

Inactivation of Primate Prefrontal Cortex Impairs Auditory and Audiovisual Working Memory

Bethany Plakke, Jaewon Hwang, and Lizabeth M. Romanski

University of Rochester School of Medicine and Dentistry, Department of Neurobiology and Anatomy, Rochester, New York 14642

The prefrontal cortex is associated with cognitive functions that include planning, reasoning, decision-making, working memory, and communication. Neurophysiology and neuropsychology studies have established that dorsolateral prefrontal cortex is essential in spatial working memory while the ventral frontal lobe processes language and communication signals. Single-unit recordings in nonhuman primates has shown that ventral prefrontal (VLPFC) neurons integrate face and vocal information and are active during audiovisual working memory. However, whether VLPFC is essential in remembering face and voice information is unknown. We therefore trained nonhuman primates in an audiovisual working memory paradigm using naturalistic face-vocalization movies as memoranda. We inactivated VLPFC, with reversible cortical cooling, and examined performance when faces, vocalizations or both faces and vocalization had to be remembered. We found that VLPFC inactivation impaired subjects' performance in audiovisual and auditory-alone versions of the task. In contrast, VLPFC inactivation did not disrupt visual working memory. Our studies demonstrate the importance of VLPFC in auditory and audiovisual working memory for social stimuli but suggest a different role for VLPFC in unimodal visual processing.

Key words: monkey; multisensory; visual working memory; vocalization

Significance Statement

The ventral frontal lobe, or inferior frontal gyrus, plays an important role in audiovisual communication in the human brain. Studies with nonhuman primates have found that neurons within ventral prefrontal cortex (VLPFC) encode both faces and vocalizations and that VLPFC is active when animals need to remember these social stimuli. In the present study, we temporarily inactivated VLPFC by cooling the cortex while nonhuman primates performed a working memory task. This impaired the ability of subjects to remember a face and vocalization pair or just the vocalization alone. Our work highlights the importance of the primate VLPFC in the processing of faces and vocalizations in a manner that is similar to the inferior frontal gyrus in the human brain.

Introduction

The idea that the ventral frontal lobe plays a pivotal role in communication harkens back to the time of Paul Broca's work on patients suffering from aphasia. Functional neuroimaging studies have broadened our understanding of the ventral frontal lobe, confirming its involvement in language, and highlighting its role in working memory (WM), attention, decision-making, and reward processing (Burns and Fahy, 2010; D'Esposito and Postle,

2015). With its widespread afferent input and its established role in speech and gestures, the ventral lobe is perfectly poised to integrate sensory, motor and limbic information during communication. The inferior frontal gyrus (IFG) is part of a larger network, including areas in the temporal lobe, which selectively respond to faces (Bruce et al., 1981; Perrett et al., 1982; Desimone et al., 1984; Baylis et al., 1987; Tsao et al., 2006; Tsao et al., 2008) and voices (Belin et al., 2000; Petkov et al., 2008; Perrodin et al., 2011). Indeed, the IFG is active during the joint processing and integration of face or gestural information and vocal stimuli (von Kriegstein et al., 2005; von Kriegstein and Giraud, 2006; Xu et al., 2009). Neurophysiological recordings have shown that single cells in the macaque ventrolateral prefrontal cortex (VLPFC), an area considered homologous with the human IFG, respond to complex sounds, including species-specific vocalizations (Romanski and Goldman-Rakic, 2002; Romanski et al., 2005; Cohen et al., 2007; Russ et al., 2008b) and are selectively responsive to faces (O'Scalaidhe et al., 1997; Scalaidhe et al., 1999; Romanski and Diehl, 2011). Furthermore, single VLPFC neurons integrate

Received March 27, 2015; revised May 21, 2015; accepted May 27, 2015.

Author contributions: B.P. and L.M.R. designed research; B.P. and J.H. performed research; B.P. analyzed data; B.P. and L.M.R. wrote the paper.

This work was supported by National Institutes of Health Grant DC04845, Training and Hearing Balance and Spatial Orientation Grant DC009974, Schmitt Program on Integrative Brain Research, and the Center for Visual Science. We thank William E. O'Neill and Chris Petkov for comments on the manuscript.

The authors declare no competing financial interests.

Correspondence should be addressed to Dr. Bethany Plakke, University of Rochester School of Medicine, Department of Neurobiology and Anatomy, Rochester, NY 14642. E-mail: Bethany_plakke@urmc.rochester.edu.

DOI:10.1523/JNEUROSCI.1218-15.2015

Copyright © 2015 the authors 0270-6474/15/359666-10\$15.00/0

faces and vocalizations (Sugihara et al., 2006), detect mismatched face and vocalization pairs (Diehl and Romanski, 2014), and maintain face-vocalization information in WM (Hwang and Romanski, 2015). Together, it suggests that this evolutionarily conserved frontal lobe region may be involved in communication in both humans and animals. We must then ask whether VLPFC is essential to the process of integrating and/or remembering face and vocal information.

The essential role of the prefrontal cortex (PFC) in WM and sensory discrimination has previously been assessed with lesions in a variety of paradigms. Lesions limited to the dorsolateral prefrontal cortex (DLPFC) impair visuospatial WM (Goldman et al., 1971; Funahashi et al., 1993), whereas larger lesions of the lateral PFC disrupt auditory discrimination (Gross and Weiskrantz, 1962; Iversen and Mishkin, 1973), rule learning (Rygula et al., 2010), and some forms of decision making (Baxter et al., 2009). However, the effects of VLPFC lesions on WM have been equivocal with some showing impairments (Passingham, 1975; Mishkin and Manning, 1978), whereas other studies describe only mild effects on performance (Rushworth et al., 1997; Bussey et al., 2002). Importantly, the effect of selective lesions of VLPFC on the processing and remembering of auditory or combined face and vocal stimuli has not been carefully examined. To determine whether VLPFC is essential in audiovisual (AV) WM, we used a modified form of reversible cortical cooling. This method is similar to that developed and used by Lomber and Malhotra (2008) and Hussein et al. (2014) to dissociate function in auditory cortical areas, and has been used in both human and nonhuman primates (Chafee and Goldman-Rakic, 2000; Bakken et al., 2003). We inactivated VLPFC while nonhuman primates performed a nonmatch-to-sample AV WM task (AV NMTS) where a face and vocalization pair had to be remembered. Because PFC has been tied to mnemonic processing and VLPFC neurons encode and integrate face and vocal information, we hypothesized that inactivation of the VLPFC would interfere with AV WM. We further assessed performance using simpler variants of our task requiring only faces or only vocalizations as the remembered stimuli. Interestingly, we found that VLPFC inactivation disrupted AV and auditory WM but surprisingly did not affect WM when only faces had to be remembered. This has important implications for the role of VLPFC in communication and WM.

Materials and Methods

The research subjects were two rhesus monkeys (*Macaca mulatta*; a 5-year-old male, 8.0 kg; and an 8-year-old female, 6.8 kg) previously trained on a delayed NMTS task. All methods were in accordance with National Institutes of Health standards and were approved by the University of Rochester Care and Use of Research Animals committee. Before training, a titanium head post was surgically implanted under general anesthesia on the skull. After completion of training, recording cylinders were implanted over VLPFC (centered ~31–32 mm anterior to the interaural line and 20–21 mm lateral to midline on skull, at an angle of 25°–30° from the vertical to maximize an orthogonal approach to VLPFC, areas 12/47 and 45 (Preuss and Goldman-Rakic, 1991; Petrides and Pandya, 2002; Romanski and Goldman-Rakic, 2002) (see Fig. 1). Although the cooling of the entire chamber included a small portion of DLPFC (area 46), the majority of the cortical surface that was inactivated was within the inferior convexity and included areas 12/47 and 45 where previous studies have documented neuronal responses to face and vocal stimuli (Romanski et al., 2005; Sugihara et al., 2006). An MR image of the location of Subject 1's titanium chamber is shown in Figure 1D, and the lateral view of Subject 2's brain is shown in Figure 1C.

Apparatus and stimuli. The apparatus and stimuli are identical to that used in Hwang and Romanski (2015). All training and cooling sessions were performed in a sound-attenuated room, lined with Sonex (Acous-

tical Solutions). AV movie clips were presented to the subjects on a computer monitor (NEC MultiSync LCD1830, 1280 × 1024, 60 Hz), which was at 75 cm distance from the monkey's eyes. Speakers (Yamaha MSP5; frequency response, 50 Hz to 40 kHz) were located on either side of the monitor at the height of the subject's head. The auditory stimuli ranged from 65 to 80 dB SPL (35–112 mPa) measured at the level of the monkey's ear with a B and K sound level meter. The NMTS task used dynamic movie clips of rhesus macaques vocalizing, as dynamic movies enhance discrimination (Ghazanfar et al., 2005; Chandrasekaran et al., 2013). These movie clips are part of a library of movies of monkeys filmed in our home colony. The video was captured at 30 fps with a size of 320 × 240 pixels (6.8° × 5.1° in visual angle), and the audio was recorded with a 48 kHz sampling rate and 16-bit resolution. The digitized movies were processed with Corel Video Studio (Corel), VirtualDub (www.virtualdub.org), and GoldWave software (GoldWave). The length of the video tracks was 467–1367 ms (892 ms on average), and that of the audio ranged from 145 to 672 ms (308 ms on average). The mismatching components (face or vocalization) were inserted so that there was no asynchrony between them (Romanski, 2012). Mismatching stimuli were chosen to be incongruent in vocalization type, making them easy to discriminate.

Movies were presented at eye level. Eye position was continuously monitored using an ISCAN infrared pupil monitoring system (ISCAN). Behavioral data (eye position and button press) was collected on a PC via PCI interface boards (NI PCI-6220 and NI PCI-6509; National Instruments). The timing of stimulus presentation and reward delivery was controlled with in-house C++ software, which was built based upon Microsoft DirectX Technologies.

Cooling apparatus and procedure. We inactivated VLPFC while subjects performed the NMTS task by cooling the cortical surface of VLPFC <20°C. The cortical surface was cooled by inserting a sealed, stainless-steel 18.5 mm doughnut-shaped chamber with a 0.01 cm hole in the center into the 19 mm recording cylinder, on top of the dural surface. During cooling sessions, ice-cold ethanol (−40°C to −70°C) was circulated into the cooling chamber to cool the cortex beneath the dural surface and decrease synaptic activity (see Fig. 1). Brain temperature was measured by lowering a hypodermic temperature micro probe (Omega) through the hole in the center of the cooling chamber. The target temperature range was 15°C–20°C at a depth of 3 mm from the cortical surface. A temperature probe was also affixed to the bottom of the cooling chamber, which measured the temperature at the surface of the dura. Temperature of the circulating ethanol and dural surface were continuously monitored throughout the task.

Experimental procedure. Each day, the monkeys were brought to the testing room and prepared for behavioral training or a cooling session. The recording cylinders were cleaned and lidocaine was applied to the dura for 10 min and then rinsed with saline. The cooling chamber and temperature probes were then lowered into the recording cylinders, and room temperature ethanol was circulated through the cooling chambers. To begin a trial, animals were required to fixate on a central point for 500 ms, then the sample stimulus was presented followed by the delay period (1.25 s). On half of the trials, trial Type 1, the second stimulus presented after the delay was a nonmatching stimulus (auditory or visual). On the other half of the trials, trial Type 2, the sample stimulus and delay period were repeated (match stimulus) before the nonmatch occurred. Randomization of these two types of trials made the presentation of the nonmatch stimulus unpredictable. Subjects were required to detect the nonmatch with a button press to receive a juice reward, which occurred 0.5 s after a correct button press. Subjects maintained their gaze within a large window that held the movie stimulus (10 × 10 degrees) throughout the entire trial. Eye movements outside this window terminated the trial. All trial conditions (auditory or visual nonmatch, trial Type 1, 2) were presented in a pseudo-random fashion and counterbalanced across trials.

There were 20 cooling sessions for each subject conducted using the AV NMTS task as described in Experiment 1, where half of the presentations were an audio nonmatch and the other half consisted of video nonmatch. We assessed performance accuracy and reaction time in the auditory and visual nonmatch trials of the AV NMTS task during WARM

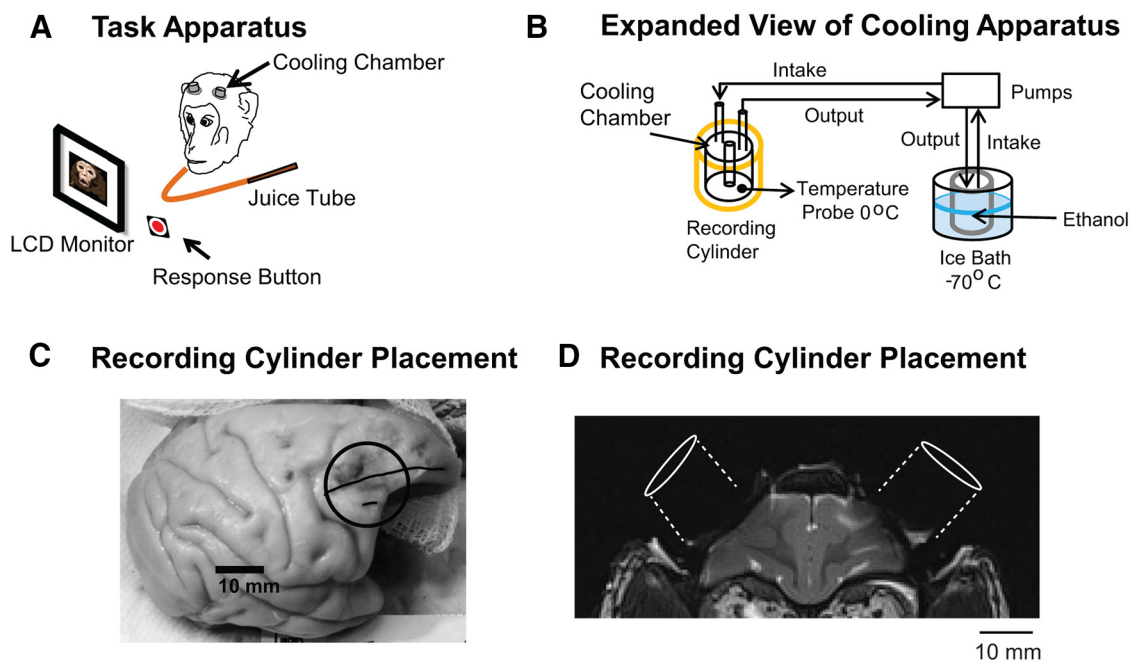


Figure 1. Experimental setup. **A**, The monitor, response button, juice tube, position of the subject, and the bilateral recording cylinders, which were positioned over the PFC, are shown. **B**, Expanded schematic of the cooling system. The cooling chambers were placed inside the recording cylinders on the subject's head and lie apposed directly to the dural surface overlying VLPFC. Cooled ethanol is pumped through tubing into the cooling chambers. The temperature of the dura is measured with a temperature probe fixed to the bottom of the cooling chamber. Cortical temperature is measured by a different temperature probe inserted through a hole in the center of the cooling chambers into the cortex at a depth of 3 mm. **C**, Lateral view of the postperfusion brain of Subject 2, with the cylinder location shown by a black circle. Black represents the location of the principal sulcus. Black circle represents the cortical region that was inactivated by cooling on both sides in this subject, as confirmed by presence of dye markers placed into this region before perfusion. **D**, A coronal MRI section through the region of the VLPFC, which was inactivated in Subject 1. The titanium recording cylinders (which are not visible) cause a shadow and image distortion on the prefrontal cortical surface. Therefore, the region, which was cooled in Subject 1, is approximated by the dotted white outline of the inner edge of the titanium recording cylinders.

and COLD trials. Only sessions where 100 WARM and 100 COLD trials were completed were used in the analysis. Experiments 2 and 3 were done to examine only visual (Experiment 2) (10 sessions for one subject, 11 sessions for the second subject) or only auditory (Experiment 3) (10 sessions for each subject) WM. These experiments followed the same cooling procedure as the AV NMTS, but in Experiment 2 only the face was switched on nonmatch trials and in Experiment 3 only the vocalization component was changed. During the visual-only nonmatching face sessions of Experiment 2, the interstimulus interval delay was lengthened to 3 s to increase the memory requirement and to make the difficulty of the task comparable to the AV NMTS.

To control for possible order effects, 15 control sessions (Experiment 4) for each subject were run, where a warm session was followed by another warm session (WARM-WARM). In these sessions, the cooling chambers were lowered into the recording cylinders, and only room temperature ethanol was run through the system for the entire session. The first warm 100 trials were equivalent to the time frame of the "WARM" sessions in the previous experiment, and the second block of 100 warm trials was equivalent to the "COLD" time period. The number of trials and timing in between the two blocks, in this case WARM and WARM, were identical to the other experiments where temperature was manipulated.

Data analysis. Performance of the subjects, measured by percentage correct (number of correct trials/total number of answered trials, per session) was analyzed. Errors occurred as a "miss" when subjects did not press the button to the nonmatching stimulus on trial Type 1, or as a "wrong press" when subjects pressed to the match stimulus on trial Type 2. Overall percentage correct values for both animals were assessed with a one-way ANOVA using SPSS 22 software (SPSS) with temperature as a factor. To further examine the differences on trial Types 1 and 2, a three-way ANOVA (temperature, modality of nonmatch, trial type) was performed on percentage correct values across sessions for each subject during Experiment 1, the AV cooling experiment. A similar three-way ANOVA was performed on Experiment 4 (WARM, WARM Sessions)

with factors of time period, modality, and trial type. For any significant interaction, a *post hoc* Bonferroni test was computed. To examine trial type effects in the visual only nonmatch (Experiment 2) and the audio only nonmatch (Experiment 3) experiments, a two-way ANOVA (temperature by trial type) was performed. For all experiments, reaction time latency was measured from the onset of the nonmatch stimulus and was compared between the COLD and WARM trials with *t* tests (Bonferroni corrected by the number of comparisons for each condition), for each trial type (correct auditory Type 1 trials, correct visual Type 1 trials, etc.). In the AV experiment, the reaction time *p* value was set to 0.004; for the visual-only and auditory-only experiments, the *p* value was set to 0.016. The percentage of lost fixation trials from the WARM and COLD trials was analyzed with *z* tests, as we were comparing proportions ($p = 0.05$). There were no significant effects of cooling or modality on reaction time values; thus, reaction times were collapsed across temperature and modality factors and then overall values were compared between correct Type 1, correct Type 2, and false alarm error Type 2 trials ($p = 0.05$). Because the third stimulus is always the nonmatch on Type 2 trials and is predictable, the nonmatch reaction time on Type 2 trials is typically faster.

Results

Experiment 1: Inactivation of PFC impairs audiovisual working memory

To examine the role of VLPFC on AV WM, we reversibly inactivated VLPFC in two macaques while they each performed a delayed non-match-to-sample (NMTS) task. We cooled prefrontal cortex and reduced synaptic activity in the target region by lowering the temperature of the cortex $<20^\circ$. The apparatus to cool the cortical surface is shown in Figure 1 and consisted of a stainless-steel, sealed chamber, which was inserted into the recording cylinder where it lay flat against the dura over VLPFC and the ventral edge of DLPFC. Each daily session began by assessing

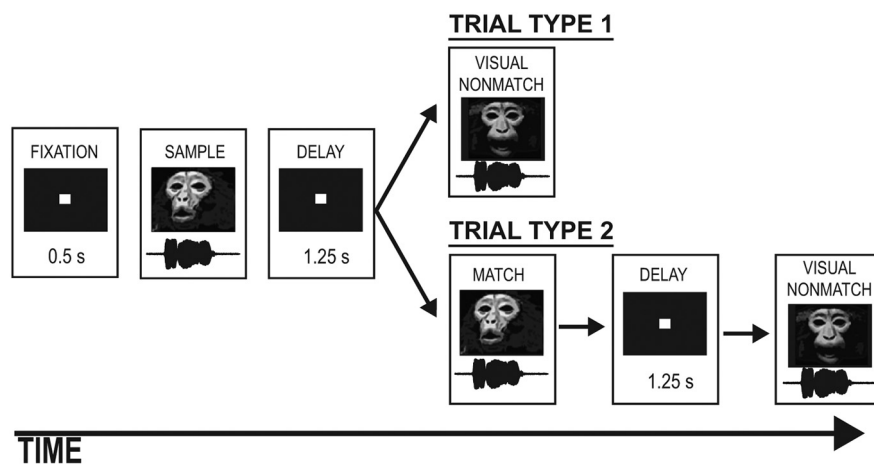


Figure 2. Schematic of the AV NMTS task. A vocalization movie (with an audio and video track) was presented as the sample stimulus, and the subject was required to remember the auditory and visual components (vocalization and accompanying facial gesture) and to detect the change of either the face or vocalization component in subsequent stimulus presentations with a button press. On Type 1 trials, in half the trials, the nonmatch occurred following the sample and delay period. On Type 2 trials, a matching stimulus intervened and the nonmatch occurred as the third stimulus. In the example shown, the face in the second stimulus of Type 1 trials and the third stimulus of Type 2 trials does not match the sample (visual nonmatch). This occurred for 50% of the trials, and for the other half of trials the audio track was altered and an incongruent vocalization replaced the sample audio track (auditory nonmatch; data not shown).

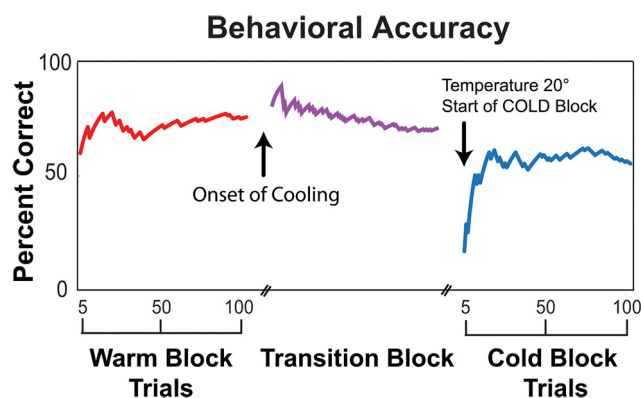


Figure 3. Data from one cooling session showing behavioral accuracy as percentage correct (calculated as cumulative performance, trial by trial, for each block) by block over time. Red line indicates behavioral performance averaging ~75% correct during the WARM block. Purple line indicates when the cooling process was initiated. Blue line indicates performance when the brain reached a steady temperature and the COLD block began, where accuracy dropped to slightly >50% correct during the COLD block. Each line starts on trial 5 of the block.

performance in the AV WM task as the cortex remained at, or near, normal body temperature and room temperature ethanol (23°C) (WARM trials) was circulated through the cooling chambers. After 100 trials were completed at normal temperature, we circulated cold ethanol through the cooling chambers to cool the cortex and inactivate VLPFC (Fig. 1). The monkeys continued to perform the task while the cortical temperature dropped toward the target range, which took ~10 min to reach. After reaching the cold target range, 100 COLD trials were recorded. Performance for each daily training session was assessed by comparing accuracy during 100 trials completed at the beginning of each session before cooling (WARM trials) to the 100 trials when the temperature had reached the target range of 15°C–20°C (COLD trials).

In Experiment 1, animals performed the AV NMTS task that we have used previously (Hwang and Romanski, 2015). Subjects fixated a spot on the monitor for 500 ms, and then an AV sample movie was presented. All of our AV movies consisted of a short

movie clip of a rhesus macaque vocalizing. The AV sample was followed by one of two trial types. In trial Type 1, a 1.25 s delay occurred followed by a nonmatching AV stimulus (Fig. 2). In trial Type 2, the sample stimulus was repeated (the match stimulus), followed by a second delay period (1.25 s), which was then followed by a nonmatching AV stimulus (Fig. 2). The subjects were required to press the response button when a nonmatching stimulus occurred. Each of the trial types occurred 50% of the time, which made the nonmatch stimulus unpredictable during the second stimulus presentation. Because our stimuli consisted of a complex AV stimulus, the nonmatch could be a change in the auditory track (auditory nonmatch) or a change in the visual track of the sample movie (visual nonmatch) (Fig. 2). The button press response was required to occur within the duration of the movie plus 900 ms to receive a juice reward. Half the trials were auditory nonmatch trials (i.e., the non-

match stimulus consisted of video track that matched the sample stimulus but an audio track that differed from the track in sample movie). The other half of the trials were visual nonmatch trials where the video track differed from the sample but the audio (vocalization) remained unchanged. Hence, subjects had to remember both the face and the vocalization that were presented as the sample to correctly detect the unpredictable nonmatching component.

We performed 20 cooling sessions in each of the subjects. As the cortex cooled to the target range, performance accuracy also decreased (Fig. 3). We assessed performance accuracy during the WARM and COLD sessions and found that bilateral cooling of VLPFC impaired performance in the AV NMTS task in both subjects (Fig. 4). Percentage correct values for performance across all trial types for each subject were assessed with a two-way ANOVA (factors were temperature and modality). In both subjects, we found a significant effect of temperature (Fig. 4) (Subject 1: $F_{(1,76)} = 12.820$, $p < 0.05$; Subject 2: $F_{(1,76)} = 30.595$, $p < 0.05$) such that percentage correct during COLD trials was significantly lower than percentage correct during WARM trials. For Subject 1, there was also a significant effect of modality ($F_{(1,76)} = 27.483$, $p < 0.05$), which indicated worse performance on auditory nonmatch trials compared with visual nonmatch trials in both WARM and COLD trials. Because our task involves two cognitive processes, actively responding when a nonmatching stimulus is detected (trial Type 1) and withholding a response when a matching stimulus is detected (trial Type 2), we analyzed the data for each animal with trial type as an additional factor (three-way ANOVA, temperature, modality, trial type). In both animals, there was a main effect of temperature (Fig. 5) (Subject 1: $F_{(1,152)} = 13.889$, $p < 0.01$; Subject 2: $F_{(1,152)} = 46.524$, $p < 0.01$) demonstrating a decrease in performance during COLD trials. In Subject 1, there was a significant effect of modality ($F_{(1,152)} = 29.774$, $p < 0.01$) where performance was worse on auditory nonmatch trials. This subject also had a significant interaction between modality and trial type ($F_{(1,152)} = 41.820$, $p < 0.01$). *Post hoc* tests (Bonferroni, $p < 0.05$) indicated that performance for this subject was worse on auditory nonmatch trials com-

AUDIOVISUAL WORKING MEMORY PERFORMANCE

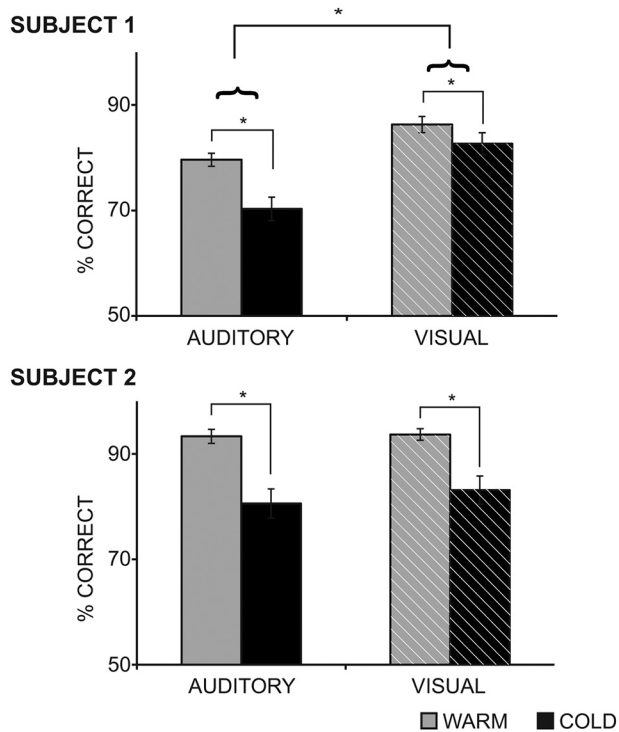


Figure 4. VLPFC inactivation by cooling significantly impairs overall performance accuracy on the AV NMTS task (Experiment 1). * $p < 0.05$. Mean performance was calculated from 20 testing sessions for each subject with 100 trials (WARM, normal temperature before cooling) and 100 trials during cooling of VLPFC (COLD). Gray bars represent WARM trials. Black bars represent COLD trials (during cooling of the VLPFC). Solid color bars represent performance during auditory component nonmatch trials. Striped bars represent performance during visual component nonmatches. For both subjects, there was a significant decrease in performance during cooling (COLD trials, black bars) compared with the control period (WARM trials, gray bars) for both auditory and visual nonmatch trials. In addition, Subject 1 demonstrated significantly worse performance on auditory trials compared with visual trials. Error bars indicate the SEM calculated across 20 sessions per subject.

pared with visual nonmatch trials during Type 1 trials, when the nonmatch was the second stimulus and required a button press (Fig. 5).

Effect of prefrontal cooling on reaction time and fixation

It is possible that impaired performance during cooling could be due to the cooling procedure, which could induce a slowing of motor responses or an inability to maintain gaze on the task stimuli. We therefore examined reaction times and fixation during the task. We found that there was no significant difference between response time during WARM or COLD trials (t tests, $p > 0.05$). We then examined the reaction times by trial type (for further details, see Materials and Methods). Overall average reaction time for Subject 1 on correct Type 1 trials was 1086.2 ms, on correct Type 2 trials it was 899.3 ms, and on false alarms it was 1192.2 ms. For Subject 2 on correct Type 1 trials reaction time was 1195.4 ms, on correct Type 2 trials it was 1104.2 ms, and on false alarms it was 1234.2 ms. For Subject 1 there was a significant difference between reaction times on correct Type 2 trials and false alarm trials ($p < 0.05$), indicating slower reaction time on error trials.

Because each trial is initiated when the subject fixates a red square for 500 ms and then maintains gaze within a viewing

window that includes the movie stimulus, inability to maintain gaze within the viewing window would cause the trial to abort. The ability to maintain fixation can be interpreted as a general measure of attention, which could also explain impairment during VLPFC cooling. Lost fixation trials were indeed rare and most frequently occurred during the first fixation period just before the onset of the sample AV movie, unrelated to any specific modality nonmatch or trial type. z tests indicated that there was no significant difference between lost fixation trials during WARM trials and COLD trials for Subject 2 (warm, 11.5%; cold, 21.5%), but there was a significant effect for Subject 1, with more aborted fixation trials occurring during cooling ($p < 0.05$; warm, 8.6%; cold, 15.8%).

Experiment 2: Prefrontal inactivation does not impair visual working memory

In our version of the AV NMTS task, subjects must encode both the face and vocal stimulus and remember them throughout the delay intervals to correctly detect the presence of a nonmatching face or voice. This is more difficult and involves a greater memory load than typical visual WM tasks due to the need to encode and maintain 2 complex stimuli, of different modalities, in mind. We investigated visual memory performance by simplifying our task and only changing the video track of the AV movie. Thus, subjects only had to remember the face presented during the AV sample, instead of both the face and the vocalization from the sample movie. We hypothesized that this easier task might not engage VLPFC and would not result in impairment during cooling.

Subjects were retrained with the same set of AV vocalization movies, but only the visual component was changed in the nonmatching stimulus. Even though the sample is bimodal, the subjects quickly learn that only the visual-face stimulus needs to be remembered for the correct response. To make this paradigm similar to previously published visual WM studies (Rushworth et al., 1997; Chafee and Goldman-Rakic, 2000) and to equate the level of difficulty with that of the AV NMTS task, we lengthened the delay period to 3 s. We performed the cortical cooling sessions in the same manner as above, collecting data while subjects performed this “visual-only” version of the NMTS task for 100 trials at normal temperature (WARM) and then cooling the prefrontal cortex to $<20^{\circ}\text{C}$ and collecting 100 COLD trials. We found that cooling VLPFC did not significantly affect performance in the “visual-only” NMTS task in either subject (Fig. 6A); one-way ANOVA (temperature) did not show a difference in performance accuracy for either subject between WARM (average percentage correct) and COLD (average percentage correct) trials (Subject 1: $F_{(1,19)} = 1.454$, $p = 0.244$; Subject 2: $F_{(1,20)} = .414$, $p = 0.527$). A two-way ANOVA (temperature \times trial type) also did not reveal any significant difference in performance (Subject 1: no effect of temperature, $F_{(1,36)} = 1.911$, $p = 0.175$; no effect of trial type, $F_{(1,36)} = 2.803$, $p = 0.103$; no interaction, $F_{(1,36)} = 0.046$, $p = 0.831$; Subject 2: no effect of temperature, $F_{(1,40)} = .650$, $p = 0.425$; no effect of trial type, $F_{(1,40)} = .031$, $p = 0.860$; no interaction, $F_{(1,40)} = 0.175$, $p = 0.678$).

There were also no significant effects of cooling on reaction times. Overall average reaction time for Subject 1 on correct Type 1 trials was 1069.9 ms, on correct Type 2 trials it was 894.1 ms, and on false alarms, which occur on Type 2 trials, it was 1214.2 ms. For Subject 2 on correct Type 1 trials reaction time was 1231.1 ms, and on correct Type 2 trials it was 1290.8 ms, and on false alarms it was 1526.7 ms. Overall reaction time between correct trials and false alarm trials were not significantly different for

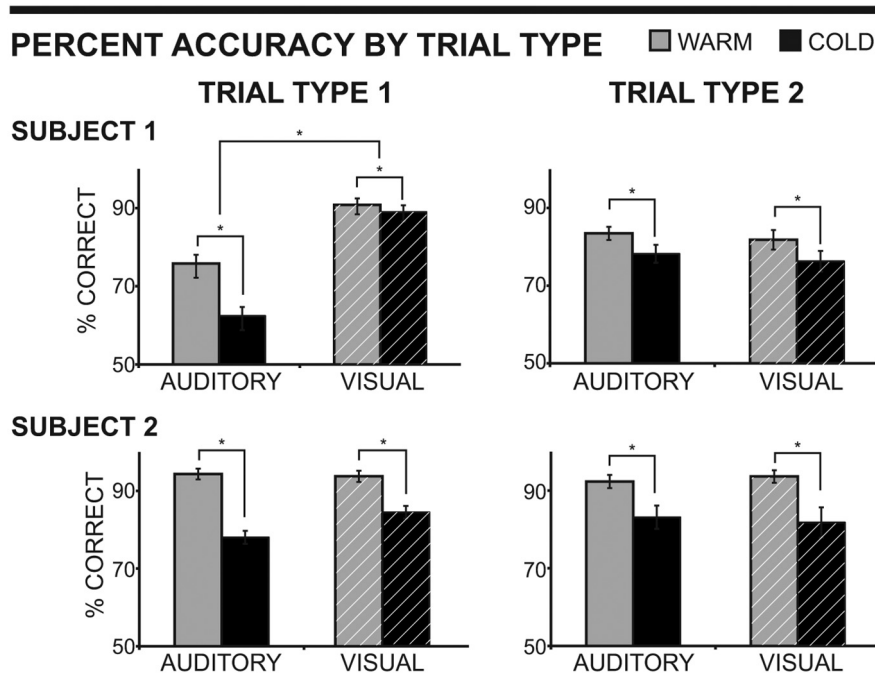


Figure 5. Percentage accuracy by trial type during AV WM. In both subjects, cooling of the prefrontal cortex resulted in a significant decrease in performance across trial types. A significant interaction between modality and trial type in Subject 1 demonstrated significantly worse performance on auditory Type 1 trials compared with visual trials. $p < 0.05$. Colors and shading as in Figure 4. Error bars indicate the SEM calculated across 20 sessions per subject.

VISUAL ONLY- WORKING MEMORY

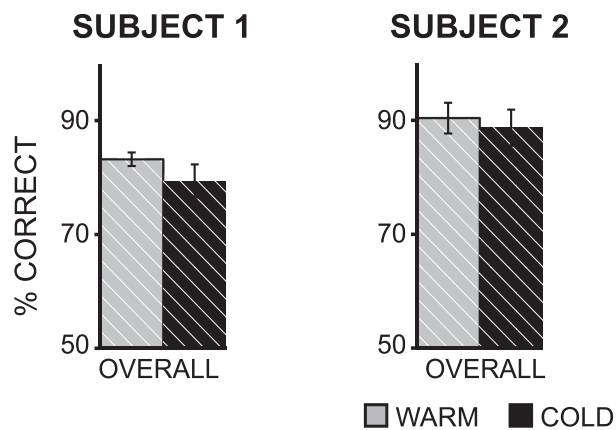


Figure 6. Inactivation of VLPFC does not significantly affect visual-only WM performance. In Experiment 2 during the visual-only condition, there was no significant change in performance accuracy during cooling of the VLPFC (COLD trials, black bars) compared with the precooling trials (WARM, gray bars) in either subject. Error bars indicate SEM. Mean performance was calculated across 10 sessions (Subject 1) and 11 sessions (Subject 2).

Subject 2. Overall reaction time between correct Type 2 and false alarm error Type 2 trials were significantly different for Subject 1 ($p < 0.05$), with values indicating slower responses on error trials, which is consistent with previous experiments. It is important to remember that on Type 2 trials the third stimulus is always the nonmatch and is predictable; thus, reaction time is typically faster on correct trial Type 2.

The overall percentage of lost fixation trials, where subjects gaze traveled outside fixation window, for Subject 1 was as follows: warm, 10.5%; cold, 18.9%; and for Subject 2 as fol-

lows: warm: 5.8% cold: 12.7%. There was a significant increase in lost fixation trials in COLD trials compared with WARM trials for Subject 1 ($p < 0.05$), similar to that in Experiment 1.

Experiment 3: Prefrontal inactivation impairs auditory working memory

Given that cooling did not affect performance in our version of a “unimodal visual WM task,” we asked whether VLPFC inactivation would affect an “auditory-only” version of the NMTS task. In contrast to visual discrimination performance, auditory WM performance is typically weaker and takes longer to train in NHPs. Given the results in the AV version of our task, we hypothesized that cooling of VLPFC would impair auditory-only WM. Animals were given a second re-training period on the same task. In this variant of the task, during the nonmatch only the auditory component, the vocalization track of the movie stimulus, would differ from the sample. As in Experiment 2, the sample stimulus was the bimodal facial gesture-vocalization movie, but in this experiment, the nonmatch was always a change in the auditory track where a different, incongruent but synchronous, vocalization was inserted (see Materials and Methods). It was necessary to use the same delay period as the AV version of the task, 1.25 s, to allow for a baseline performance of 70%–80% correct before cooling. As we predicted, transient inactivation of VLPFC by cooling resulted in a significantly impaired performance in each animal. A one-way ANOVA (temperature) on the overall performance in the task for each animal revealed a significant decrease in performance accuracy during COLD trials compared with performance in WARM trials (Fig. 7; Subject 1: $F_{(1,19)} = 18.549$, $p < 0.05$; Subject 2: $F_{(1,19)} = 5.662$, $p < 0.05$). A two-way ANOVA (temperature \times trial type) for both animals showed a significant effect of temperature (Subject 1: $F_{(1,39)} = 30.479$, $p < 0.05$; Subject 2: $F_{(1,39)} = 7.460$, $p < 0.05$). In addition, for Subject 2, there was also a significant main effect of trial type ($F_{(1,39)} = 10.599$, $p < 0.05$), which indicated greater impairment on Type 2 trials, where it was necessary to withhold responding during the match stimulus.

There was no difference in reaction time latencies between the warm and cold trials; therefore, reaction time data were collapsed across temperature conditions (warm/cold) to examine any differences in reaction during the different trial types. Overall average reaction time for Subject 1 on correct Type 1 trials was 1154.1 ms, on correct Type 2 trials it was 903.2 ms, and on false alarm errors on Type 2 trials it was 1230.4 ms. For Subject 2 on correct Type 1 trials reaction time was 1376.9 ms, on correct Type 2 trials it was 1338.0 ms, and on false alarm errors on Type 2 trials it was 1343.2 ms. We found a significant difference for Subject 1 between false alarm errors on Type 2 trials and correct Type 2 trials ($p < 0.05$) consistent with earlier results demonstrating slower responses on error trials overall (regardless of temperature).

The overall percentage of lost fixation trials for Subject 1 was 7.4% (warm) and 36.9% (cold); and for Subject 2 was 4.9% (warm) and 8.4% (cold). Inactivation of VLPFC by cooling also

AUDITORY ONLY- WORKING MEMORY

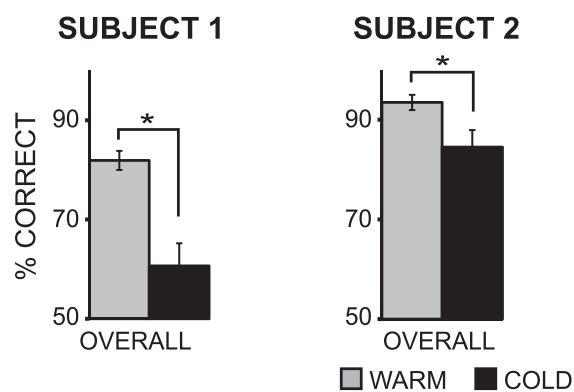


Figure 7. Inactivation of VLPFC by cooling significantly decreases auditory-only WM performance. In Experiment 3 during the auditory-only condition, there was a significant decrease in performance for both subjects during the COLD trials (black bars) compared with precooling performance (WARM, gray bars). Error bars indicate SEM. Mean performance was calculated across 10 testing sessions per subject. $*p < 0.05$.

resulted in a significant increase in aborted fixation trials for Subject 1 ($p < 0.05$) in this auditory alone task, just as it did in Experiments 1 and 2.

Experiment 4: Effect of order on NMTS performance

In each of the previous experiments (AV, visual-only, auditory-only), we completed 100 trials at normal temperature (WARM) before cooling to gauge baseline performance and then cooled prefrontal cortex and assessed performance for 100 trials with the temperature at or $<20^{\circ}\text{C}$. Because subjects often reached their satiation for juice after completing the COLD trials, it was not always possible to evaluate the postcooling return to baseline performance. Because the COLD trials followed the WARM trials, it is possible that any decrease in performance was simply due to fatigue during the second block of trials. To test this, we ran 15 AV NMTS sessions in each subject where there was no change in temperature across the two 100 trial blocks and waited the same amount of time in between blocks, as in Experiment 1. The cooling chambers were lowered into the recording cylinders, room temperature ethanol was circulated through the system, but no temperature change was introduced across the two blocks (WARM-WARM). We assessed the effect of block order performance by comparing the first 100 trials (Early) to the second set of 100 trials (Late) corresponding to when the COLD trials took place. A two-way ANOVA (time period, modality) found no significant effect of early and late blocks on performance of trials for either subject (Fig. 8). For Subject 1, there was a significant effect of modality $F_{(1,56)} = 34.26$ ($p < 0.05$), indicating that visual performance was better than auditory performance across both time periods. A three-way ANOVA (time period, trial type, modality) indicated that for Subject 1 there was a significant effect of modality and an interaction between time period and trial type, as well as a modality by trial type interaction. *Post hoc* analysis indicated an increase in performance on Type 2 trials in the later block of trials and an improvement on visual nonmatch trials during Type 1 trials, compared with auditory nonmatch trials. For Subject 2, there was a significant effect of modality indicating better performance on visual trials compared with auditory trials overall. These results are consistent with the performance during the WARM trials of the previous experiment and indicate that

AUDIOVISUAL PERFORMANCE WITHOUT COOLING

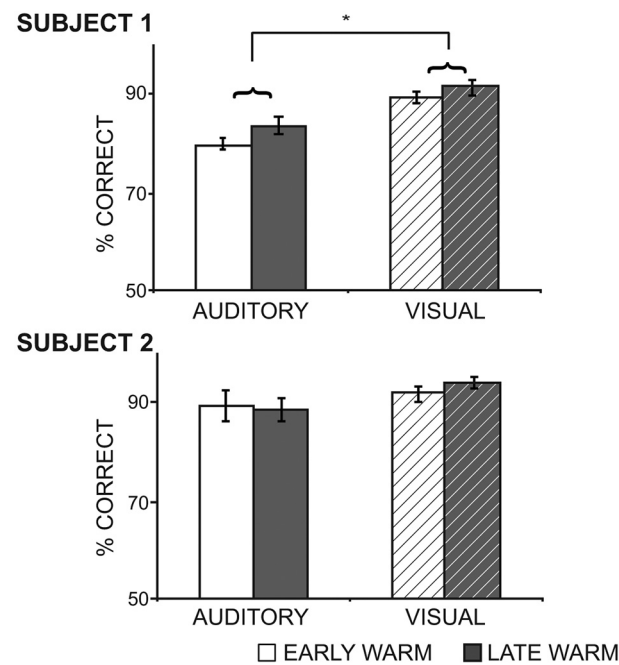


Figure 8. AV performance accuracy without cooling (no temperature change). In Experiment 4, where the temperature was not changed in early and late trial blocks (WARM-WARM), there was no significant difference in performance between the early warm and late warm time periods. Solid color bars represent performance during auditory component nonmatch trials. Striped bars represent performance during visual component nonmatches. For Subject 1, there was an effect of modality indicating better visual performance compared with auditory performance across both time periods. $*p < 0.05$. White bars represent early warm trials. Gray bars represent late warm trials. Error bars indicate SEM. Mean performance was calculated across 15 sessions per subject.

decreases during the second block of trials were not due to temporal order, fatigue, or lack of motivation.

There was no significant difference in reaction time from the early block to the late block of trials. We found a significant difference for Subject 1 between false alarm errors on Type 2 trials and correct Type 2 trials ($p < 0.05$) consistent with earlier results demonstrating slower responses on error trials overall (regardless of block). The overall percentage of lost fixation trial was low for both animals and did not differ significantly from the early to the late period. Overall average reaction time for Subject 1 on correct Type 1 trials was 1125.6 ms, on correct Type 2 trials it was 1010.3 ms, and for false alarm errors it was 1286.8 ms. For Subject 2 on correct Type 1 trials reaction time was 1288.6 ms, on correct Type 2 trials it was 1152.3 ms, and for false alarm errors it was 1766.6 ms. The overall percentage of lost fixation trial was low for both animals (Subject 1 early warm, 10.7%; late warm, 10.9%; Subject 2 early warm, 4.5%; late warm 2, 7.3%).

Discussion

We have demonstrated that the prefrontal cortex is essential in AV WM because inactivation of the VLPFC significantly impaired the ability of nonhuman primates to encode and recall a face and vocalization movie. The fact that inactivation of VLPFC impaired processing when the memoranda was an AV movie but did not affect performance when only the face from the movie needed to be recalled, underscores the role of VLPFC in multi-

sensory memory and integration. Importantly, inactivation of VLPFC also impaired performance in the auditory-alone WM version of the task, when subjects only had to detect a change in the vocalization component of a face-vocalization movie during cortical cooling. At first glance, it might appear as if both of the inactivation-induced impairments might be due to incorrect performance on the auditory trials in both the AV and the auditory-only versions of the task. However, as Figure 4 illustrates, during the AV version of the WM task, detection of both the auditory and the visual nonmatch stimuli was impaired during VLPFC inactivation. To explain the effects on both auditory and AV WM, we postulate that, when unimodal auditory WM uses complex stimuli and requires VLPFC, this mnemonic process alone taxes the capacity of WM. Adding a second memoranda to encode a face could exceed memory load capacity because auditory WM is already using these stores. In humans, WM capacity and load were correlated when subjects performed a visual-auditory detection task (Yu et al., 2014); additionally, auditory and visual memory can interfere with one another (Saults and Cowan, 2007). Hence, VLPFC may be necessary for both auditory and AV WM, especially in dual tasks.

Results from Experiments 2 and 4 rule out the possibility that the performance deficits during cooling were due to fatigue during the second testing block. There was no change in performance in the second block of trials when VLPFC was cooled during visual-only WM (Experiment 2) or during the 2-block sequential WARM-WARM sessions (Experiment 4). This suggests that the decreased performance in the AV and auditory-only paradigms were not due to general effects of cooling on motor responses, loss of attention, fatigue, or loss of motivation to perform the task. For one subject, there was a slight increase in lost fixation trials during VLPFC inactivation, which could indicate a loss of attentional control, a process that has been relegated to PFC (Rossi et al., 2009). It is also interesting to consider that subjects made different types of errors after VLPFC inactivation. In general, Subject 2 made fewer errors than Subject 1 but made more wrong press or “False Alarm” errors, when a match stimulus occurred and it was necessary to withhold a button press until the nonmatch occurred. In contrast, Subject 1 had more errors on Type 1 trials and failed to detect a nonmatching stimulus when it occurred immediately after the sample (“Missed press”). Both types of errors indicate that subjects reverted to a simpler cognitive strategy when the task became difficult during VLPFC inactivation and demonstrated their specific impairment during the critical decision period, when it was necessary to compare the second face-vocalization movie to the sample movie and decide “match” or “nonmatch.”

It is clear that the human IFG is active during auditory detection/discrimination, auditory WM, as well as during phonological, syntactic, and semantic operations (Zatorre et al., 1994; Klein et al., 1995; Stromswold et al., 1996; Buchanan et al., 2000; Caplan et al., 2000; Burton et al., 2003; Waters et al., 2003; Fecteau et al., 2005; Strand et al., 2008). Nonetheless, attempts to disambiguate the cellular processes and the precise role that the ventral frontal lobe plays in complex auditory processing and WM are lacking. A few studies demonstrated decreases in performance accuracy after large lateral PFC lesions in nonhuman primates during delay (Goldman and Rosvold, 1970) and auditory discrimination tasks (Gross and Weiskrantz, 1962; Gross, 1963; Iversen and Mishkin, 1973). Lesions or inactivation of temporal lobe regions have been shown to impair sound localization and auditory pattern discrimination (Lomber and Malhotra, 2008) as well as short-term auditory memory (Fritz et al., 2005). However, the present study

is the first to show that selective inactivation of VLPFC impairs auditory and AV WM. Single-unit recordings in macaques are predictive of the results in the present study because neurons across several regions of the lateral PFC are driven by a variety of visual stimuli, whereas in VLPFC there is a specific auditory responsive region (Romanski and Goldman-Rakic, 2002; Romanski et al., 2005; Russ et al., 2008b). These neurons are robustly responsive to complex sounds, including species-specific vocalizations (Romanski et al., 2005; Cohen et al., 2006; Plakke et al., 2013a) and are active during AV WM (Hwang and Romanski, 2015). Cooling these neurons, in the present study, disrupted auditory WM. Neurons in other areas of the lateral PFC are active in auditory decision-making paradigms (Bodner et al., 1996; Cohen et al., 2006, 2009b; Russ et al., 2008a) and show delay activity in auditory WM tasks (Plakke et al., 2013b). Whereas the area that we cooled consists of mostly VLPFC, the lower edge of DLPFC was also present in the chamber and was cooled. Further experiments are needed to delineate the precise roles of prefrontal regions in cognitive processes.

Importantly, the same VLPFC region that is responsive to faces and their corresponding vocalizations also shows evidence of AV integration (Sugihara et al., 2006). These VLPFC multisensory neurons detect mismatching or incongruent face-vocalization pairs (Diehl and Romanski, 2014) or asynchronous face-vocalization stimuli (Romanski and Hwang, 2012). Recordings in VLPFC during AV NMTS indicate that single VLPFC neurons retain both face and vocalization information during WM (Hwang and Romanski, 2015). This observation predicted that VLPFC inactivation might disrupt AV WM performance. Our findings in the macaque also substantiate data regarding the multisensory nature of WM in the human ventral frontal lobe. fMRI studies show activation of the human IFG during perceptual fusion of face and vocal stimuli, during learning about cross-modal stimuli, and when incongruent AV objects are presented (Miller and D’Esposito, 2005; von Kriegstein and Giraud, 2006; Hein et al., 2007; Naumer et al., 2009; Adam and Noppeney, 2010). The IFG likely plays a crucial role in recognition and is active when accumulating multisensory evidence to make a categorical decision (Noppeney et al., 2010), and when remembering faces or voices (Rama and Courtney, 2005). The impairment observed in the current study on both auditory and visual trials, in the AV version of our task, could be due in part, to the partial shutdown of this multisensory network, which is crucial during memory, recognition, and communication.

While the present results highlight the role of VLPFC in auditory and AV WM, they suggest that the VLPFC may not be essential in visual WM. It is possible that these findings are due in part to differences in difficulty between auditory and visual memory. This is consistent with the fact that even humans have a harder time recognizing voices compared with faces (Legge et al., 1984) and have worse memory performance for auditory recognition compared with visual recognition (Buckner et al., 1996; Cohen et al., 2009a; Bigelow and Poremba, 2014). In previous studies, lesions of the lateral PFC have resulted in WM deficits in some instances, when large lesions of lateral PFC increased errors during a visual delayed matching task (Passingham, 1975; Mishkin and Manning, 1978). However, a later study using smaller lesions in VLPFC demonstrated initial impaired color discrimination but no impairment at longer delays (Rushworth et al., 1997). Additional studies have demonstrated minor or no impairment on visual WM tasks when VLPFC was lesioned (Kowalska et al., 1991; Rushworth et al., 1997; Bussey et al., 2002). In contrast, inactivation of DLPFC decreases performance in visual

WM and cognitive control paradigms (Fuster et al., 1985; Chafee and Goldman-Rakic, 2000; Hussein et al., 2014). It is possible that DLPFC regions, which were not inactivated in the present study, play an essential role in some types of visual WM, whereas VLPFC may be required for auditory WM. Interestingly, inactivation of the human IFG with transcranial magnetic stimulation was shown to impair memory for faces on a difficult continuous visual WM paradigm (Lee and D'Esposito, 2012) and during selective attention for color (Zanto et al., 2011). In both studies, IFG deactivation was associated with impaired performance when there were high attentional or memory demands. In the current study, when memory load was increased by requiring maintenance of a cross-modal stimulus (i.e., remembering both a face and vocalization), VLPFC was required. Data from the present study, as well as previous lesion and inactivation studies in humans and NHPs, support the hypothesis that the ventral frontal lobe may be involved in visual processing for complex tasks, such as when memory load is large or under conditions of cross-modal memory. Consequently, our combined data indicate that VLPFC is essential for cross-modal WM as well as auditory WM, both of which are crucial components of AV communication.

References

- Adam R, Noppeney U (2010) Prior auditory information shapes visual category-selectivity in ventral occipito-temporal cortex. *Neuroimage* 52:1592–1602. [CrossRef Medline](#)
- Bakken HE, Kawasaki H, Oya H, Greenlee JD, Howard MA 3rd (2003) A device for cooling localized regions of human cerebral cortex: technical note. *J Neurosurg* 99:604–608. [CrossRef Medline](#)
- Baxter MG, Gaffan D, Kyriazis DA, Mitchell AS (2009) Ventrolateral prefrontal cortex is required for performance of a strategy implementation task but not reinforcer devaluation effects in rhesus monkeys. *Eur J Neurosci* 29:2049–2059. [CrossRef Medline](#)
- Baylis GC, Rolls ET, Leonard CM (1987) Functional subdivisions of the temporal lobe neocortex. *J Neurosci* 7:330–342. [Medline](#)
- Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B (2000) Voice-selective areas in human auditory cortex. *Nature* 403:309–312. [CrossRef Medline](#)
- Bigelow J, Poremba A (2014) Achilles' ear? inferior human short-term and recognition memory in the auditory modality. *PLoS One* 9:e89914. [CrossRef Medline](#)
- Bodner M, Kroger J, Fuster JM (1996) Auditory memory cells in dorsolateral prefrontal cortex. *Neuroreport* 7:1905–1908. [CrossRef Medline](#)
- Bruce C, Desimone R, Gross CG (1981) Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *J Neurophysiol* 46:369–384. [Medline](#)
- Buchanan TW, Lutz K, Mirzazade S, Specht K, Shah NJ, Zilles K, Jäncke L (2000) Recognition of emotional prosody and verbal components of spoken language: an fMRI study. *Brain Res Cogn Brain Res* 9:227–238. [CrossRef Medline](#)
- Buckner RL, Raichle ME, Miezin FM, Petersen SE (1996) Functional anatomic studies of memory retrieval for auditory words and visual pictures. *J Neurosci* 16:6219–6235. [Medline](#)
- Burns MS, Fahy J (2010) Broca's area: rethinking classical concepts from a neuroscience perspective. *Top Stroke Rehabil* 17:401–410. [CrossRef Medline](#)
- Burton H, Diamond JB, McDermott KB (2003) Dissociating cortical regions activated by semantic and phonological tasks: an fMRI study in blind and sighted people. *J Neurophysiol* 90:1965–1982. [CrossRef Medline](#)
- Bussey TJ, Wise SP, Murray EA (2002) Interaction of ventral and orbital prefrontal cortex with inferotemporal cortex in conditional visuomotor learning. *Behav Neurosci* 116:703–715. [CrossRef Medline](#)
- Caplan D, Alpert N, Waters G, Olivieri A (2000) Activation of Broca's area by syntactic processing under conditions of concurrent articulation. *Hum Brain Mapp* 9:65–71. [CrossRef Medline](#)
- Chafee MV, Goldman-Rakic PS (2000) Inactivation of parietal and prefrontal cortex reveals interdependence of neural activity during memory-guided saccades. *J Neurophysiol* 83:1550–1566. [Medline](#)
- Chandrasekaran C, Lemus L, Ghazanfar AA (2013) Dynamic faces speed up the onset of auditory cortical spiking responses during vocal detection. *Proc Natl Acad Sci U S A* 110:E4668–E4677. [CrossRef Medline](#)
- Cohen YE, Hauser MD, Russ BE (2006) Spontaneous processing of abstract categorical information in the ventrolateral prefrontal cortex. *Biol Lett* 2:261–265. [CrossRef Medline](#)
- Cohen YE, Theunissen F, Russ BE, Gill P (2007) Acoustic features of rhesus vocalizations and their representation in the ventrolateral prefrontal cortex. *J Neurophysiol* 97:1470–1484. [Medline](#)
- Cohen MA, Horowitz TS, Wolfe JM (2009a) Auditory recognition memory is inferior to visual recognition memory. *Proc Natl Acad Sci U S A* 106:6008–6010. [CrossRef Medline](#)
- Cohen YE, Russ BE, Davis SJ, Baker AE, Ackelson AL, Nitecki R (2009b) A functional role for the ventrolateral prefrontal cortex in non-spatial auditory cognition. *Proc Natl Acad Sci U S A* 106:20045–20050. [CrossRef Medline](#)
- Desimone R, Albright TD, Gross CG, Bruce C (1984) Stimulus-selective properties of inferior temporal neurons in the macaque. *J Neurosci* 4:2051–2062. [Medline](#)
- D'Esposito M, Postle BR (2015) The cognitive neuroscience of working memory. *Annu Rev Psychol* 66:115–142. [CrossRef Medline](#)
- Diehl MM, Romanski LM (2014) Responses of prefrontal multisensory neurons to mismatching faces and vocalizations. *J Neurosci* 34:11233–11243. [CrossRef Medline](#)
- Fecteau S, Armony JL, Joanette Y, Belin P (2005) Sensitivity to voice in human prefrontal cortex. *J Neurophysiol* 94:2251–2254. [CrossRef Medline](#)
- Fritz J, Mishkin M, Saunders RC (2005) In search of an auditory engram. *Proc Natl Acad Sci U S A* 102:9359–9364. [CrossRef Medline](#)
- Funahashi S, Bruce CJ, Goldman-Rakic PS (1993) Dorsolateral prefrontal lesions and oculomotor delayed-response performance: evidence for mnemonic "scotomas." *J Neurosci* 13:1479–1497.
- Fuster JM, Bauer RH, Jervey JP (1985) Functional interactions between inferotemporal and prefrontal cortex in a cognitive task. *Brain Res* 330:299–307. [CrossRef Medline](#)
- Ghazanfar AA, Maier JX, Hoffman KL, Logothetis NK (2005) Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. *J Neurosci* 25:5004–5012. [CrossRef Medline](#)
- Goldman PS, Rosvold HE (1970) Localization of function within the dorsolateral prefrontal cortex of the rhesus monkey. *Exp Neurol* 27:291–304. [CrossRef Medline](#)
- Goldman PS, Rosvold HE, Vest B, Galkin TW (1971) Analysis of the delayed-alternation deficit produced by dorsolateral prefrontal lesions in the rhesus monkey. *J Comp Physiol Psychol* 77:212–220. [CrossRef Medline](#)
- Gross CG (1963) A comparison of the effects of partial and total lateral frontal lesions on test performance by monkeys. *J Comp Physiol Psychol* 56:41–47. [CrossRef](#)
- Gross CG, Weiskrantz L (1962) Evidence for dissociation of impairment on auditory discrimination and delayed response following lateral frontal lesions in monkeys. *Exp Neurol* 5:453–476. [CrossRef Medline](#)
- Hein G, Doehrmann O, Müller NG, Kaiser J, Muckli L, Naumer MJ (2007) Object familiarity and semantic congruency modulate responses in cortical audiovisual integration areas. *J Neurosci* 27:7881–7887. [CrossRef Medline](#)
- Hussein S, Johnston K, Belbeck B, Lomber SG, Everling S (2014) Functional specialization within macaque dorsolateral prefrontal cortex for the maintenance of task rules and cognitive control. *J Cogn Neurosci* 26:1918–1927. [CrossRef Medline](#)
- Hwang J, Romanski LM (2015) Prefrontal neuronal responses during audiovisual mnemonic processing. *J Neurosci* 35:960–971. [CrossRef Medline](#)
- Iversen SD, Mishkin M (1973) Comparison of superior temporal and inferior prefrontal lesions on auditory and non-auditory tasks in rhesus monkeys. *Brain Res* 55:355–367. [CrossRef Medline](#)
- Klein D, Milner B, Zatorre RJ, Meyer E, Evans AC (1995) The neural substrates underlying word generation: a bilingual functional-imaging study. *Proc Natl Acad Sci U S A* 92:2899–2903. [CrossRef Medline](#)
- Kowalska DM, Bachevalier J, Mishkin M (1991) The role of the inferior prefrontal convexity in performance of delayed nonmatching-to-sample. *Neuropsychologia* 29:583–600. [CrossRef Medline](#)
- Lee TG, D'Esposito M (2012) The dynamic nature of top-down signals orig-

- inating from prefrontal cortex: a combined fMRI-TMS study. *J Neurosci* 32:15458–15466. [CrossRef Medline](#)
- Legge GE, Grossmann C, Pieper CM (1984) Learning unfamiliar voices. *J Exp Psychol* 10:298–303.
- Lomber SG, Malhotra S (2008) Double dissociation of ‘what’ and ‘where’ processing in auditory cortex. *Nat Neurosci* 11:609–616. [CrossRef Medline](#)
- Miller LM, D’Esposito M (2005) Perceptual fusion and stimulus coincidence in the cross-modal integration of speech. *J Neurosci* 25:5884–5893. [CrossRef Medline](#)
- Mishkin M, Manning FJ (1978) Non-spatial memory after selective prefrontal lesions in monkeys. *Brain Res* 143:313–323. [CrossRef Medline](#)
- Namer MJ, Doebrmann O, Müller NG, Muckli L, Kaiser J, Hein G (2009) Cortical plasticity of audio-visual object representations. *Cereb Cortex* 19:1641–1653. [CrossRef Medline](#)
- Noppeney U, Ostwald D, Werner S (2010) Perceptual decisions formed by accumulation of audiovisual evidence in prefrontal cortex. *J Neurosci* 30:7434–7446. [CrossRef Medline](#)
- O’Scalaidhe SP, Wilson FA, Goldman-Rakic PS (1997) Areal segregation of face-processing neurons in prefrontal cortex. *Science* 278:1135–1138. [CrossRef Medline](#)
- Passingham R (1975) Delayed matching after selective prefrontal lesions in monkeys (*Macaca mulatta*). *Brain Res* 92:89–102. [CrossRef Medline](#)
- Perrett DI, Rolls ET, Caan W (1982) Visual neurones responsive to faces in the monkey temporal cortex. *Exp Brain Res* 47:329–342. [CrossRef Medline](#)
- Perrodin C, Kayser C, Logothetis NK, Petkov CI (2011) Voice cells in the primate temporal lobe. *Curr Biol* 21:1408–1415. [CrossRef Medline](#)
- Petkov CI, Kayser C, Steudel T, Whittingstall K, Augath M, Logothetis NK (2008) A voice region in the monkey brain. *Nat Neurosci* 11:367–374. [CrossRef Medline](#)
- Petrides M, Pandya DN (2002) Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and cortico-cortical connection patterns in the monkey. *Eur J Neurosci* 16:291–310. [CrossRef Medline](#)
- Plakke B, Diltz MD, Romanski LM (2013a) Coding of vocalizations by single neurons in ventrolateral prefrontal cortex. *Hear Res* 305:135–143. [CrossRef Medline](#)
- Plakke B, Ng CW, Poremba A (2013b) Neural correlates of auditory recognition memory in primate lateral prefrontal cortex. *Neuroscience* 244:62–76. [CrossRef Medline](#)
- Preuss TM, Goldman-Rakic PS (1991) Architectonics of the parietal and temporal association cortex in the strepsirrhine primate galago compared to the anthropoid primate macaca. *J Comp Neurol* 310:475–506. [CrossRef Medline](#)
- Rämä P, Courtney SM (2005) Functional topography of working memory for face or voice identity. *Neuroimage* 24:224–234. [CrossRef Medline](#)
- Romanski LM (2012) Integration of faces and vocalizations in ventral prefrontal cortex: implications for the evolution of audiovisual speech. *Proc Natl Acad Sci U S A* 109[Suppl 1]:10717–10724.
- Romanski LM, Diehl MM (2011) Neurons responsive to face-view in the primate ventrolateral prefrontal cortex. *Neuroscience* 189:223–235. [CrossRef Medline](#)
- Romanski LM, Goldman-Rakic PS (2002) An auditory domain in primate prefrontal cortex. *Nat Neurosci* 5:15–16. [CrossRef Medline](#)
- Romanski LM, Hwang J (2012) Timing of audiovisual inputs to the prefrontal cortex and multisensory integration. *Neuroscience* 214:36–48. [CrossRef Medline](#)
- Romanski LM, Averbach BB, Diltz M (2005) Neural representation of vocalizations in the primate ventrolateral prefrontal cortex. *J Neurophysiol* 93:734–747. [CrossRef Medline](#)
- Rossi AF, Pessoa L, Desimone R, Ungerleider LG (2009) The prefrontal cortex and the executive control of attention. *Exp Brain Res* 192:489–497. [CrossRef Medline](#)
- Rushworth MF, Nixon PD, Eacott MJ, Passingham RE (1997) Ventral prefrontal cortex is not essential for working memory. *J Neurosci* 17:4829–4838. [Medline](#)
- Russ BE, Orr LE, Cohen YE (2008a) Prefrontal neurons predict choices during an auditory same-different task. *Curr Biol* 18:1483–1488. [CrossRef Medline](#)
- Russ BE, Ackelson AL, Baker AE, Cohen YE (2008b) Coding of auditory-stimulus identity in the auditory non-spatial processing stream. *J Neurophysiol* 99:87–95. [CrossRef Medline](#)
- Rygula R, Walker SC, Clarke HF, Robbins TW, Roberts AC (2010) Differential contributions of the primate ventrolateral prefrontal and orbitofrontal cortex to serial reversal learning. *J Neurosci* 30:14552–14559. [CrossRef Medline](#)
- Saults JS, Cowan N (2007) A central capacity limit to the simultaneous storage of visual and auditory arrays in working memory. *J Exp Psychol Gen* 136:663–684. [CrossRef Medline](#)
- Scalaidhe SP, Wilson FA, Goldman-Rakic PS (1999) Face-selective neurons during passive viewing and working memory performance of rhesus monkeys: evidence for intrinsic specialization of neuronal coding. *Cereb Cortex* 9:459–475. [CrossRef Medline](#)
- Strand F, Forssberg H, Klingberg T, Norrelgen F (2008) Phonological working memory with auditory presentation of pseudo-words: an event related fMRI study. *Brain Res* 1212:48–54. [CrossRef Medline](#)
- Stromswold K, Caplan D, Alpert N, Rauch S (1996) Localization of syntactic comprehension by positron emission tomography. *Brain Lang* 52:452–473. [CrossRef Medline](#)
- Sugihara T, Diltz MD, Averbach BB, Romanski LM (2006) Integration of auditory and visual communication information in the primate ventrolateral prefrontal cortex. *J Neurosci* 26:11138–11147. [CrossRef Medline](#)
- Tsao DY, Freiwald WA, Tootell RB, Livingstone MS (2006) A cortical region consisting entirely of face-selective cells. *Science* 311:670–674. [CrossRef Medline](#)
- Tsao DY, Moeller S, Freiwald WA (2008) Comparing face patch systems in macaques and humans. *Proc Natl Acad Sci U S A* 105:19514–19519. [CrossRef Medline](#)
- von Kriegstein K, Giraud AL (2006) Implicit multisensory associations influence voice recognition. *PLoS Biol* 4:e326. [CrossRef Medline](#)
- von Kriegstein K, Kleinschmidt A, Sterzer P, Giraud AL (2005) Interaction of face and voice areas during speaker recognition. *J Cogn Neurosci* 17:367–376. [CrossRef Medline](#)
- Waters G, Caplan D, Alpert N, Stanczak L (2003) Individual differences in rCBF correlates of syntactic processing in sentence comprehension: effects of working memory and speed of processing. *Neuroimage* 19:101–112. [CrossRef Medline](#)
- Xu J, Gannon PJ, Emmorey K, Smith JF, Braun AR (2009) Symbolic gestures and spoken language are processed by a common neural system. *Proc Natl Acad Sci U S A* 106:20664–20669. [CrossRef Medline](#)
- Yu JC, Chang TY, Yang CT (2014) Individual differences in working memory capacity and workload capacity. *Front Psychol* 5:1465. [CrossRef Medline](#)
- Zanto TP, Rubens MT, Thangavel A, Gazzaley A (2011) Causal role of the prefrontal cortex in top-down modulation of visual processing and working memory. *Nat Neurosci* 14:656–661. [CrossRef Medline](#)
- Zatorre RJ, Evans AC, Meyer E (1994) Neural mechanisms underlying melodic perception and memory for pitch. *J Neurosci* 14:1908–1919. [Medline](#)