented on a grey background approximately 20 degrees in width

and 15 degrees in height at refresh frequency of 60Hz using LCD goggles from Resonance Technology Inc. (Los Angeles, CA).

Eight millimeter diameter Ag/AgCl electrodes containing a conductive paste (MedAssociates TD-246, 0.5M NaCl suspension) were attached to the top of the subject's right foot for stimulation. The shock stimulus was 60Hz alternating current of constant amplitude. Current levels for shock stimulation were chosen by each subject before the experiment, using an ascending rating method where the current amplitude was raised until the subject gave a rating of 9 on a scale of 1-10. A rating of 1 indicated the subject could barely feel the shock and 10 indicated the shock was painful and could not be used in the experiment. Subject shock levels ranged from 1.2 to 4.8mA with a median of 1.8mA.

### 4.2.4 Skin Conductance Conditioning

Skin conductance was recorded throughout the acquisition and extinction phases of learning. Eight millimeter diameter Ag/AgCl electrodes containing a conductive paste (MedAssociates TD-246, 0.5M NaCl suspension) were attached to the arch of the subject's left foot (Boucsein, 1992). These electrodes were attached to carbon fiber leads (Biopac, Goleta, CA) that ran through a wave guide into the control room, where they were filtered using the equipment described in Chapter 6. The signal was then recorded using equipment from Contact Precision Instruments (Boston, MA). Synchronization pulses were sent from the parallel port of the stimulation computer to the CPI standalone module. Special care was taken to ensure minimal radiofrequency induced contamination of the SCR signal (see Chapter 6).

Skin conductance responses (SCRs) were defined as the maximum amplitude response initiated no earlier than 1second with a peak no later than 5seconds after the CS onset. SCR amplitudes were range corrected by the maximum response for that subject (Lykken, 1972). Raw skin conductance data was filtered using a second-order Savitzky-Golay smoothing algorithm with a window of four seconds. This filtering removed any remaining radiofrequency noise in the skin conductance trace.

### 4.2.5 fMRI

Thirty-two slice EPI images were acquired at a resolution of 3 cubic millimeters using a Siemens Trio scanner and 8 channel phased array head coil (www.siemens.com). Slices were acquired in ascending interleaved order. These settings resulted in a TR of 2 seconds. Functional images were then preprocessed and analyzed using SPM5 (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK). Preprocessing included motion realignment, slice timing acquisition correction, normalization to the MNI template and smoothing to a final value of 8mm FWHM.

The functional analysis for acquisition modeled each of the four CS types separately. Parametric modulations of each of the CSs were used to test for changes over time and any effects due to the presence of the shock. In controlling for the presence or absence of the shock, the reinforcement schedule for the CS+ images was mirrored for CS- trials. Time changes were modeled using an exponential decay (see inset Figure 4-5) with a half life of ¼ the length of the experiment (Quirk et al., 1997; Buchel et al., 1998b). When looking for events that show an increase to plateau over time, we test for areas anticorrelated with this term (see inset Figure 4-6).

The functional analysis for extinction was the same as acquisition, except for the term to model differences between reinforced and nonreinforced trials (since no trials were reinforced).

## *4.3 Results*

The primary focus of this study was to look at changes in the representation of stimuli as a result of conditioning. We first describe areas of the brain that over all trials show a greater activation to the CS+ than CS-. We then focus on conditioning to faces and specifically two areas of the brain, the amygdala and a functionally defined cluster in the fusiform region thought to represent faces (Allison et al., 1994; Haxby et al., 1994; Kanwisher et al., 1997). We focus on faces rather than abstract images because of the wealth of information concerning face representation.

### 4.3.1 Skin Conductance Conditioning

On average, subjects used in our analysis demonstrate a conditioned aversion to the CS+. We compared skin conductance responses (SCRs) on nonreinforced CS+ trials (there were 15) to matched CS- trials. Individual plots of each subject's average response to the nonreinforced CS+ and matched CS- trials are shown in left and middle portions of Figure 4-2. It is important that CS+ vs CS- comparisons occur in a pair-wise fashion. While each subject has a larger average response to the CS+ than CS- it is not always true that each subject's average CS+ response is larger than all other subject's average response to the CS-. The box plots in the right portion of Figure 4-2 show the distribution of average differences between nonreinforced CS+ and CS- trials. Subjects show a

greater average response to the CS+ than CS-, indicating they were conditioned ($P<0.0001$ for abstract and face images; single sample t-test)

### 4.3.2 Brain Activity during Acquisition

We compared BOLD activity during CS+ trials with CS- trials. Regions of the brain that exhibit stronger responses to the face CS+ compared to the face CS- include the insula, operculum, caudate, globus pallidus, inferior parietal lobule, thalamic nuclei including superior colliculus, cingulate cortex and areas of the cerebellum (see Figure 4-3 and Table 4-1). These results are similar to those described in past studies (Buchel et al., 1998b; Carter et al., 2006). Similar areas were active when comparing responses for the abstract CS+ > CS-.

**Differences in Face and Abstract Image Representation**

BOLD responses to the images that were never reinforced, the CS-, were compared for faces and abstract images. Contrasts were masked for positive responses to the image type of interest. This comparison identifies those areas that preferentially respond to one particular class of stimuli. It does so without the shock or association formation as potential confounds. There is, however, the possibility that the CS- serves as a safety signal and could therefore be considered rewarding (Seymour et al., 2005). Similar results are obtained if the CS+ trials are used.

Areas of the brain that preferentially respond to faces include the fusiform face area, superior temporal sulcus, and medial occipital cortex (Table 4-2a; also see the left half of Figure 4-4). Areas of the brain that preferentially respond to abstract figures

include the occipital pole and parahippocampus (Table 4-2b also see the right half of Figure 4-4).

**BOLD interactions with time**

Previous studies (Buchel et al., 1998b) find that activity in the amygdala fits an exponential decay function. Changes in BOLD activity over time were therefore modeled using an exponential decay function (see methods, Figure 4-5 inset) in an effort to reproduce this result. We tested for an exponential decay over CS+ trials when compared to CS- trials. In line with the study by Buchel et al., we find that BOLD activity in the amygdala correlates with an exponential decay function for conditioning to both faces and abstract images (Figure 4-5). For faces, the peaks of amygdala activity were located at MNI coordinates 24, -6, -21 (right) and -15, -6, -24 (left). For abstract images, the peaks of amygdala activity were located at 30, -6, -15 (right) and -21, -9, -15 (left). These amygdala peaks were all significant at $P<0.05$ (family wise error corrected for multiple comparisons using a 10 millimeter diameter sphere around the peak voxel).

We hypothesized that reinforced stimuli would come to elicit stronger responses in areas that were capable of differentiating within that class of stimuli. We therefore tested for areas of the brain that showed a constrained increase in activity over time when compared to the CS- trials. The modeled increase is the inverted decay function of the previous section (Figure 4-6 inset). We specifically examined the areas identified in Table 4-2 (Figure 4-4) that preferentially represented each class of stimuli. We found that none of the areas preferring abstract figures correlated with an increase in activity over time (see discussion). However, the hypothesis held for faces. We found that activity in the fusiform area (peak at 42, -63, -15) showed increasing responses to CS+ trials over

time (see Figure 4-6; P<0.05 corrected for small volumes using a 10 millimeter diameter sphere centered at the peak). This cluster of voxels (P<0.01 uncorrected) fell within the face sensitive ROI (P<0.001 uncorrected). Figure 4-8 shows the extent of overlap between the fusiform face area and those regions that exhibited larger responses to the CS+ over time. Increasing activity to the face CS+ is not apparent in the superior temporal sulcus or medial occipital cortex using the thresholds described above. Using lower thresholds, activity becomes apparent in both the superior temporal sulcus and right medial occipital lobe. However, these clusters are a sizeable distance away from the ROI peak in another anatomical region ( > 20 millimeters) or do not survive correction for multiple comparisons. An increase in activity in the fusiform is consistent with an increase in representation of the CS+ image compared to the CS-.

### 4.3.3 Extinction

Following the acquisition of conditioning, subjects were told they would no longer receive any shocks. They were then shown 20 presentations of each CS with no shocks (extinction). Extinction of the previous association can be confirmed by the absence of differential skin conductance responses (SCRs). Testing of SCRs to both face and abstract images revealed no significant differences between CS+ and CS-, indicating an effective extinction of the previous association (see Figure 4-7a).

In contrast to the loss of differential skin conductance responses during extinction, the fusiform face area (peak at 36, -45, -24) remains more responsive to the CS+ than CS- (see Figure 4-7b; P<0.05 small volume corrected for a 10 millimeter diameter sphere

centered at the peak). At a threshold of P<0.001 uncorrected, it is clear from Figure 4-7b that the majority of the visual pathway shows a greater response to the face CS+.

The bottom section of Figure 4-7b shows activity in the geniculate nucleus. The left half is an SPM showing activity that is greater for the CS+ than CS- during extinction (peaks: left -21, -27, -6; right 24, -24, -3; P<0.05 small volume corrected for a 10 mm diameter sphere centered at the peak). The right half displays the overlap with visual hemifield stimulation using a standard rotating wedge (Wunderlich et al., 2005), confirming the responsiveness of the area to visual activity in one hemifield in a separate session. We find it surprising that activity in an area so early in the visual stream responds to one face more strongly than another.

There is no activity in the fusiform face area correlated with a differential increase or decrease over extinction, as might be expected if activity in perceptual areas mirrored association strength.

Activity in the fusiform face area increasingly responds to the CS+ face during acquisition and retains that differential response after extinction of the conditioned association. Figure 4-8 displays the overlap of regions that preferentially respond to faces (red), increasingly respond to the CS+ during conditioning (yellow), and consistently respond more strongly to the CS+ face in spite of extinction of conditioning.

## *4.4 Discussion*

We sought to provide evidence of modification of perceptual areas representing visual stimuli similar to established work in auditory conditioning (Bakin and Weinberger, 1990; Morris et al., 1998b; Bao et al., 2001). We used an aversive conditioning paradigm

to condition subjects to faces and abstract images. Examination of activity in the fusiform face area during conditioning to faces indicates that responses to the CS+ increase over time. Tests in the other face responsive areas listed in Table 4-2 do not provide strong evidence for increasing responses to the CS+ face (see Results). We therefore restrict our discussion to the FFA. In addition to acquisition, responses in the FFA were also examined during extinction phase, where conditioned stimuli are presented without reinforcement. Extinction is meant to form a new association that the previously reinforced stimuli are now safe. Larger responses to the CS+ trials persist through this separate extinction phase, indicating that changes in representation are present after the elimination of conditioned responses. We find no evidence of differential SCRs during extinction and also see little, if any, activity in areas associated with the conditioned response such as the amygdala or insula. This leads us to believe that there is little remaining affective component to the FFA activity in extinction.

With some differences noted due to task, familiarity, and emotion, repeated presentations of visual stimuli generally result in decreasing activity in visual areas (Henson et al., 2002; Ishai et al., 2004). A reduction in responses to known stimuli is consistent with the need to maintain a sparse representation, thereby allowing a greater number of stimuli to be represented. Results in auditory conditioning show both a shift in peak neuronal responses (Edeline et al., 1993) and a greater number of neurons responding to the conditioned stimulus even without a peak shift (Kisley and Gerstein, 2001). We know of no studies demonstrating persistent BOLD response changes for visual stimuli, although recent work in rats has shown a shift in timing (Shuler and Bear, 2006).

We find that as significance is attached to an image of a particular face, BOLD responses in the fusiform face area increase. The fusiform face area is posited to be an area of the brain with enhanced face representation (Puce et al., 1996; Kanwisher et al., 1997). The innateness and degree of this specificity are often questioned (Gauthier et al., 2000; Tarr and Gauthier, 2000), but, in general, this area of the brain does seem to be involved in the representation and identification of faces. It is plausible that an increase in activity in this portion of the brain to a specific face reflects a change in the likelihood that this face will be identified (an increase in saliency). Future experiments will test for behavioral correlates that could reflect saliency increases such as reduced search times when looking for CS+ faces. Another potential behavioral manifestation of increased representation by the FFA is that any given face could be more likely to be identified as the CS+ face. When shown an image of two faces mixed by morphing we would expect that previously conditioned subjects would have an equilibrium point, the point at which subjects are equally likely to identify the image as either individual, further from the CS+ face. That is, they would require a smaller CS+ identity percentage in order to describe the image as belonging to the CS+ individual.

One possible cause for the increased responsiveness in the FFA is changes in either top-down or bottom-up attentional processes. If the result is due to top-down attention, we would expect to see that activity in brain areas known to be involved in attention networks reflect the same changes we see in the FFA. In fact, during acquisition, bilateral BOLD activity in the intraparietal sulcus (IPS, right 21,-60, 36; left -27, -57, 45) also increases over time for the CS+ ($P<0.05$ small volume corrected for a

10mm diameter sphere centered on the peak). The IPS is thought to be involved in the control of attention; for review, see Pessoa 2003 (Pessoa et al., 2003). However, there is no evidence that this activity in the IPS is persistent throughout extinction as is the activity in the FFA. This leads us to hypothesize that connectivity analysis would reveal interaction between the FFA and IPS during conditioning in order to cause perceptual changes. The lack of persistent activity in attentional or emotional networks during extinction leads us to believe post-extinction changes in the FFA represent a bottom-up or saliency process contained in the FFA.

Also of interest is that increased responses to the CS+ face during extinction are not limited to the FFA. Areas as early as the LGN (see Figure 4-7) show similar differences. Face recognition can occur in small groups of neurons such as the brain of the honey bee (Dyer et al., 2005). However, it is more likely that LGN activity in extinction is a result of top-down influences from the FFA, since it both preferentially responds to faces and shows a pattern of learning during acquisition that was not present in the LGN. Connectivity analysis in future experiments may also be useful in disentangling driving effects in this situation.

In contrast to the results for faces, we find no evidence of representation changes in visual areas that preferentially respond to abstract images. There are many potential causes for this. First, the images chosen may not fit into an image class represented in a specific brain area. This explanation does not seem highly likely given that there is significant overlap between areas shown to prefer abstract images when comparing CS- trials and areas shown to prefer abstract images when comparing CS+ trials. The abstract images do appear to show some overlap in representation. Second, since the abstract

images are novel stimuli, there may not be adequate representation of the stimuli in the brain for differentiation to be localized. Third, as in all imaging studies, it may be that this study lacks the sensitivity to pick out a change in representation that differs between CS+ and CS- for the abstract images. The attempt to represent the novel images may overwhelm any difference between the two.

We have shown that learned changes in face representation due to conditioning are persistent throughout extinction. This persistent activity may reflect changes in saliency. Future experiments should include an assessment of the relative saliency of the conditioned face as well as connectivity analysis that could disambiguate the top-down and bottom-up effects taking place.

## *4.5 Tables*

**Table 4-1**

| Cluster Region | MNI Coord. | Cluster size | P value |
|---|---|---|---|
| a) Faces CS+ > CS- | | | |
| Left Posterior Insula | (-36,-12,15) | 25 | P<0.001 |
| Left Insula | (-39,0,6) | 111 | P<0.001 |
| Left Medial Thalamus | (-6,-27,-9) | 79 | P=0.001 |
| Right Medial Thalamus | (15,-12,12) | 30 | P<0.01 |
| Right Posterior Insula | (39,-18,-3) | 11 | P<0.01 |
| Right Insula | (39,-3,-12) | 16 | P<0.01 |
| Left Inferior Parietal Lobule | (-66,-27,27) | 10 | P<0.01 |
| Left Globus Pallidus | (-12,3,-6) | 8 | P<0.05 |
| b) Abstract CS+ > CS- | | | |
| Left Thalamus | (-12,-12,0) | 30 | P<0.01 |
| Left Posterior Insula | (-33,-21,12) | 10 | P<0.01 |
| Left Inferior Parietal Lobule | (-60,-18,18) | 11 | P<0.01 |
| Left Insula | (-33,-9,6) | 22 | P<0.01 |
| Right Medial Thalamus | (3,-15,6) | 25 | P<0.01 |
| Left Geniculate Nucleus | (-21,-21,-6) | 5 | P<0.05 |
| Left Inferior Insula | (-33,-12,-12) | 6 | P<0.05 |
| Right Inferior Insula | (-33,-21,12) | 7 | P<0.05 |
| | | | |

Table 4-1 Areas where brain activity is greater for CS+ trials than CS- trials for faces (a) and abstract figures (b). This table lists the region of activation followed by the MNI coordinates of its statistical peak, the number of voxels (3 $mm^3$) in the cluster and the peak whole brain corrected P value. Cluster size is those voxels with P<0.05 family wise error (FWE) corrected. Clusters containing 5 or less voxels are not listed (see also Figure 4-3).

**Table 4-2**

| Cluster Region | MNI Coord. | Learning | Sustained |
|---|---|---|---|
| a) Faces CS- > Abs. CS- | | | |
| Right Superior Temporal Sulcus | (57,-45,0) | No | Yes |
| Right Superior Temporal Sulcus | (45,-42,15) | No | Yes |
| Left Medial Occipital Cortex | (-12,-54,-3) | No | No |
| Right Fusiform | (39,-48,-24) | No | Yes |
| Right Medial Occipital Cortex | (6,-78,0) | Yes | Yes |
| b) Abs. CS- > Face CS- | | | |
| Right Occipital Pole | (30,-90,9) | No | No |
| Left Parahippocampus | (-27,-60,-12) | No | No |

Table 4-2 Areas of the brain preferentially responding to faces (a) or abstract (b) images (FWE corrected $P<0.05$, comparing CS- trials). This table lists the region of activation followed by the MNI coordinates of its statistical peak, whether or not it was active in learning during acquisition ("Learning") and if the region remained more responsive to the CS+ than CS- during extinction ("Sustained"). An ROI was considered active if statistically significant voxels ($P<0.01$ uncorrected) were found within the ROI mask ($P<0.001$ uncorrected) for the given CS type (faces > abstract or abstract > faces).

## 4.6 Figures

**Figure 4-1**

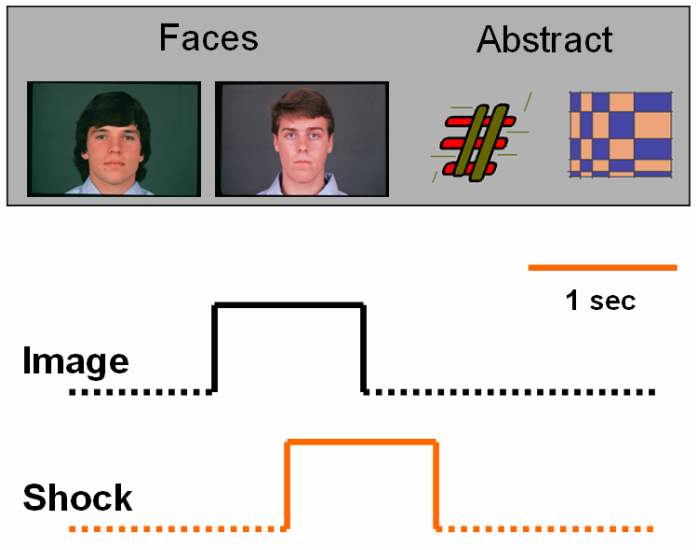



Figure 4-1 Experimental Setup – The top half of the figure shows the four images used as conditioned stimuli (CS). Each colored image was presented 60 times during acquisition and 20 times during extinction. One face and one abstract image were chosen randomly to be reinforced for each subject (the CS+). During acquisition, CS+ trials were reinforced on 75% of trials. The bottom half of the image shows the timing of stimulus presentation. Each image was shown for one second. On trials that were reinforced, a one second shock, the unconditioned stimulus (US) was given overlapping with the last half second of the image presented.

**Figure 4-2**

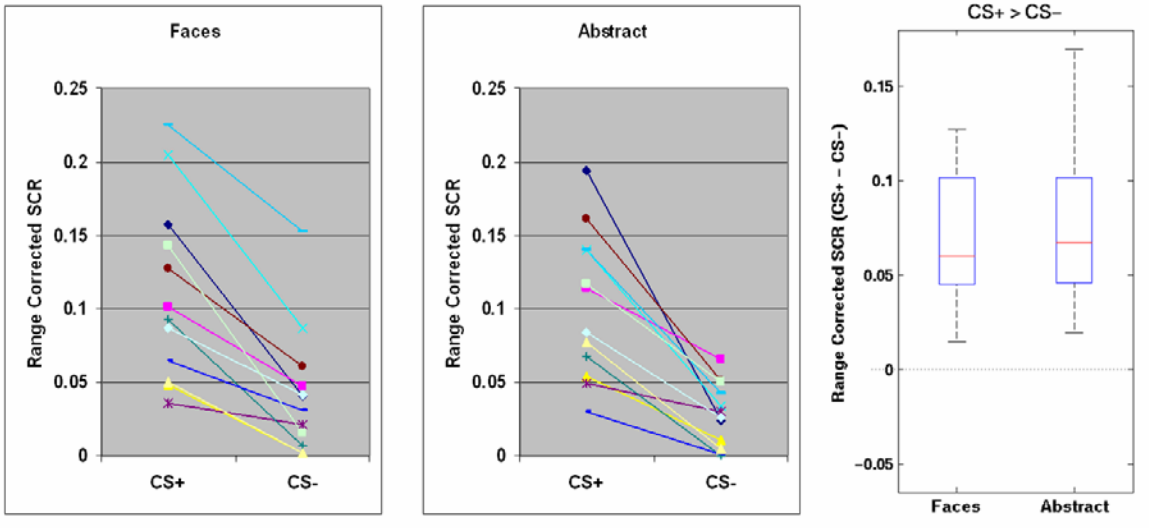



Figure 4-2 Skin conductance responses during acquisition show subjects learned to associate the CS+ stimulus with the shock for both faces and abstract images. The left two plots show average responses to nonreinforced CS+ trials and matched CS- trials for each subject. The leftmost plot is for face stimuli and the middle plot is for abstract stimuli. Each line represents a subject. Each subject shows that, on average, they had stronger responses to the CS+ than CS- and were therefore conditioned. The rightmost plot is a box and whisker plot of the average difference between CS+ and CS- stimuli for each subject, indicating that on average the group showed conditioning ($P<0.0001$ for both abstract and face images). The median of each group is marked by a red line. The bottom and top edges of the box represent the edge of the lower and upper quartiles, respectively. The whiskers indicate the full range of data. Any outliers are indicated by a '+'.

**Figure 4-3**

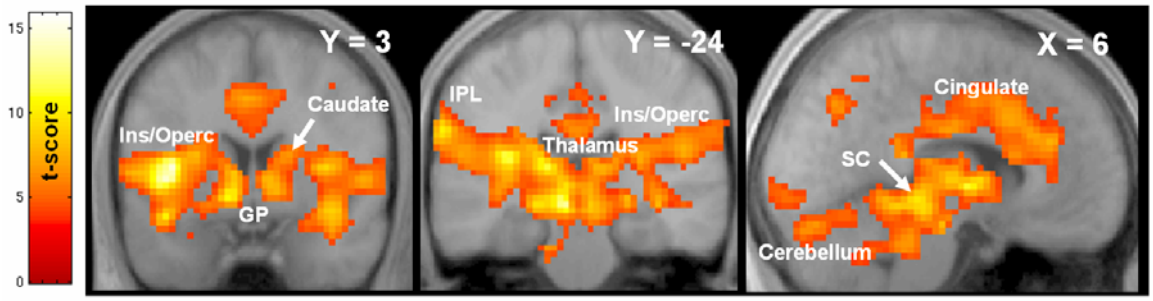



Figure 4-3 Statistical Parametric Maps (SPMs) of those brain regions that have greater activity on CS+ trials than on CS- trials during conditioning to faces. Prominent clusters include the insula (Ins), operculum (Operc), globus pallidus (GP), caudate, inferior parietal lobule (IPL), thalamus, cingulate cortex, superior colliculus (SC) and cerebellum. These results are consistent with previous studies. SPMs are shown at a threshold of $P<0.001$ uncorrected, overlaid on an average T1 structural image that has been normalized to the MNI template. Table 4-1a lists clusters of significant activity after whole brain correction for multiple comparisons.

**Figure 4-4**

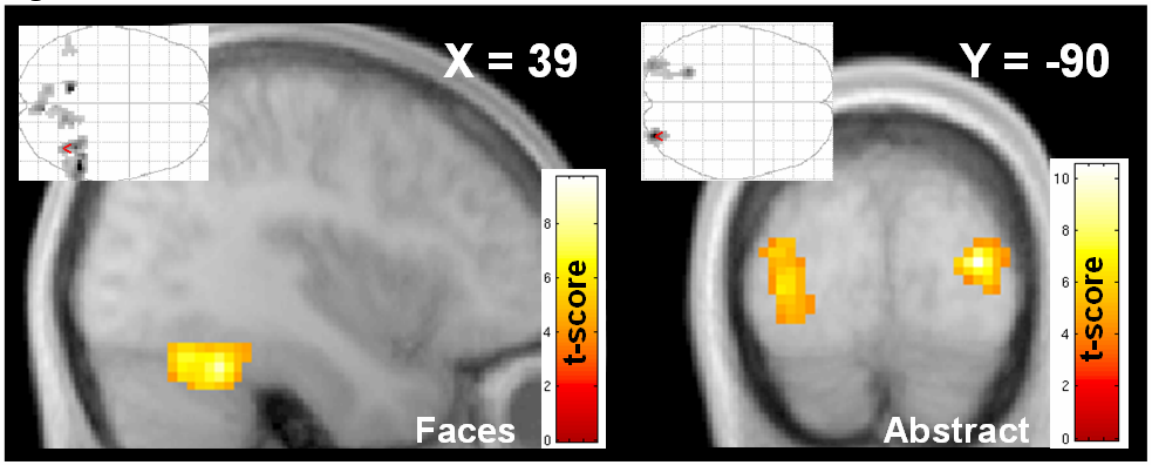



Figure 4-4 Differences in representation of the CS- during acquisition (faces > abstract left; abstract > faces right; P<0.001 uncorrected). Areas of the brain that respond more strongly to faces than abstract figures include the fusiform face area shown in the left image above. Abstract images elicit stronger responses than faces in extra-striate cortex as well as the para-hippocampal region. SPMs are shown overlaid on an average structural image that has been normalized to the MNI template. Glass brain images (inset) provide a means of visualizing any activity not in the slice displayed.

**Figure 4-5**

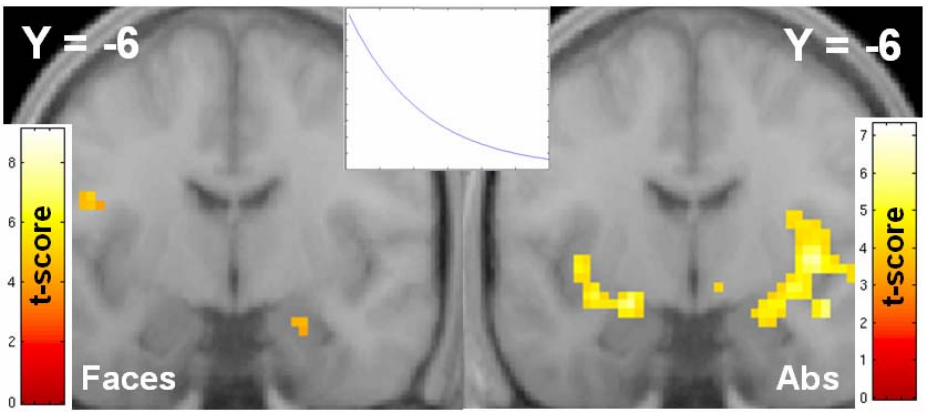



Figure 4-5 Differential responses (CS+ > CS-) in the amygdala correlate with an exponential decay function (inset) for faces (left) and abstract images (right). The SPM is shown at a threshold of P<0.001 uncorrected.

**Figure 4-6**

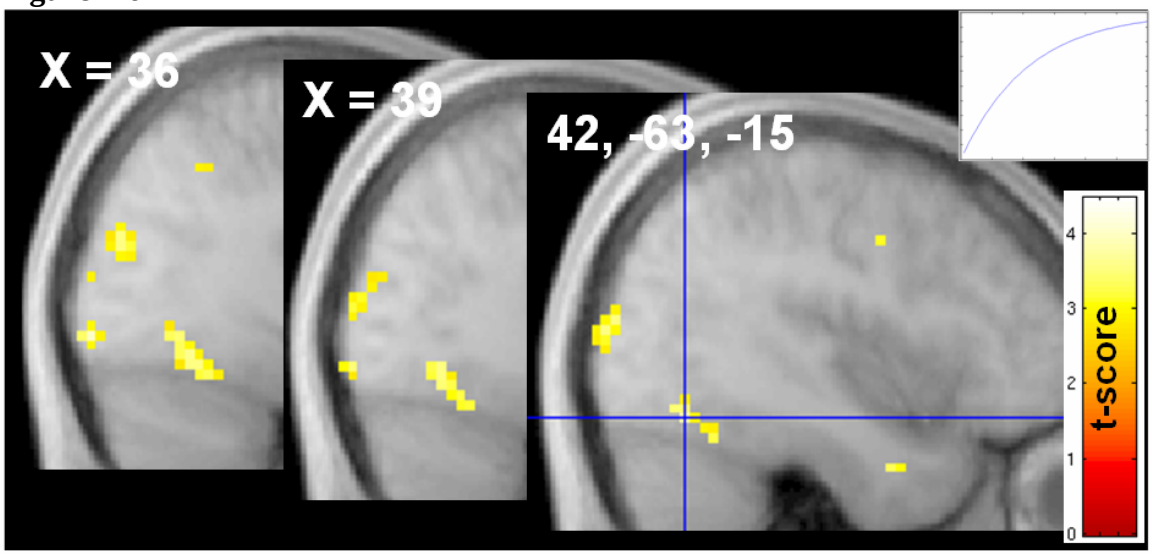



Figure 4-6 Differential responses to face presentations (CS+ > CS-) in the fusiform gyrus increase over time (inset). Coordinates for the cluster of activity marked by the crosshairs are given in the upper left corner. This SPM is shown at a threshold of P<0.01 uncorrected to illustrate the extent of activity in the fusiform gyrus.

**Figure 4-7**

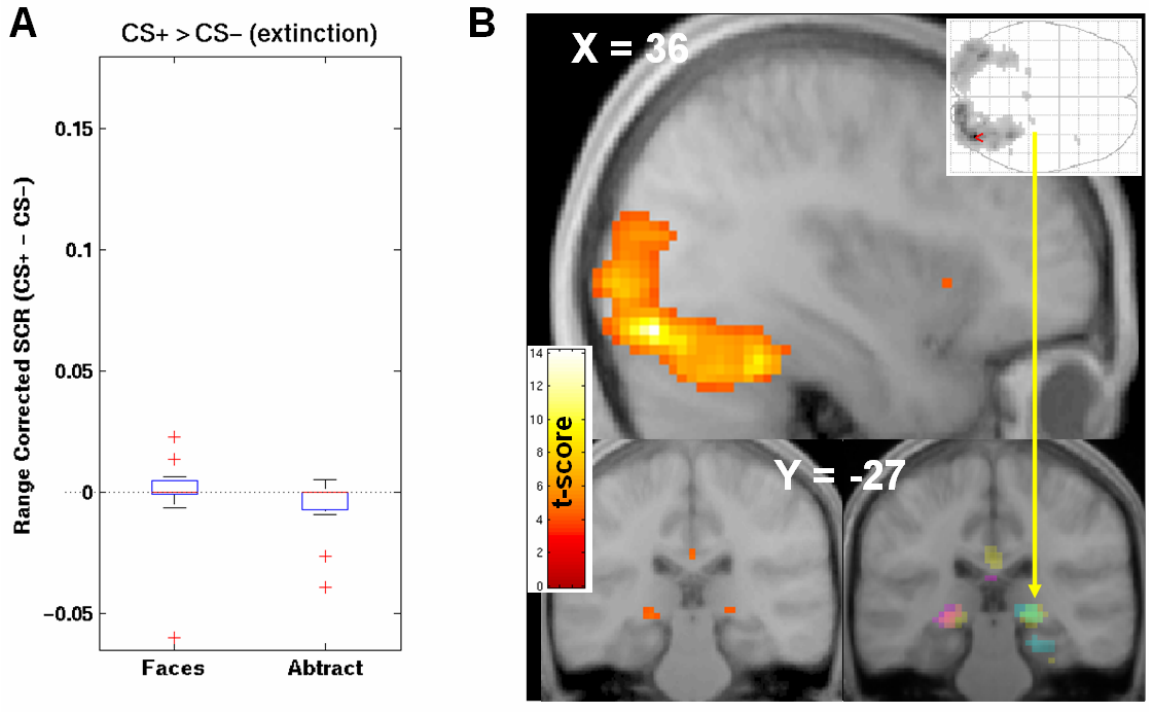



Figure 4-7 Skin conductance and brain activity during extinction. **A**) a box plot (see Figure 4-2) indicating subjects no longer show differential responses to the CS. The conditioned response has been extinguished. **B**) SPMs indicating areas that respond more strongly to the presentation of the face CS+ than face CS- during extinction. These areas include the majority of the visual stream for objects up to and including the fusiform face area (upper SPM, P<0.001 uncorrected). The bottom SPM (P<0.001 uncorrected) on the left is a slice centered on activity in the lateral geniculate nucleus (LGN). Localization of activity to the LGN is confirmed in the bottom right image (P<0.01 uncorrected) that maps the overlap between learned responses (yellow) and activity due to presentation of a wedge stimulus in each visual hemifield (magenta – right; cyan – left).

**Figure 4-8**

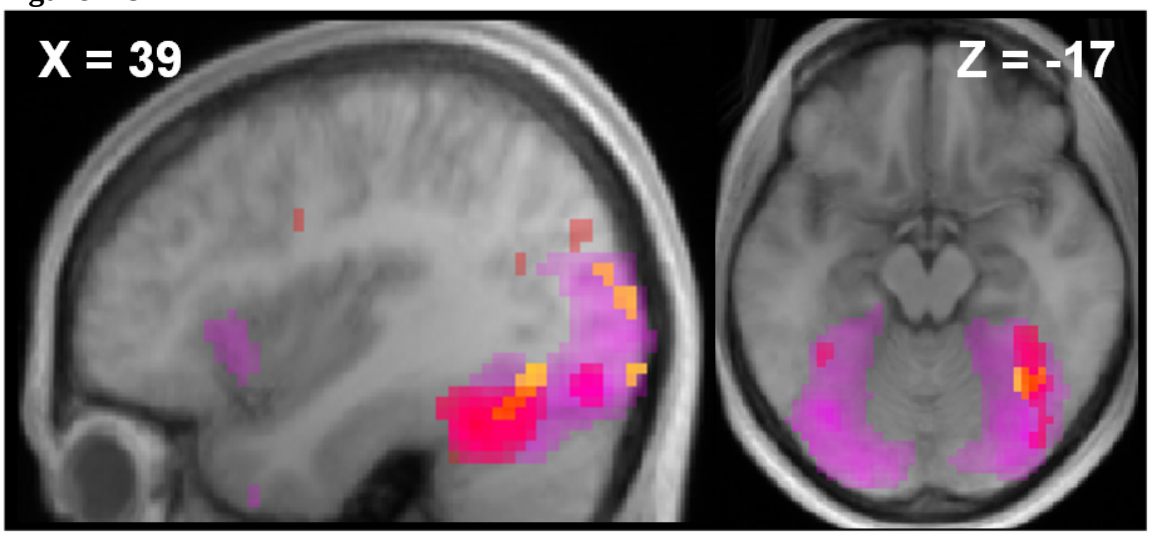



Figure 4-8 Summary of changes in face representation over conditioning and extinction. BOLD activity in the fusiform face area (displayed above in red; see Figure 4-4 and Table 4-2a) increases during acquisition of conditioning (displayed above in yellow; see Figure 4-6). This activity is persistent throughout extinction of the conditioned association (displayed above in magenta; see Figure 4-7). Activity is displayed using a lower threshold than previous figures ($P<0.01$ uncorrected) to indicate the extent of overlap.

# 5 Discussion

The previous three chapters provide a description of the interaction between explicit and implicit learning systems. Chapter 2 described the effects of performing a working memory task during aversive conditioning. Concurrent performance of a working memory task affects both delay and trace conditioning protocols used. Chapter 3 describes the areas of the brain involved in the explicit and implicit aspects of conditioned aversion. Activation in the amygdala correlates with implicit measures of learning, while activation in bilateral middle frontal gyri correlates with the accuracy of expectancy, our explicit measure of learning. Chapter 4 describes changes in visual cortex due to conditioning. These changes are persistent through extinction, reflecting long term changes in the representation, and potentially perception, of the face image used.

## *5.1 Explicit Influences on Implicit Processes*

Explicit knowledge exerts a strong influence over implicit processes. Results from Chapter 2 indicate that a high level task can result in less effective conditioning. This is true even for delay conditioning which, for eye-blink conditioning, is described by Manns et al. as being independent of awareness (Manns et al., 2002). A reduction in conditioning as a result of the concurrent working memory demand argues that working memory may be necessary for both delay and trace conditioning. However, the experimental manipulations that compensated for the drop in conditioning differ for delay and trace protocols.

In trace conditioning, the deficit can be compensated for by the combination of simplifying the protocol (using only one CS instead of two) and providing explicit

knowledge of the CS/US relationship. It does not seem to be affected by simplification of the protocol alone. The same deficit in delay condition is compensated for when the protocol is simplified. We did not test the effects of informing subjects in a delay protocol due to the levels of conditioning in a single cue protocol being near ceiling. It is possible that this difference in effective compensatory mechanisms reflects two different methods of working memory interference with conditioning.

During delay conditioning, prefrontal resources involved in working memory may not be necessary for conditioning, but strong prefrontal activity could still cause suppression of amygdala activity (Rosenkranz and Grace, 2001). Amygdalar suppression may be compensated for by increased activity in the amygdala when the number of stimuli is reduced. It may be true that the same process occurs during trace conditioning. However, it is also true that for trace conditioning, recovery is dependent on also receiving explicit knowledge of the CS/US relationship. This dependence on explicit knowledge for trace conditioning alone is consistent with the original results of Clark and Squire (Clark and Squire, 1998). The compensating effects for trace conditioning could be disentangled by providing explicit knowledge of the CS/US relationship to subjects in a differential (two cue) trace conditioning experiment.

Whether it is due to suppression of the amygdala or the lack of prefrontal resources, the deficit in delay conditioning due to a working memory task indicates a substantial influence of explicit activity in implicit processes – an influence that may not be necessary for the simplest associative learning, but most certainly takes place.

## 5.2 Implicit Influences on Explicit Processes

Conscious experience is a synthesis of implicit processes. Some implicit processes are not experienced explicitly or are greatly hindered when explicit focus is brought to them (Beilock et al., 2002). One clear example of implicitly trained differences in perception is the phenomenon called 'cue recruitment' (Haijiang et al., 2006). In this example of cue recruitment, subjects are conditioned to perceive a bistable stimulus in a certain way whenever a given cue is present. Eventually, the cue itself is capable of biasing the perception of the bistable stimulus.

Similar to the perceptual bias brought about by cue recruitment, Chapter 4 provided evidence that conditioning modifies the way a particular stimulus is represented. The stimulus that was behaviorally relevant elicited a greater response post training. We proposed that a modification of BOLD activity in FFA, without strong activity in any areas like the amygdala or insula, indicated a change to the saliency of the previously reinforced stimulus. An alternative hypothesis (proposed by committee member Shinsuke Shimojo) was that, similar to the somatic marker hypothesis (Damasio et al., 1991), the activity in FFA is reflective of remaining emotional association. Generally speaking, it might be expected that there would be BOLD activity in an emotional or associative area (such as the amygdala or insula) that showed the same characteristics as the visual network activated by the CS+. However, it is possible that the low level activity seen in the insula may be sufficient to maintain the network of activity as an emotional response. The connectivity analysis proposed in Chapter 4 would provide a means of differentiating these two explanations. The somatic marker proposal would be supported if the FFA still showed strong linkage to responses in the insula during extinction. If the FFA was not

strongly linked to the insula, but only other visual areas, it would argue for the change in saliency proposal.

One method often used to examine differences (and potentially overlaps) between explicit and implicit learning systems is that of conditioning to masked stimuli. In these experiments, a stimulus is presented to the subject for a short duration (30 ms) and then followed by a mask image presented for a longer period of time (45 ms). The second image prevents the first image from being processed by the visual system, and the subject will often report not having seen the first image. A number of studies performed analysis of masked vs. visible conditioned expression (Morris et al., 1998a; Critchley et al., 2002) as well as comparing visible stimuli that had been previously conditioned as masked or unmasked (Morris et al., 2001). These studies identified some areas that are classically thought to be implicit (such as the amygdala and insula). They also identified some regions of the brain that are normally thought to be precursors for explicit representation (such as the FFA). This is surprising, since most discussions of non-conscious visual processing center around a potential pathway involving the superior colliculus and pulvinar nucleus (McIntosh and Gonzalez-Lima, 1998; Morris et al., 1999).

In an effort to examine the time course of learning and identify those areas directly involved in the learning process, we performed a set of experiments conditioning subjects both implicitly and explicitly using similar parameters. The protocol chosen conditioned greater than 50% of the subjects used in pilot experiments without fMRI scanning. After fMRI data collection, it became clear that the subjects being conditioned during fMRI acquisition did not develop a strong enough association to justify any fMRI results. The difference in conditioning between our pilot studies and fMRI results led us

to believe that the increased number of stimuli present in the fMRI environment resulted in reduced conditioning. In hindsight, the results look similar to those described in Chapter 2. Subjects placed in a distracting environment were no longer easily conditioned. Given that the reduced conditionability depends on cross-modal factors (such as auditory noise in an MRI scanner) and that past implicit conditioning studies show changes in cortical representations, it seems likely that the classical model of a distinct explicit and implicit systems is inadequate. In fact, incorporating sensitivity measures available in an ROC analysis, amygdala responses to masked faces don't appear to be automatic but rather related to the face's visibility (Pessoa et al., 2006). These results argue that the difference between explicit and implicit may be more continuous than originally proposed.

## *5.3 Continuum or Separate Systems*

Evidence from HM first led researchers to explore the potential separation between explicit and implicit learning systems (Scoville and Milner, 1957). Evidence for a non-conscious visual pathway came from blind-sight patients who seemed to be performing visual tasks without explicit knowledge (Weiskrantz et al., 1974). Work by Clark and Squire (Clark and Squire, 1998) described delay conditioning as occurring independent of awareness (Manns et al., 2002). This in spite of previous work arriving at a theoretical agreement called the "necessary gate hypothesis;" that explicit knowledge of the CS/US relationship was necessary but not sufficient for conditioning (Dawson and Furedy, 1976).

What has been described as non-conscious conditioning has been shown using masking to hide the CS and a variety of techniques to assess explicit knowledge (Soares

and Ohman, 1993). Explicit assessment techniques range from expectancy reports to familiarity questionnaires. Using these techniques, groups have even described non-conscious trace conditioning when the stimuli are fear relevant (Ohman and Soares, 1998), Whether or not masked conditioning is a non-conscious process has been questioned. Arguments against masked conditioning being non-conscious mostly concern the ability to assess awareness of the masked stimuli (Lovibond and Shanks, 2002). Difficulties include memory requirements for the test, whether the test is objective or subjective and partial identification of conditioned stimuli. Recent work has also questioned whether traditional statistical assessments have the sensitivity to identify subtle differences due to conscious recognition (Pessoa et al., 2006). This debate surrounding what is conscious and what is non-conscious is suggestive of a subtle grade between the two systems rather than a sharp separation.

The work presented in this thesis fits with the possibility of a continuum of conscious experience rather than two distinct subsystems. At minimum we find interaction between the two subsystems in each study. There is the caveat that all of this work has been in healthy subjects. It is entirely possible that the explicit and implicit systems can't be disentangled when all brain areas are functioning normally, that they function independently only when large portions of the brain have been damaged. In and of itself, this would be interesting since it would allow anatomical distinction without necessarily requiring functional separation.

## *5.4 Work on the Conscious Mouse*

This work was begun in an attempt to better characterize the processes and substrates involved in conscious experience. Collaboration was initiated between my advisor

Christof Koch (Caltech), David Anderson (Caltech), Michael Fanselow (UCLA), and later Henry Lester (Caltech) and Tad Blair (UCLA). Beginning with Larry Squire's result that trace conditioning was related to explicit learning while delay was not, the group initiated studies of the differences between the two types of conditioning. Since trace and delay conditioning are studied both in humans and a large number of animal models, this provided a great opportunity to study the substrates required for trace conditioning with an eye toward analogues of explicit learning in humans. Efforts from human studies provided better information about the nature of explicit processes while animal model systems allowed for the use of tools (molecular, genetic, and lesion) not available when studying human subjects. The explicit distraction results described in Chapter 2 led to the use of a distraction task in mice during conditioning. Results indicated that the presentation of a cross-modal stimulus made trace conditioning less effective but did not affect delay (Han et al., 2003). The focus of the collaboration then shifted to the use of genetic tools to reversibly silence specific populations of neurons (Slimko et al., 2002). Using a combinatorial technique, the silencing could be directed toward specific neuron types that may be spatially inter-mixed with other neurons. Initial targets for silencing after verification include different nuclei in the amygdala as well as the anterior cingulate and prefrontal cortex. When these tools are fully developed, they should yield a great deal of information about implicit and explicit processes in conditioning.

## *5.5 Conclusion*

There are robust interactions between explicit and implicit processes that provide a method of balance between the two learning systems, and, most importantly, seem to provide a method for forgetting. Specifically, areas of the prefrontal cortex are involved

in extinction of conditioned associations (Milad et al., 2006; Sotres-Bayon et al., 2006). One study of functional connectivity in post traumatic stress disorder patients indicates that the disorder may be a result of excessive amygdala activity (Gilboa et al., 2004). Without suppression from prefrontal areas, the amygdala causes increasing activity in higher sensory areas, creating a feedback loop that becomes unmanageable. The necessity of these interactions is also indicated by the relative success of cognitive therapy in treating anxiety disorders.

There is evidence from a large number of areas that indicate it may be possible to sometimes separate implicit and explicit learning, the hippocampal patient HM described earlier is a good example. However, both the successful and unsuccesful work presented here indicates that explicit and implicit learning systems display little independence in practice.

# 6 Appendix – SCR Recording During FMRI Acquisition

This work was conducted by Antoine Bruguier, R. McKell Carter, Christof Koch and Steven Quartz. Experiments were carried out at the Caltech Biological Imaging Center (CBIC) by AB, CK and RMC. Steve Flaherty and J. Michael Tyszka from the CBIC were also very helpful in conducting the experiments. Analysis and interpretation were conducted by all authors. The first draft of the text below was prepared by AB. Figures were prepared by AB and RMC. All authors were involved in the review of this manuscript. We also received assistance during the filter design process, specifically in how to ensure subject safety while using grounded filters, from Alan Macy of Biopac (Goleta, CA).

## *6.1 Abstract*

Investigative methods in neuroscience increasingly combine functional magnetic resonance imaging (fMRI) with other measurement and stimulus-delivery systems. Many of these, such as electroencephalography (EEG), electrocardiography (ECG) and skin conductance response (SCR) measurements, attach electrodes to subjects inside the strong variable magnetic field of the scanner. This may induce dangerous voltages on the leads that often go unassessed. While burn injuries and electric shocks have been reported, there is surprisingly little available research describing these risks. This paper presents a simple model of the human body and a filtering system that aims to assess these burn risks and prevent electrical shocks. The electrical properties of this setup and the induced voltages on the leads as measured in a variety of configurations, including the effect of fMRI transmitting and receiving coils and lead composition, are presented. Since these combined methods introduce noise that requires additional filtering, we also

studied the safety constraints of various filters. Even though the design methods and measurements are applied to a skin-conductance/shock delivery setup, they can be generalized to other systems for assessing and preventing risks associated with similar combined methods.

## 6.2 Introduction

Many recording methods often combined with fMRI, such as EEG, GSR, and ECG, involve attaching leads to subjects. There are two main risks of having leads attached to subjects during MRI: inducing currents in leads that may cause sufficient heat to burn subjects, and creating an electrical current inside the subjects themselves. Despite these risks, we have not found any satisfactory studies of the risk inherent to attaching electrodes, since most references, such as (Shellock, 2000b, a), concentrate on safety regarding the specific absorption rate (SAR – the amount of radio frequency energy absorbed in tissue, usually watts per kilogram for a given volume) and implanted devices. The limitation on the SAR was implemented in order to reduce the heating of the subject's tissue, and is now regulated by the FDA. Ferromagnetic implants will experience an attractive force and may cause physical harm, and numerous cases of injuries and even deaths have been reported. For this reason, most research institutions screen subjects for implanted devices and virtually ban most of them. Burn injuries are not, however, limited to implanted devices, as the presence of electrodes is in itself a hazard. The FDA has reported excessive heating resulting in third degree burns in the case of an ECG connection (see for example report M765635 in their Medical Device Reporting database). Finally, the voltages created by the scanner create noise in the

recording devices. Given that the strong variable magnetic field inside a scanner may induce a voltage in any attached leads, such methods raise two important issues: 1) what are the direct safety consequences to the subject and 2) how can the noise such leads introduce be eliminated without causing further safety concerns?

As indicated above, the variable magnetic fields inside the scanner create substantial noise in the various recordings that should be filtered out to obtain a usable signal. A first step is to use analog filters before the actual recording takes place. Unfortunately, these filters have their own safety requirements. First, they usually require a ground connection, and one should be concerned about connecting a subject to a ground lead because there can be a voltage difference between a room ground and the ground conductor of a medical device. Subjects who are in contact with two unequal ground references may experience a leakage current. Second, these filters modify the recording circuit itself; therefore, the safety of the subject and the quality of the recordings should be jointly studied.

We here investigate these issues through two conjunctive methods: skin conductance recording, and the delivery of shocks in a scanner. Skin conductance recording is a fairly common conjunctive method that is used for peripheral correlates of emotional states, while shock delivery is increasingly used for behavioral conditioning and pain experiments (Carter et al., 2006). Although we focus on these two applications, these methods carry over to other applications.

## *6.3 Material and methods*

### 6.3.1 Equipment

The full schematic of our equipment is shown in Figure 6-1. Two devices from Psylab (Psylab SAM, Boston MA), a skin conductance instrument and a shock delivery device, were powered through a PC-managed controller. In between the devices and the simulated subject we placed a low-pass filter that is described in more detail below. On the subject side of the filter, various types of leads were connected and attached to the simulated subject. A wave-guide served as the interface between the control room and the room where the scanner is located.

Since we use a 3T scanner (Siemens Trio), its Larmor frequency is ~123 MHz; thus, the values of components in the filters presented here are designed for such frequencies. Since there is variability in different institutions' hardware, only the methods can be generalized and the results should be investigated on a case-by-case basis.

### 6.3.2 Body simulation

To simulate the electrical properties of a human body attached to pairs of leads, we used conductive electric paste. This paste, Med-Associates TD-246 (0.5M NaCl suspension), was originally designed to create contact between a subject and electrodes and has similar conductive properties to human perspiration. To mimic both hands, paste was placed in two plastic dishes on a strip approximately 4 inches long and 0.5 inch wide, resulting in a resistivity of approximately 30kΩ. These two dishes were then connected together by another strip of paste (0.25 inch wide and 16 inches long) to simulate a subject's chest (Figure 6-1). Since the magnetic fields increase with the proximity to the center of the

scanner, we placed this model in the approximate location a real subject's hands would be located in relation to the scanner center.

By using this "dummy", we tried to mimic the various loops created by four leads. It should be noted that the model described above should be modified for other setups that include more leads.

### 6.3.3 Leads and electrodes

Since the behavior of various lead materials within the scanner is not firmly established, we used a number of different leads (custom made by InVivo Metric, Healdsburg, CA, and Biopac Systems Inc, Santa Barbara, CA) to investigate the extent of induced voltages in them. We tested regular copper wires (30 foot, 16 AWG), short (6 foot) carbon fiber leads extended by 24 foot-long regular copper wires, and full-length (30 foot) carbon fiber leads. Some of the leads were shielded (standard copper coaxial shielding) and we tested both when this shielding was connected to the common ground reference and when it was not. While carbon fiber has the advantage of being radio-translucent and is, therefore, less likely to experience induced currents, it is also more expensive and not as readily available. The end-connections to electrodes were either regular or snap-on, a type of connector that can be snapped on a socket pasted on a subject's skin. Given the large number of possible lead/electrode/end-connection combinations, we restricted ourselves to a smaller set of electrode types. However, the length of the electrode lead was fixed to 30 feet, since a variation in the length of the wire would modify the resonance properties of the whole system.

In order to limit the effect of variable geometry, the placement of the leads was also fixed. Across all experiments, leads immediately descended to the floor, then directly to the wave-guide, and into the control room.

### 6.3.4 Heat insulation

Even though our work was designed to prevent any risk of burn on the subject, we implemented additional safety measures. Because the induced voltage is directly proportional to the surface between the conductive loop (Faraday's law of induction), we twisted together each pair of wires and stuffed them into standard window insulation foam. In addition to keeping the two wires close together and reducing the risk of accidental coiling, it prevented direct contact of the wires on the subject's skin.

### 6.3.5 Filter design

We used two types of filters, a simple capacitor filter and a third-order pi filter. These filters were placed in line with the electrode leads inside the MRI control room (see the box marked 'filters' in Figure 6-1). The simple filter type consisted of a 10pF capacitor between each wire of a lead pair (Figure 6-2). At high frequencies, the capacitor behaves as a simple wire and effectively short circuits the two leads. This results in a low-pass filter, rejecting differential mode noise signal while keeping the low-frequency GSR recording. This type of filter; however, proved to be insufficient for our noise constraints, as the signal of interest was not clearly visible. A large amount of noise was common between the leads with this filter.

The second type of filter was a standard third-order pi design diagrammed in Figure 6-3. The filter response with loadings of 30 kΩ on both ends is displayed on Figure 6-4. The main safety issue with this design is the need to have a connection to the ground. Connecting a subject to a non-isolated ground is regarded as dangerous because a voltage differential between two references can result in electric shock. We therefore used high voltage capacitors (rated 3kV) to prevent such risks. This practice was suggested by the international norm IEC 60601-1.

For the second type of filter, we used two types of ground connection, the cage surrounding the scanner room, or a ground common to the Psylab hardware. In the first case, the filter should reduce the noise on the skin-conductance measurements, but in the second case there is additional electrical isolation (discussed in more detail in results below). It is most important that equipment electrically connected to the measuring equipment be connected to the same ground to minimize any ground reference differences. We also minimized any line noise by using band-pass filtered power strips.

### 6.3.6 Head coils

We used two types of head coils, as they could potentially modify the currents in the leads by modifying the characteristics of the magnetic field. The first type was the standard “bird cage” coil (CP Head, receive and transmit, Siemens Medical, Munich Germany). We also used a custom “8-channel” coil (receive only, MRI Devices, Orlando, FL). Because the scanner body coil is used as a transmitter while the 8-channel coil is used as a receiver only, this setup yields a better image quality but induces greater noise in attached electrodes.

### 6.3.7 Resonance testing

The first safety test was performed outside the scanner utilizing a network analyzer (Agilent 8712ET 300kHz-1300MHz RF Network Analyzer, Agilent, Palo Alto, CA). Different parts of the installation were connected together except for the power supply, which was disconnected in order to test the passive properties of the circuit. By connecting the probe electrodes to the leads on the paste human model, we could sweep across a wide range of frequencies in order to detect resonances. Our rationale for this safety test was that the network analyzer injects frequencies in a fashion similar to the scanner magnetic field. If a circuit had presented a resonance at the scanner's Larmor frequency, it would be considered unsafe. Results are described below.

### 6.3.8 Measurement of induced voltages

The second set of measurements was performed with all the equipment turned on. After placing the paste model into the scanner, we ran an EPI scan (T2*-weighted PACE EPI TR=2000ms, TE=30ms, 64x64, 3.28125x3.28125 mm2, 32 3.0mm slices, no gap, field of view = 210) and measured the voltages between leads with a digital oscilloscope (TDS 5104 Digital Oscilloscope 1GHz 5GS/s, Tektronix, Richardson, TX). We took three measurements; the first was between the two leads of the SCR electrodes, the second one was between the two leads of the shocking electrodes, and the third one was between one lead of each type.

Since the sequence does not produce voltages between the leads continuously, a direct measurement cannot be taken. We increased the trigger level of the oscilloscope

until the trace was stable and then read both the peak and the root mean square (RMS) values directly. The peak values reflect the maximum instantaneous voltage received and the RMS values are a direct measure of the energy induced in the dummy.

It should be noted that we limited ourselves to EPI sequences during our measurements on the dummy, and that the leads should be disconnected when a human subject is scanned with another sequence or during shimming.

## *6.4 Results*

### 6.4.1 Resonances

Figure 6-5 shows one example of a network analysis plot. The lack of a sharp dip around 123 MHz reveals that the circuit does not show specific resonance around the Larmor frequency, and that most of the energy injected into the circuit at that frequency is not absorbed (in this example the absorption is 2.4dB). We observed several other dips at other frequencies, but since they are far away from our operating range, we concluded that they presented no safety risks.

None of our various configurations presented any resonance around the Larmor frequency, and we therefore proceeded to the next step.

### 6.4.2 Recorded waveform

Figure 6-6 shows a typical waveform recorded during an EPI sequence. One can see a group of two pulses that occur at repeated intervals. We matched this frequency with the number of slices acquired every second. The first pulse of the group is the fat-saturation pulse, while the second narrower pulse corresponds to a slice selection pulse.

By increasing the time resolution, we can look into the larger amplitude of the two, the slice selection pulse. The measured frequency matches ~123 MHz. This confirms that the signal we recorded is induced by the scanner and is the one to be investigated to test the safety of the installation.

### 6.4.3 Recorded voltages

We then proceeded by repeatedly recording the voltages induced during the slice selection sequence. Three main parameters were identified: type of filter, type of head coil and type of lead. Among all combinations tested, no measurement was above 3000mV, which, with a skin conductance of about 30kΩ, would create a current of 0.1mA, generally considered below detectable limits.

Two types of head coils were used: standard “bird cage” and high quality “8 channel,” as shown in Table 1, Section 1. Results indicate that modification of the magnetic fields greatly changes the induced voltages on the leads. The values recorded when using the bird cage coil (top three rows) are significantly below ($p < 0.001$ in all cases) the ones when using the 8-channel coil (bottom three rows). The bird cage is a receive/transmit coil that probably confines most of the variable magnetic field to a region close to the head. The 8-channel coil, being only a receiver, uses the magnet’s coils as transmitter and therefore yields a higher variable field near the hands. Even though the 8-channel coil yields higher voltages, the values are minimal and its use is still safe. Therefore, we chose to use it over the bird cage, as it provides superior fMRI recordings.

Section 2 of Table 6-1 shows the effect of the different leads. The carbon fiber leads seem to display the lowest induction and we believe that, unless one is concerned with their relatively low conductance (resistance of 200 Ω for 1 m) or their higher cost, they should be used. We can also note that the shielding lowers the inducted voltage if properly grounded.

We measured voltages (see Table 6-1, Section 3) for two types of filters. Even though the filters were designed to improve the quality of the recordings, they are a parameter when it comes to subjects' safety. The two ground connections for the type-2 filter do not modify the recorded values significantly. This may have been due to remaining ground reference differences. As the type-1 filter neither provides better quality signal nor lower induced voltages, we do not recommend its use unless one does not want a connection to the ground at any cost.

### 6.4.4 GSR recording quality

Figure 6-7 displays a typical GSR recording showing the onset of EPI scans. The first part of the figure shows a typical skin conductance recording while the second part depicts a recording during an EPI scan, the onset of the scanning being marked with a vertical line. The first recordings were of poor quality due to the presence of interference from other electronic equipment and bad lead connections. However, a careful set-up leads to much higher signal quality, and this configuration yielded an SCR signal with very little noise.

## 6.5 Discussion

Many investigation techniques in neuroscience, such as EEG and ECG recording, skin conductance measurements, or the delivery of shocks, are useful for investigations in neuroscience. However, recording in conjunction with fMRI scanning presents safety risks and adds noise that requires signal filtering. In this chapter, we presented a method to evaluate the safety of a complete recording system. All values point toward induced currents that are well below safety requirements. In addition, the filter presented eliminates most of the noise induced by the scanner, although further digital filtering can be applied.

This procedure for safety testing can be easily reproduced for other systems. Even though it appears that this type of analysis is rarely done, the effect of the leads, filters, or head coil shown above prove that any system should be tested prior to use on human subjects. The measurements can be reproduced to provide early testing of any biological recording system.

## 6.6 Tables

**Table 6-1**

| | | Peak | | RMS | |
|---|---|---|---|---|---|
| | | Mean | STD | Mean | STD |
| **Bird cage** | **SHK** | 76.1 | 7.9 | 19.8 | 5.6 |
| | **GSR** | 89.8 | 5.1 | 39.2 | 5.4 |
| | **XRS** | 206.4 | 12.6 | 86.9 | 2.7 |
| **8 channel coil** | **SHK** | 733 | 155.8 | 383.3 | 89.8 |
| | **GSR** | 588 | 82.9 | 283.2 | 41.1 |
| | **XRS** | 305 | 24.3 | 117.6 | 15.7 |

*filter 1 / carbon fiber leads*

| | | Peak | | RMS | |
|---|---|---|---|---|---|
| | | Mean | STD | Mean | STD |
| **Copper leads** | **SHK** | 1264 | 23.9 | 466.6 | 27.3 |
| | **GSR** | 382.4 | 13.1 | 185.1 | 8 |
| | **XRS** | 2176 | 61.4 | 982 | 39.3 |
| **Carbon fiber extension** | **SHK** | 998 | 42 | 474.4 | 30.2 |
| | **GSR** | 485 | 41.5 | 240.7 | 31.3 |
| | **XRS** | 579.6 | 11.7 | 264.2 | 4.9 |
| **Shielded snap leads** | **SHK** | 2446 | 111.8 | 1523.8 | 62.7 |
| | **GSR** | 1174 | 26.8 | 588.5 | 26.4 |
| | **XRS** | 784 | 46.6 | 350.8 | 23.8 |
| **Snap leads** | **SHK** | 732 | 71.2 | 326 | 40.4 |
| | **GSR** | 1512 | 20.9 | 781.8 | 28.2 |
| | **XRS** | 307.4 | 26.7 | 241.5 | 7.3 |
| **Carbon fiber leads** | **SHK** | 733 | 155.8 | 383.3 | 89.9 |
| | **GSR** | 588 | 82.9 | 283.2 | 41.1 |
| | **XRS** | 305 | 24.3 | 117.6 | 15.7 |

*8 channel coil / filter 1*

| | | Peak | | RMS | |
|---|---|---|---|---|---|
| | | Mean | STD | Mean | STD |
| **Filter 1** | **SHK** | 773 | 155.8 | 383.3 | 89.8 |
| | **GSR** | 558 | 82.9 | 283.2 | 41.1 |
| | **XRS** | 305 | 24.3 | 117.6 | 15.7 |
| **Filter 2 - room ground** | **SHK** | 336.6 | 23.2 | 164 | 32.5 |
| | **GSR** | 266.8 | 6.5 | 128.8 | 5.4 |
| | **XRS** | 224.8 | 19.4 | 103.8 | 10.3 |
| **Filter 2 - isolated ground** | **SHK** | 335.2 | 6.8 | 164.6 | 4.3 |
| | **GSR** | 316.4 | 13.6 | 149.2 | 7.7 |
| | **XRS** | 188 | 20.4 | 89.4 | 9.6 |

*8 channel coil / carbon fiber leads*

Table 6-1 Comparison of the effect of the head coil: We measured the peak and RMS voltages (in milivolts) for different configurations. The probe leads were connected between the two shock leads (SHK), the two skin conductance leads (GSR), or between one shock lead and one skin conductance lead (XRS). Measures were taken repeatedly and we reported the mean value (left column) and the standard deviation (right column). The first part of the table describes the effect of the head-coils, the second part the effect of the leads, and the third part the effect of the filters.

Figures

**Figure 6-1**

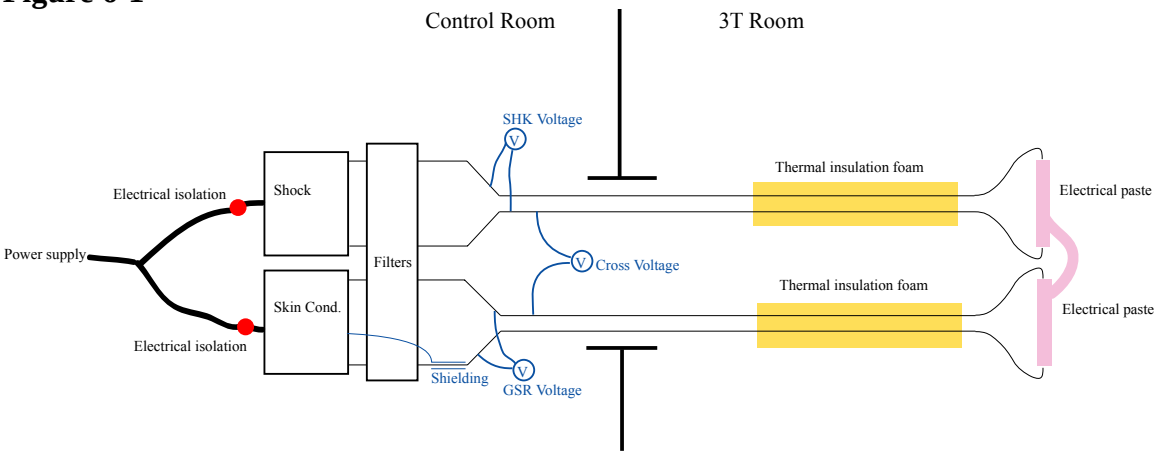
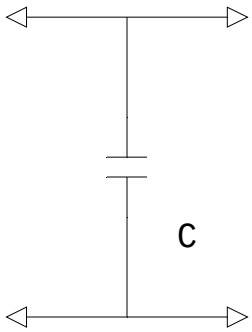


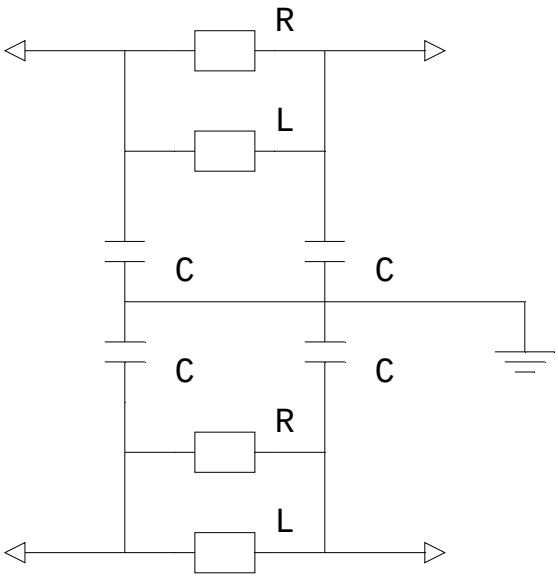

Diagram of the experimental setup

**Figure 6-2**



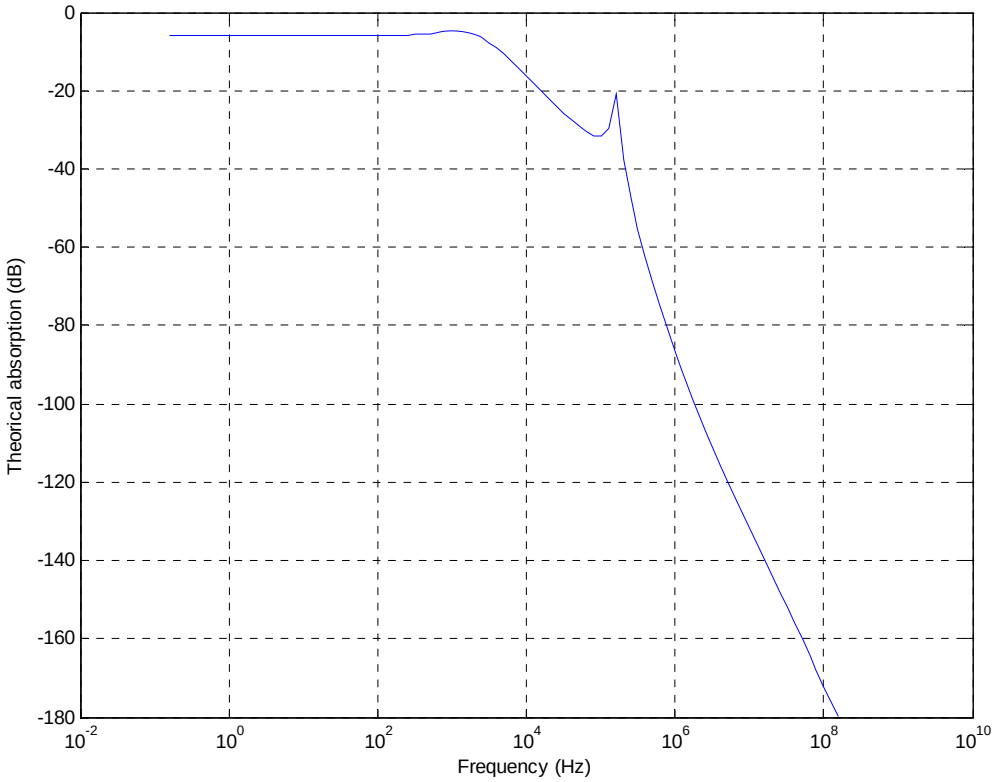

Filter 1: Simple filter connected between the two leads that are connected to the skin conductance device. An identical filter is also used between the leads of the shocking device (C=10pF). Filter positions in the experimental setup are indicated in Figure 6-1 in the box marked “filters”. One filter pair would be located in the top half of the box and one in the lower half of the box. Filter orientation is such that the simulated subject would be on the right and the control system would be on the left.

**Figure 6-3**



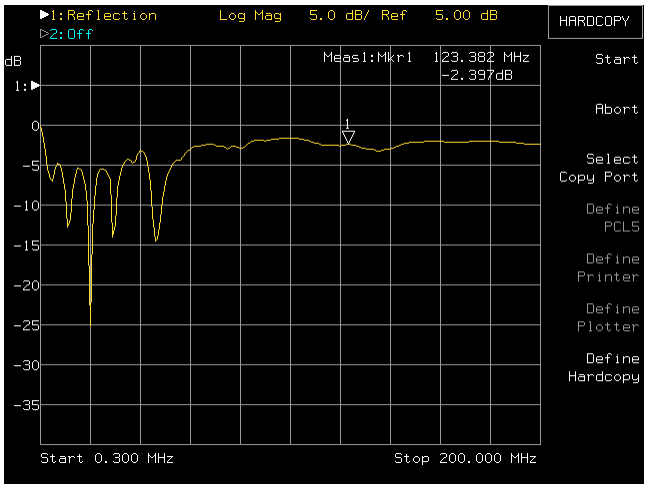

Filter 2: Pi-filter connected between the two leads that are connected to the skin conductance device. An identical filter is also used between the leads of the shocking device. The grounds of both filters are connected together and then connected to an isolated ground on the Psylab box (C = 1nF, R=10kΩ, L=10mH). Filter positions in the experimental setup are indicated in Figure 6-1 in the box marked “filters”. One filter pair would be located in the top half of the box and one in the lower half of the box. Filter orientation is such that the simulated subject would be on the right and the control system would be on the left.

**Figure 6-4**



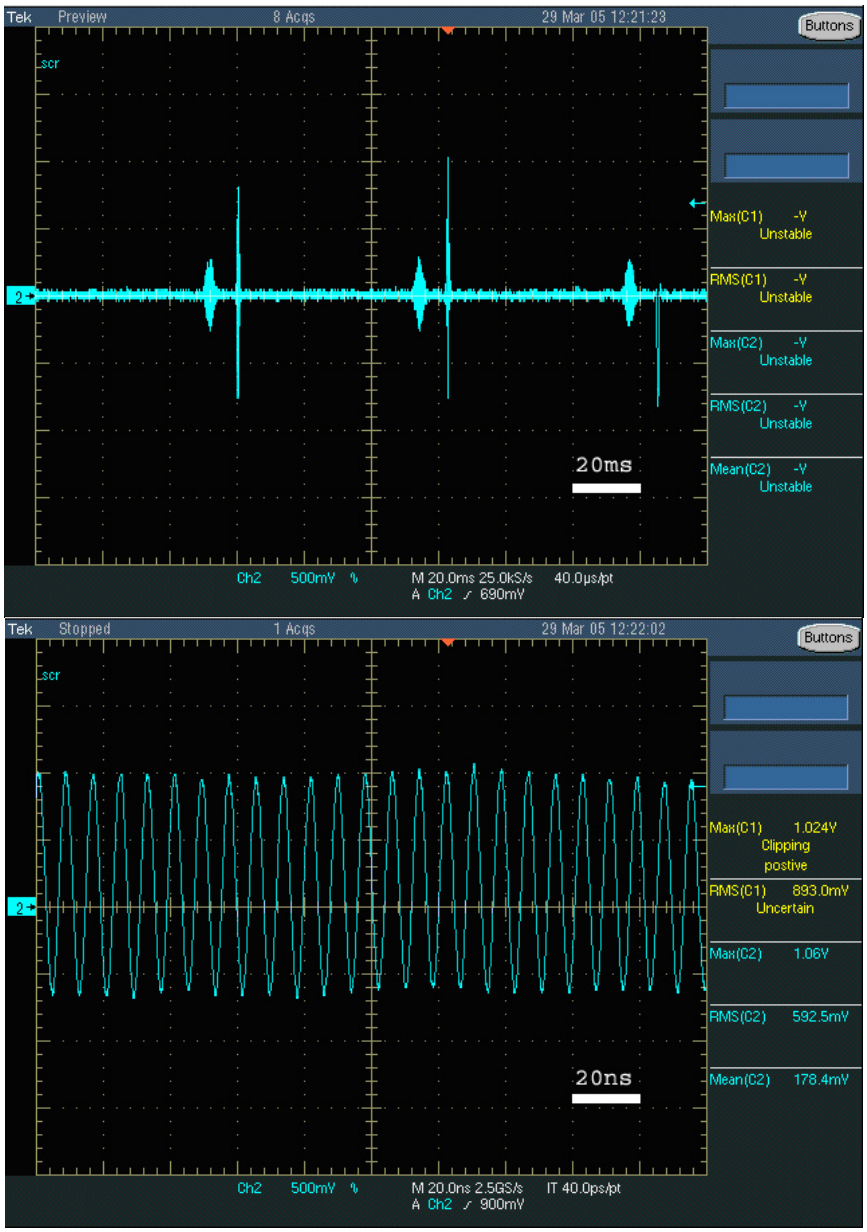

Theoretical response of the type-2 filter

**Figure 6-5**



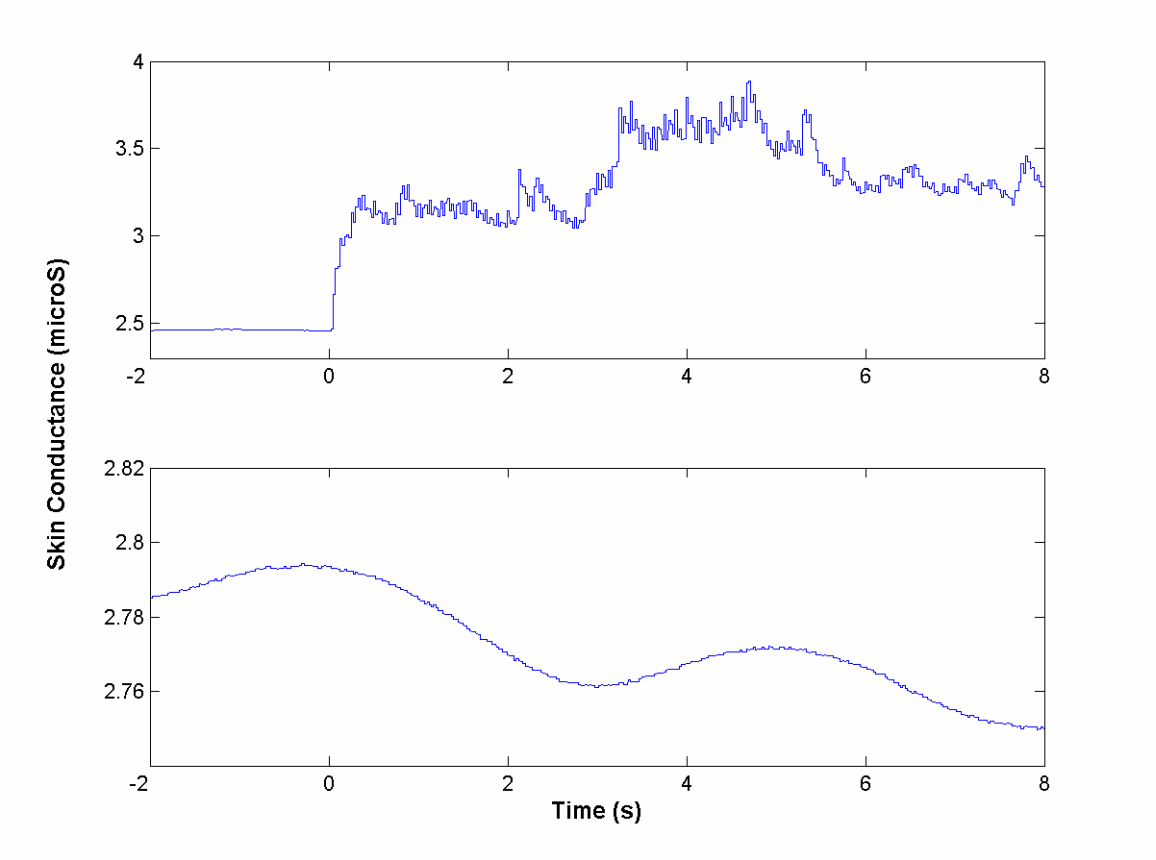

Typical resonance response, GSR with snap leads: The cursor is located at the Larmor frequency (horizontal axis ranging from 0.3 MHz to 200 MHz) and the reflected power is measured in dB (vertical axis ranging from 1 to -40 dB)

**Figure 6-6**




Typical induced voltages on the leads: The top snapshot displays the repetition of a slice acquisition. The bottom snapshot is a time-magnification that shows the oscillations of the magnetic fields during the slice selection pulse exhibited at the Larmor frequency.

Figure 6-7




Skin conductance recording during EPI scanning, using a full carbon fiber electrode configuration with (bottom) and without optimizations (top). Optimizations of the skin conductance trace shown in the lower half included the use of the low pass filter in Figure 6-3, connecting all components to a common ground reference, and attention to placement of components within the control room. The onset of EPI scanning occurs approximately at zero seconds.

# 7 References

Allison T, Ginter H, McCarthy G, Nobre AC, Puce A, Luby M, Spencer DD (1994) Face recognition in human extrastriate cortex. J Neurophysiol 71:821-825.

Asaad WF, Rainer G, Miller EK (2000) Task-specific neural activity in the primate prefrontal cortex. J Neurophysiol 84:451-459.

Baer PE, Fuhrer MJ (1982) Cognitive factors in the concurrent differential conditioning of eyelid and skin conductance responses. Mem Cognit 10:135-140.

Bakin JS, Weinberger NM (1990) Classical conditioning induces CS-specific receptive field plasticity in the auditory cortex of the guinea pig. Brain Res 536:271-286.

Bao S, Chan VT, Merzenich MM (2001) Cortical remodelling induced by activity of ventral tegmental dopamine neurons. Nature 412:79-83.

Bechara A, Tranel D, Damasio H, Adolphs R, Rockland C, Damasio AR (1995) Double dissociation of conditioning and declarative knowledge relative to the amygdala and hippocampus in humans. Science 269:1115-1118.

Beilock SL, Carr TH, MacMahon C, Starkes JL (2002) When paying attention becomes counterproductive: impact of divided versus skill-focused attention on novice and experienced performance of sensorimotor skills. J Exp Psychol Appl 8:6-16.

Boucsein W (1992) Electrodermal Activity. New York: Plenum Press.

Brainard DH (1997) The Psychophysics Toolbox. Spatial Vision 10:433-436.

Buchel C, Dolan RJ (2000) Classical fear conditioning in functional neuroimaging. Curr Opin Neurobiol 10:219-223.

Buchel C, Holmes AP, Rees G, Friston KJ (1998a) Characterizing stimulus-response functions using nonlinear regressors in parametric fMRI experiments. Neuroimage 8:140-148.

Buchel C, Morris J, Dolan RJ, Friston KJ (1998b) Brain systems mediating aversive conditioning: an event-related fMRI study. Neuron 20:947-957.

Buchel C, Dolan RJ, Armony JL, Friston KJ (1999) Amygdala-hippocampal involvement in human aversive trace conditioning revealed through event-related functional magnetic resonance imaging. J Neurosci 19:10869-10876.

Carrillo M, Gabrieli J, Disterhoft J (2000) Selective effects of division of attention on discrimination conditioning. Psychobiology 28:293-302.

Carter RM, Hofstotter C, Tsuchiya N, Koch C (2003) Working memory and fear conditioning. Proc Natl Acad Sci U S A 100:1399-1404.

Carter RM, O'Doherty JP, Seymour B, Koch C, Dolan RJ (2006) Contingency awareness in human aversive conditioning involves the middle frontal gyrus. Neuroimage 29:1007-1012.

Castner SA, Goldman-Rakic PS, Williams GV (2004) Animal models of working memory: insights for targeting cognitive dysfunction in schizophrenia. Psychopharmacology (Berl) 174:111-125.

Chun MM, Phelps EA (1999) Memory deficits for implicit contextual information in amnesic subjects with hippocampal damage. Nat Neurosci 2:844-847.

Clark RE, Squire LR (1998) Classical conditioning and brain systems: the role of awareness. Science 280:77-81.

Clark RE, Squire LR (1999) Human eyeblink classical conditioning: Effects of manipulating awareness of the stimulus contingencies. Psychol Sci 10:14-18.

Clark RE, Manns JR, Squire LR (2001) Trace and delay eyeblink conditioning: contrasting phenomena of declarative and nondeclarative memory. Psychol Sci 12:304-308.

Clark RE, Manns JR, Squire LR (2002) Classical conditioning, awareness, and brain systems. Trends Cogn Sci 6:524-531.

Cole LE (1939) A Comparison of Factors of Practice and Knowledge of Experimental Procedure in Conditioning the Eye-lid Response of Human Subjects. J Gen Psychol 20:349-373.

Colloca L, Benedetti F (2005) Placebos and painkillers: is mind as real as matter? Nat Rev Neurosci 6:545-552.

Compton DM, Griffith HR, McDaniel WF, Foster RA, Davis BK (1997) The flexible use of multiple cue relationships in spatial navigation: a comparison of water maze performance following

hippocampal, medial septal, prefrontal cortex, or posterior parietal cortex lesions. Neurobiol Learn Mem 68:117-132.

Connolly JB, Roberts IJ, Armstrong JD, Kaiser K, Forte M, Tully T, O'Kane CJ (1996) Associative learning disrupted by impaired Gs signaling in Drosophila mushroom bodies. Science 274:2104-2107.

Cook SW, Harris RE (1937) The Verbal Conditioning of the Galvanic Skin Response. J Exp Psychol 21:202-210.

Corlett PR, Aitken MR, Dickinson A, Shanks DR, Honey GD, Honey RA, Robbins TW, Bullmore ET, Fletcher PC (2004) Prediction error during retrospective revaluation of causal associations in humans: fMRI evidence in favor of an associative model of learning. Neuron 44:877-888.

Critchley HD, Mathias CJ, Dolan RJ (2002) Fear conditioning in humans: the influence of awareness and autonomic arousal on functional neuroanatomy. Neuron 33:653-663.

D'Esposito M, Aguirre GK, Zarahn E, Ballard D, Shin RK, Lease J (1998) Functional MRI studies of spatial and nonspatial working memory. Cogn Brain Res 7:1-13.

Damasio AR, Tranel D, Damasio H (1991) Somatic Markers and the Guidance of Behaviour: Theory and Preliminary Testing. In: Frontal Lobe Function and Dysfunction (Levin HS, Eisenberg HM, Benton AL, eds), pp 217-229. New York: Oxford University Press.

Dawson ME, Furedy JJ (1976) The role of awareness in human differential autonomic classical conditioning: the necessary-gate hypothesis. Psychophysiology 13:50-53.

Degonda N, Mondadori CR, Bosshardt S, Schmidt CF, Boesiger P, Nitsch RM, Hock C, Henke K (2005) Implicit associative learning engages the hippocampus and interacts with explicit associative learning. Neuron 46:505-520.

Deichmann R, Gottfried JA, Hutton C, Turner R (2003) Optimized EPI for fMRI studies of the orbitofrontal cortex. Neuroimage 19:430-441.

Dyer AG, Neumeyer C, Chittka L (2005) Honeybee (Apis mellifera) vision can discriminate between and recognise images of human faces. J Exp Biol 208:4709-4714.

Edeline JM, Pham P, Weinberger NM (1993) Rapid development of learning-induced receptive field plasticity in the auditory cortex. Behav Neurosci 107:539-551.

Efron B, Tibshirani RJ (1998) An Introduction to the Bootstrap, 2 Edition. New York: Chapman and Hall.

Eichenbaum H (1997) NEUROSCIENCE: How Does the Brain Organize Memories? Science 277:330-332.

Eichenbaum H, Schoenbaum G, Young B, Bunsey M (1996) Functional organization of the hippocampal memory system. Proc Natl Acad Sci U S A 93:13500-13507.

Fanselow MS (2000) Contextual fear, gestalt memories, and the hippocampus. Behav Brain Res 110:73-81.

Fendt M, Fanselow MS (1999) The neuroanatomical and neurochemical basis of conditioned fear. Neurosci Biobehav Rev 23:743-760.

Friston KJ, Holmes AP, Price CJ, Buchel C, Worsley KJ (1999) Multisubject fMRI studies and conjunction analyses. Neuroimage 10:385-396.

Fritz J, Shamma S, Elhilali M, Klein D (2003) Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. Nat Neurosci 6:1216-1223.

Gallistel CR (1990) The Organization of Learning. Cambridge, MA: Bradford Books, MIT Press.

Gauthier I, Tarr MJ (1997) Becoming a "Greeble" expert: exploring mechanisms for face recognition. Vision Res 37:1673-1682.

Gauthier I, Skudlarski P, Gore JC, Anderson AW (2000) Expertise for cars and birds recruits brain areas involved in face recognition. Nat Neurosci 3:191-197.

Gilboa A, Shalev AY, Laor L, Lester H, Louzoun Y, Chisin R, Bonne O (2004) Functional connectivity of the prefrontal cortex and the amygdala in posttraumatic stress disorder. Biol Psychiatry 55:263-272.

Goldman-Rakic P (1987) Circuitry of primate prefrontal cortex and regulation of behaviour by representational memory. In: Handbook of physiology (Plum F, Mouncastle U, eds), pp 373-417. Washington, DC: The American Physiological Society.

Goldstone RL (1998) Perceptual learning. Annu Rev Psychol 49:585-612.

Grillon C (2002) Startle reactivity and anxiety disorders: aversive conditioning, context, and neurobiology. Biol Psychiatry 52:958-975.

Haijiang Q, Saunders JA, Stone RW, Backus BT (2006) Demonstration of cue recruitment: Change in visual appearance by means of Pavlovian conditioning. PNAS 103:483-488.

Han CJ, O'Tuathaigh CM, van Trigt L, Quinn JJ, Fanselow MS, Mongeau R, Koch C, Anderson DJ (2003) Trace but not delay fear conditioning requires attention and the anterior cingulate cortex. Proc Natl Acad Sci U S A 100:13087-13092.

Haxby JV, Horwitz B, Ungerleider LG, Maisog JM, Pietrini P, Grady CL (1994) The functional organization of human extrastriate cortex: a PET-rCBF study of selective attention to faces and locations. J Neurosci 14:6336-6353.

Henson RN, Shallice T, Gorno-Tempini ML, Dolan RJ (2002) Face repetition effects in implicit and explicit memory tests as measured by fMRI. Cereb Cortex 12:178-186.

Hilgard ER, Campbell RK, Sears WN (1937) Conditioned discrimination: The effect of knowledge of stimulus relationships. Am J Psychol 51:498-506.

Ishai A, Pessoa L, Bikle PC, Ungerleider LG (2004) Repetition suppression of faces is modulated by emotion. PNAS 101:9827-9832.

Kanwisher N, McDermott J, Chun MM (1997) The fusiform face area: a module in human extrastriate cortex specialized for face perception. J Neurosci 17:4302-4311.

Kaube H, Katsarava Z, Kaufer T, Diener HC, Ellrich J (2000) A new method to increase nociception specificity of the human blink reflex. Clin Neurophysiol 111:413-416.

Khurana B, Carter RM, Watanabe K, Nijhawan R (2006) Flash-lag chimeras: The role of perceived alignment in the composite face effect. Vision Res 46:2757-2772.

Kim JJ, Fanselow MS (1992) Modality-specific retrograde amnesia of fear. Science 256:675-677.

Kisley MA, Gerstein GL (2001) Daily variation and appetitive conditioning-induced plasticity of auditory cortex receptive fields. Eur J Neurosci 13:1993-2003.

Knight DC, Smith CN, Stein EA, Helmstetter FJ (1999) Functional MRI of human Pavlovian fear conditioning: patterns of activation as a function of learning. Neuroreport 10:3665-3670.

Knight DC, Cheng DT, Smith CN, Stein EA, Helmstetter FJ (2004) Neural substrates mediating human delay and trace fear conditioning. J Neurosci 24:218-228.

Knuttinen MG, Power JM, Preston AR, Disterhoft JF (2001) Awareness in classical differential eyeblink conditioning in young and aging humans. Behav Neurosci 115:747-757.

Kocorowski LH, Helmstetter FJ (2001) Calcitonin gene-related peptide released within the amygdala is involved in Pavlovian auditory fear conditioning. Neurobiol Learn Mem 75:149-163.

Kronforst-Collins MA, Disterhoft JF (1998) Lesions of the caudal area of rabbit medial prefrontal cortex impair trace eyeblink conditioning. Neurobiol Learn Mem 69:147-162.

LaBar KS, Disterhoft JF (1998) Conditioning, awareness, and the hippocampus. Hippocampus 8:620-626.

LaBar KS, Gatenby JC, Gore JC, LeDoux JE, Phelps EA (1998) Human amygdala activation during conditioned fear acquisition and extinction: a mixed-trial fMRI study. Neuron 20:937-945.

Law E, Nuttley WM, van der Kooy D (2004) Contextual taste cues modulate olfactory learning in C. elegans by an occasion-setting mechanism. Curr Biol 14:1303-1308.

LeDoux J (2003) The emotional brain, fear, and the amygdala. Cell Mol Neurobiol 23:727-738.

Leung HC, Gore JC, Goldman-Rakic PS (2002) Sustained mnemonic response in the human middle frontal gyrus during on-line storage of spatial memoranda. J Cogn Neurosci 14:659-671.

Lovibond PF, Shanks DR (2002) The role of awareness in Pavlovian conditioning: empirical evidence and theoretical implications. J Exp Psychol Anim Behav Process 28:3-26.

Luria AR (1968) The mind of a mnemonist; a little book about a vast memory [by] A. R. Luria. Translated from the Russian by Lynn Solotaroff. With a foreword by Jerome S. Bruner. New York: Basic Books.

Lykken DT (1972) Range correction applied to heart rate and to GSR data. Psychophysiology 9:373-379.

Mackintosh NJ (1983) Conditioning and Associative Learning. Oxford, England: Clarendon Press.

Manns JR, Clark RE, Squire LR (2000a) Awareness predicts the magnitude of single-cue trace eyeblink conditioning. Hippocampus 10:181-186.

Manns JR, Clark RE, Squire LR (2000b) Parallel acquisition of awareness and trace eyeblink classical conditioning. Learn Mem 7:267-272.

Manns JR, Clark RE, Squire LR (2002) Standard delay eyeblink classical conditioning is independent of awareness. J Exp Psychol Anim Behav Process 28:32-37.

Maren S (2001) Neurobiology of Pavlovian fear conditioning. Annu Rev Neurosci 24:897-931.

Maren S, Quirk GJ (2004) Neuronal signalling of fear memory. Nat Rev Neurosci 5:844-852.

Maren S, Aharonov G, Fanselow MS (1997) Neurotoxic lesions of the dorsal hippocampus and Pavlovian fear conditioning in rats. Behav Brain Res 88:261-274.

Mayer MJ, Ross LE (1969) Effects of stimulus complexity, interstimulus interval, and masking task conditions in differential eyelid conditioning. J Exp Psychol 81:469-474.

McEchron MD, Bouwmeester H, Tseng W, Weiss C, Disterhoft JF (1998) Hippocampectomy disrupts auditory trace fear conditioning and contextual fear conditioning in the rat. Hippocampus 8:638-646.

McIntosh AR, Gonzalez-Lima F (1998) Large-scale functional connectivity in associative learning: interrelations of the rat auditory, visual, and limbic systems. J Neurophysiol 80:3148-3162.

McLaughlin J, Skaggs H, Churchwell J, Powell DA (2002) Medial prefrontal cortex and pavlovian conditioning: trace versus delay conditioning. Behav Neurosci 116:37-47.

Medina JF, Christopher Repa J, Mauk MD, LeDoux JE (2002) Parallels between cerebellum- and amygdala-dependent conditioning. Nat Rev Neurosci 3:122-131.

Milad MR, Rauch SL, Pitman RK, Quirk GJ (2006) Fear extinction in rats: Implications for human brain imaging and anxiety disorders. Biol Psychol.

Morris JS, Ohman A, Dolan RJ (1998a) Conscious and unconscious emotional learning in the human amygdala. Nature 393:467-470.

Morris JS, Friston KJ, Dolan RJ (1998b) Experience-dependent modulation of tonotopic neural responses in human auditory cortex. Proc Biol Sci 265:649-657.

Morris JS, Ohman A, Dolan RJ (1999) A subcortical pathway to the right amygdala mediating "unseen" fear. Proc Natl Acad Sci U S A 96:1680-1685.

Morris JS, Buchel C, Dolan RJ (2001) Parallel neural responses in amygdala subregions and sensory cortex during implicit fear conditioning. Neuroimage 13:1044-1052.

Ohman A, Soares JJ (1998) Emotional conditioning to masked stimuli: expectancies for aversive outcomes following nonrecognized fear-relevant stimuli. J Exp Psychol Gen 127:69-82.

Olsson A, Phelps EA (2004) Learned fear of "unseen" faces after Pavlovian, observational, and instructed fear. Psychol Sci 15:822-828.

Pavlov IP (1906) The scientific investigation of the psychical faculties or processes in higher animals. Science 24:613-619.

Pearce JM, Redhead ES, Aydin A (1997) Partial reinforcement in appetitive Pavlovian conditioning with rats. Q J Exp Psychol B 50:273-294.

Pessoa L, Kastner S, Ungerleider LG (2003) Neuroimaging Studies of Attention: From Modulation of Sensory Processing to Top-Down Control. J Neurosci 23:3990-3998.

Pessoa L, Gutierrez E, Bandettini P, Ungerleider L (2002) Neural correlates of visual working memory: fMRI amplitude predicts task performance. Neuron 35:975-987.

Pessoa L, Japee S, Sturman D, Ungerleider LG (2006) Target visibility and visual awareness modulate amygdala responses to fearful faces. Cereb Cortex 16:366-375.

Petrides M (2000) The role of the mid-dorsolateral prefrontal cortex in working memory. Exp Brain Res 133:44-54.

Phillips RG, LeDoux JE (1992) Differential contribution of amygdala and hippocampus to cued and contextual fear conditioning. Behav Neurosci 106:274-285.

Pochon JB, Levy R, Fossati P, Lehericy S, Poline JB, Pillon B, Le Bihan D, Dubois B (2002) The neural system that bridges reward and cognition in humans: an fMRI study. Proc Natl Acad Sci U S A 99:5669-5674.

Poggio T, Fahle M, Edelman S (1992) Fast perceptual learning in visual hyperacuity. Science 256:1018-1021.

Puce A, Allison T, Asgari M, Gore JC, McCarthy G (1996) Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study. J Neurosci 16:5205-5215.

Quinn JJ, Oommen SS, Morrison GE, Fanselow MS (2002) Post-training excitotoxic lesions of the dorsal hippocampus attenuate forward trace, backward trace, and delay fear conditioning in a temporally specific manner. Hippocampus 12:495-504.

Quirk GJ, Gehlert DR (2003) Inhibition of the amygdala: key to pathological states? Ann N Y Acad Sci 985:263-272.

Quirk GJ, Armony JL, LeDoux JE (1997) Fear conditioning enhances different temporal components of tone-evoked spike trains in auditory cortex and lateral amygdala. Neuron 19:613-624.

Quirk GJ, Likhtik E, Pelletier JG, Pare D (2003) Stimulation of medial prefrontal cortex decreases the responsiveness of central amygdala output neurons. J Neurosci 23:8800-8807.

Rankin CH, Beck CD, Chiba CM (1990) Caenorhabditis elegans: a new model system for the study of learning and memory. Behav Brain Res 37:89-92.

Rosenkranz JA, Grace AA (2001) Dopamine attenuates prefrontal cortical suppression of sensory inputs to the basolateral amygdala of rats. J Neurosci 21:4090-4103.

Schendan HE, Searl MM, Melrose RJ, Stern CE (2003) An FMRI study of the role of the medial temporal lobe in implicit and explicit sequence learning. Neuron 37:1013-1025.

Scoville WB, Milner B (1957) Loss of recent memory after bilateral hippocampal lesions. J Neurol Neurosurg Psychiatry 20:11-21.

Seymour B, O'Doherty JP, Koltzenburg M, Wiech K, Frackowiak R, Friston K, Dolan R (2005) Opponent appetitive-aversive neural processes underlie predictive learning of pain relief. Nat Neurosci 8:1234-1240.

Shamma SA (2004) Topographic organization is essential for pitch perception. Proc Natl Acad Sci U S A 101:1114-1115.

Shellock FG (2000a) Magnetic Resonance Procedures: Health Effects and Safety. New York: CRC Press.

Shellock FG (2000b) Radiofrequency energy-induced heating during MR procedures: a review. J Magn Reson Imaging 12:30-36.

Shuler MG, Bear MF (2006) Reward timing in the primary visual cortex. Science 311:1606-1609.

Slimko EM, McKinney S, Anderson DJ, Davidson N, Lester HA (2002) Selective electrical silencing of mammalian neurons in vitro by the use of invertebrate ligand-gated chloride channels. J Neurosci 22:7373-7379.

Soares JJ, Ohman A (1993) Backward masking and skin conductance responses after conditioning to nonfeared but fear-relevant stimuli in fearful subjects. Psychophysiology 30:460-466.

Sotres-Bayon F, Cain CK, Ledoux JE (2006) Brain Mechanisms of Fear Extinction: Historical Perspectives on the Contribution of Prefrontal Cortex. Biol Psychiatry.

Squire LR, Kandel ER (1999) From Mind to Molecules. New York, New York: Scientific American Library, Freeman.

Tarr MJ, Gauthier I (2000) FFA: a flexible fusiform area for subordinate-level visual processing automatized by expertise. Nat Neurosci 3:764-769.

Thompson RF, Krupa DJ (1994) Organization of memory traces in the mammalian brain. Annu Rev Neurosci 17:519-549.

Tully T (1998) Toward a molecular biology of memory: the light's coming on! Nat Neurosci 1:543-545.

Tully T, Quinn WG (1985) Classical conditioning and retention in normal and mutant Drosophila melanogaster. J Comp Physiol [A] 157:263-277.

Wallis JD, Anderson KC, Miller EK (2001) Single neurons in prefrontal cortex encode abstract rules. Nature 411:953-956.

Watanabe T, Nanez JE, Koyama S, Mukai I, Liederman J, Sasaki Y (2002) Greater plasticity in lower-level than higher-level visual motion processing in a passive perceptual learning task. Nat Neurosci 5:1003-1009.

Weible AP, McEchron MD, Disterhoft JF (2000) Cortical involvement in acquisition and extinction of trace eyeblink conditioning. Behav Neurosci 114:1058-1067.

Weinberger NM (2004) Specific long-term memory traces in primary auditory cortex. Nat Rev Neurosci 5:279-290.

Weiskrantz L, Warrington EK, Sanders MD, Marshall J (1974) Visual capacity in the hemianopic field following a restricted occipital ablation. Brain 97:709-728.

Wen JY, Kumar N, Morrison G, Rambaldini G, Runciman S, Rousseau J, van der Kooy D (1997) Mutations that prevent associative learning in C. elegans. Behav Neurosci 111:354-368.

Wiens S, Ohman A (2002) Unawareness is more than a chance event: comment on Lovibond and Shanks (2002). J Exp Psychol Anim Behav Process 28:27-31.

Wilensky AE, Schafe GE, LeDoux JE (1999) Functional inactivation of the amygdala before but not after auditory fear conditioning prevents memory formation. J Neurosci 19:art. no.-RC48.

Wunderlich K, Schneider KA, Kastner S (2005) Neural correlates of binocular rivalry in the human lateral geniculate nucleus. Nat Neurosci 8:1595-1602.

Yang TT, Gallen C, Schwartz B, Bloom FE, Ramachandran VS, Cobb S (1994) Sensory maps in the human brain. Nature 368:592-593.