Journal Section: Behavioral/Systems/Cognitive Neuroscience

Title: Functional connectivity of PAG with core limbic system and laryngeal cortico-motor structures during human phonation

Jessica Galgano[a,b], Spiro Pantazatos[c,d], Kachina Allen[e,f], Ted Yanagihara[g,h], Joy Hirsch[i,j]

[a]Program for Imaging & Cognitive Sciences (PICS), Columbia University, New York, NY, USA [b]Department of Rehabilitation, New York University Langone School of Medicine, New York, NY, USA [c]Department of Psychiatry, Columbia University, New York, NY, USA [d]Molecular Biology and Neuropathology Division, New York Psychiatric Institute, New York, NY, USA [e]Department of Psychology, Princeton University, Princeton, NJ, USA [f]Department of Psychology, Rutgers University, Newark, NJ, USA [g]Department of Radiation Oncology, Weill Cornell Medical College, New York, NY USA [h]New York Presbyterian – Brooklyn Methodist Hospital, Brooklyn, NY, USA [i]Departments of Psychiatry, Neuroscience, and Comparative Medicine, Yale School of Medicine, New Haven, CT, USA [j]Department of Medical Physics and Biomedical Engineering, University College London, UK

Corresponding author to whom reprint requests should be sent:
Jessica Galgano, Ph.D.
NYU Langone School of Medicine
Department of Rehabilitation
Open Lines Speech and Communication
252 West 76th Street #1A
New York, NY 10023
Phone: 212.430.6800; Fax: 212.430.6550
jessica.galgano@nyumc.org
jgphd@openlinesny.com

This manuscript contains:
3 figures, 2 tables, 25 total pages.

Key words: human phonation (movement: motion, motor activity), PAG, fMRI, DTI, connectivity, vocal emotion

All figures should be printed in color.

ABSTRACT

Previous studies in animals and humans suggest the periaqueductal grey region (PAG) is a final integration station between the brain and laryngeal musculature during phonation. To date, a limited number of functional magnetic neuroimaging (fMRI) studies have examined the functional connectivity of the PAG during volitional human phonation. An event-related, stimulus-induced, volitional movement paradigm was used to examine neural activity during sustained vocalization in neurologically healthy adults and was compared to controlled exhalation through the nose. The contrast of vocalization greater than controlled expiration revealed activation of bilateral auditory cortex, dorsal and ventral laryngeal motor areas (dLMA and vLMA) ($p<0.05$, corrected), and suggested activation of the cerbellum, insula, dorsomedial prefrontal cortex (dmPFC), amygdala, and PAG. The functionally defined PAG cluster was used as a seed region for psychophysiological interaction analysis (PPI) to identify regions with greater functional connectivity with PAG during volitional vocalization, while the above functionally defined amygdala cluster was used in an ROI PPI analysis. Whole-brain results revealed increased functional connectivity of the PAG with left vLMA during voicing, relative to controlled expiration, while trend-level evidence was observed for increased PAG/amygdala coupling during voicing ($p=0.07$, uncorrected). Diffusion tensor imaging (DTI) analysis confirmed structural connectivity between PAG and vLMA. The present study sheds further light on neural mechanisms of volitional vocalization that include multiple inputs from both limbic and motor structures to PAG. Future studies should include investigation of how these neural mechanisms are affected in individuals with voice disorders during volitional vocalization.

1. INTRODUCTION

There is significant evidence to support the role of the periaqueductal grey matter (PAG) in eliciting vocalizations produced by lower mammalian species (Jürgens, 2009; Larson & Kistler, 1986; Adamec, 1997; Bandler & Shipley, 1994; Graeff, Silveira, Nogueira, Audi, & Oliveira, 1993; Holstege, 1989; Holstege, Kerstens, Moes, & Vanderhorst, 1997; Kanai, & Wang, 1962; Subramanian, Balnave, & Holstege, 2008; Vanderhorst, Terasawa, Ralston, & Holstege, 2000; Zhang, Bandler, & Davis, 1995; Zhang, Davis, Bandler, & Carrive, 1994). It has been suggested that this region is an integral part of a midline network of neural areas which oversee coordinated activity in lower brainstem neurons that dictate vocal fold movement and respiratory function necessary for phonation (Davis, Zhang, Winkworth, & Bandler, 1996; Holstege, 2014; Subramanian et al., 2008; Zhang et al., 1995). Vocalizations typically convey information about the animal's level of stimulation and emotional state, supporting the idea that PAG activity is governed by the level of participation from other regions in or related to the limbic system (An, Bandler, Ongur, & Price, 1998; Holstege, 1987; Holstege, Meiners, & Tan, 1985; Hopkins & Holstege, 1978; Jürgens, 2002; Jürgens & Zwirner, 1996; Kuipers, Mensinga, Boers, Klop, & Holstege, 2006; Saleem, Kondo, & Price, 2008).

Though a principal problem has been transferring these findings to the human phonatory system, it has been shown that PAG activity is also directly involved in human speech production (Blank, Scott, Murphy, Warburton, & Wise, 2002; Botez & Barbeau, 1971; Esposito, Demeurisse, Alberti, & Fabbro, 1999; Holstege & Subramanian, 2015, Ludlow, 2015, Simonyan & Horwitz, 2011, Schulz et al., 2005; Steriade, Botex, & Petrovici, 1961), which is a result of both the limbic and volitional motor systems (Holstege, 1992; Holstege, 1996; Holstege & Subramanian, 2015). PAG activity mediates movement patterns due, in part, to efferent fibers

projecting from the lateral portion of the PAG to the nucleus retroambigualis (NRA). The NRA then projects neurons to intercostal and abdominal spinal motoneuronal pools (Zhang et al., 1995), as well as laryngeal, pharyngeal, and oral motoneuronal pools (Holstege, 1989; Holstege, Kerstens, Moes, & Vanderhorst, 1997), labeling it as the final common pathway for regulation of vocalization in non-human species (Davis, 1998; Davis et al., 1996; Holstege, 1989; Zhang et al., 1995; Vanderhorst et al., 2000) and humans (Schulz et al., 2005). Neurons in the NRA play a crucial role in emotional vocalizations, such as laughter, as well as speech production (Holstege & Subramanian, 2015).

Using Positron Emission Tomography (PET), Schulz et al., (2005) described the mesencephalic PAG and the visceromotor neurons it serves, as mechanisms that work in a synchronized manner with higher cortical motor systems involved in volitional speech in humans. Findings also revealed a functional connectivity of regions within these systems during vocalization, including, but not limited to the PAG and Anterior Cingulate Cortex (ACC) (Schulz et al., 2005). Lesion studies have demonstrated the ACC's involvement in hierarchical regulation of vocal utterances induced by the PAG (Jürgens & Lu, 1993a, 1993b; Jürgens & Zwirner, 1996); however, the significant role the PAG plays as an integrative emotional center, and how this function influences its responsibility as the final common pathway for human vocalization, remains unclear.

In Jürgens' (1998) model, it was proposed that the PAG acts as a relay center and is also thought to be responsible for coupling the vocal utterance and the emotional state. Considering these findings, we hypothesized that the PAG would exhibit a significant positive change in functional connectivity with the laryngeal motor area during voicing, greater than controlled exhalation. While PAG-amygdala structural connectivity has been previously observed

(Hadjipavlou, Dunckley, Behrens, & Tracey, 2006; Hopkins & Holstege, 1978; Jürgens, 1998: Simonyan & Horwitz; 2011), a limited number of studies have examined PAG-laryngeal motor structural connectivity (Hopkins & Holstege, 1978). Therefore, we employed Diffusion Tensor Imaging (DTI) as well as functional magnetic neuroimaging (fMRI) to further examine and demonstrate the existence of white-matter pathways between these areas.

2. RESULTS

The primary contrast of interest (voicing greater than controlled expiration: M > O) averaged across all subjects revealed significant activation in the bilateral auditory cortices, bilateral ventral and dorsal laryngeal motor areas (LMA) ($p<0.05$, corrected, Table 1, Figure 1), while an exploratory threshold putative activation was observed in left cerebellum, left superior frontal gyrus (dorsomedial Prefrontal Cortex), amygdala, insula, and PAG during vocalization ($p<0.05$, uncorrected), Table 1, Figure 1. In the present study, we used exploratory, uncorrected thresholds but enforced a cluster size requirement ($k > 30$) that was greater than the smoothing kernel (smoothing kernel = 8 mm^3 FWHM and minimum activation volume = 1,080 mm^3) to help reduce false positive rates due to the relative difficulty in detecting activations in paradigms involving overt vocal productions and head and neck movements (Birn et al., 1999; 2004). An additional reason we used a more lenient threshold was that the primary objective of this analysis was to identify midbrain structures to be used as seeds for subsequent functional connectivity analyses.

Given the central role of the PAG during vocalization in mammals, and our observation of greater activity during voicing relative to controlled expiration, we used this area as a seed region to identify other areas of the brain that are functionally connected to the PAG during vocalization. For this, we conducted a PPI analysis, which identifies regions whose contribution

to another area changes significantly as a function of experimental context (Friston et al., 1997). In our case, we searched for regions whose functional coupling to the PAG increased during the voicing condition, relative to the controlled expiration condition. In a whole-brain search, the region with the most significant positive change in coupling to the PAG during the vocalization relative to controlled expiration over all subjects was left vLMA ($p<0.05$, corrected, Table 2, Figure 2). At exploratory, uncorrected thresholds ($p<0.005$, uncorrected, $k>30$), other regions included bilateral visual cortex, dorsal LMA, and insula (Table 2, Figure 2). Given the amygdala activation resulting from our exploratory analysis above, we also restricted our PAG functional connectivity analysis to this functionally defined cluster using an ROI analysis. Using activation averaged across this functionally defined amygdala cluster (amygdala cluster listed in Table 1 and shown in Figure 1) we observed trend level evidence for functional coupling of this cluster with PAG during voicing relative to expiration ($t(19)=1.54$, $p=0.07$). Note this is not double dipping since the effects of task, which was initially used to define the region, were removed (adjusted for) during the PPI parameter estimation within each subject.

DTI analysis revealed substantial white matter connectivity between our functionally-defined midbrain/PAG cluster and left ventral LMA, which also displayed the strongest functional connectivity with this midbrain/PAG cluster (Figure 3). Tractography between PAG and amygdala was not conducted as this has already been demonstrated previously (Hadjipavlou, Dunckley, Behrens, & Tracey, 2006) and revealed structural connectivity between the PAG and amygdala.

3. DISCUSSION

In a whole-brain analysis, the region with the most significant positive change in coupling to the PAG during volitional phonation (vs. exhalation) was the left ventral LMA.

Results from DTI analysis to trace axonal trajectories provide additional evidence of structural connectivity, which complements the finding of functional connectivity between the PAG and left ventral LMA. Considered together, these data support the notion that an emotional motor system and a volitional or somatic motor system are involved in the initiation and control of vocalizations produced from the LMA via the PAG (see Figure 2). It it important to note the PAG has direct projections to and from the nucleus retroambiguus (NRA), which is the only region with direct connections to the motoneuron groups innervating the muscles responsible for producing speech. The prefrontal-PAG-motoneuronal pathway is responsible for generating vocalizations and simultaneously modulating or manipulating vocalizations into words and sentences (i.e., language) by then activating the volitional motor system and its corticobulbar fibers with direct connections to motoneurons that innervate the muscles of the vocal tract needed to produce speech.

Regions active in the voicing condition included bilateral auditory areas, ventral and dorsal LMAs, as well as the left cerebellum, left superior frontal gyrus (dorsomedial Prefrontal Cortex), amygdala, insula, and PAG. Ventral and dorsal LMA clusters were in agreement to those areas found in a similar study (Brown, Ngan, & Liotti, 2007), which revealed that activity in these regions represents movements of the intrinsic muscles of the larynx and controls both spoken and sung vocalizations. Activation of the dorsomedial Prefrontal Cortex (dmPFC) and functional coupling of this area with the PAG during phonation is in agreement with Schultz et al., 2005 and Blank et al., 2002, and suggests that these regions work in an integrated manner during human phonation. Further, functional coupling of the PAG with the vLMA as well as the insula suggests a central mechanism for quick pitch adjustments of the laryngeal system,

necessary for vocal expression (Augustine, 1996; Bamiou, Musiek, & Luxon, 2003; Burnett, Freedland, Larson, & Hain, 1998).

In the current examination, we also found trend-level evidence that the amygdala, well-known for its critical role in emotion and a key limbic region known to process emotional salience cues and determine appropriate behavioral responses via efferent projections to the midbrain, PAG, and laryngeal cortico-motor regions, was functionally connected with the PAG during volitional phonation (vs. controlled exhalation). The coupling of these regions is in accordance with findings from studies on relationships between the PAG and amygdala in felines and rodents (Adamec, 1997; Graeff et al., 1993). In these species, the amygdalo-PAG pathway is critical for the integration of physiological and behavioral expressions of anxiety and panic due to innate or learned threats (Graeff et al., 1993). Furthermore, this pathway is responsible for maintaining the persistent increase in synaptic strength induced by anxiety following traumatic stress (Adamec, 1997). The PAG and the amygdala have been found to reliably co-activate in investigations of affect and emotion (Kober et al., 2008); thus, consideration of the integral role the amygdalo-PAG pathway plays in regulating behavioral and psychological manifestations of different emotions or affective states (Adamec, 1997; Graeff et al., 1993; Bandler & Shipley, 1994; Bandler & Carrive, 1988; Davis, 1998; Davis et al., 1996; Dietrich, Andreatta, Jiang, Joshi, & Stemple, 2012; Wattendorf et al., 2013; Wild, Rodden, Grodd, & Ruch, 2003) is important and supports the idea that the processing of positive or negative emotions or reacting to stressful situations can affect the PAG's ability to also mediate movement patterns during voicing. We speculate this may be due to the PAG's inefficiency in simultaneously regulating heightened states as well as coordinated movements necessary for human phonation. It is also possible that the PAG prioritizes its numerous responsibilities, placing priority on pain reduction,

emotional processing, or autonomic regulation. A greater understanding of the neural circuitry of human voicing and the influence limbic system structures have on those regions governing voicing is necessary and has the potential to provide improved therapies for people with voice disorders whose symptoms worsen secondary to psychogenic and psychosocial factors (e.g., Adductor Type Spasmodic Dysphonia (ADSD)).

Understanding the importance of the roles the PAG plays as both a core limbic system structure and an integral part of human phonation can provide insight into emotionally-triggered vocal behaviors such as discomfort and difficulty speaking when expressing grief, or loss of voice in contexts in which the speaker judges as stressful or fearful. Findings may help us better understand the nature of voice changes secondary to heightened states of affect or emotion, and the brain's ability to regulate different states a person experiences while maintaining vocal control. The integration of cognitive, sensorimotor, and emotional processes is also likely to be disrupted in psychogenic and neurogenic diseases, which affect, in part, areas in the limbic system. Consequently, major feedback pathways to regions that regulate human phonation can be interrupted. There is a considerable body of literature describing the exacerbation of clinical symptoms when experiencing anxiety or heightened emotional states in people with voice disorders such as functional dysphonia and spasmodic dysphonia (Baker, 2003, 2008; Seifert & Kollbruner, 2005; Cannito, 1991). In disorders such as these, symptoms manifest themselves during voiced versus unvoiced speech and emotionally-charged versus unemotionally-charged vocal contexts and states, which directly influence phonation and speech.

The significant impact emotional and affective states have on human voicing has been reported extensively in many bodies of literature, ranging from medicine to the arts (Baker, 2003, 2008; Seifert & Kollbruner, 2005; Cannito, 1991; Davis, 1998; Davis, et al, 1996; Sapir &

Aronson, 1987; Johnstone, Van Reekum, Hird, Kirsner, & Scherer, 2005; Johnstone, Van Reekum, Oakes, & Davidson, 2006). The human voice is, in part, a reflection of a person's level of trait or state anxiety, emotional state and reactivity, as well as their ability to regulate their emotions. Thus, researchers and clinicians commonly refer to the voice as a "barometer of emotion," to emphasize the significant influence psychogenic and psychosocial stress has on the phonatory system (Roy & Bless, 2000; Seifert & Kollbrunner, 2005). Given this, future work could examine the effect of trait and state anxiety on limbic interactions with the PAG and other regions critical to human voicing. Emotional reactivity in conjunction with a person's ability to regulate emotions also needs to be considered in investigations of this nature. Clinical considerations also include examining the effect medicinal, psychological, and alternative treatments, in conjunction with voice therapy, have on disordered voice production.

Limitations

Because the activation analyses were thresholded leniently, we cannot be certain that the midbrain cluster reflects PAG activitation during the voicing condition. However, 1) the cluster's location and vicinity to the PAG and 2) its co-activation with regions expected to be activated during vocalization (i.e. auditory cortex, laryngeal motor cortex), suggests the hypothesis that it reflects PAG activation during voicing. This hypothesis was tested using task-based functional connectivity (PPI) and structural connectivity analyses. The results, phonation-related functional connectivity between PAG and laryngeal motor area (LMA) that survives correction for multiple comparisons and PAG-LMA tractography, provide additional, converging lines of evidence in support of this hypothesis. In other words, if the midbrain cluster defined by the activation results was a false positive

(i.e. noise), then it would not have shown evidence for functional coupling with laryngeal motor regions during the voicing condition.

Summary

The current study demonstrates functional connectivity of the PAG with vLMA (and putative evidence for functional coupling with amygdala) during voicing, relative to controlled expiration. Our results lend support for a model in which limbic center regions interact with midbrain and cortical laryngeal motor regions involved in human voicing/phonation, and identify these putative regions using fMRI. Future studies should examine how these pathways/interactions may be altered in voice disorders, such as ADSD, which are thought to have both a cognitive-emotional as well as neurophysiological component.

4. EXPERIMENTAL PROCEDURE

4.1.1 *Subjects*

Twenty neurologically healthy adults (11 females, 9 males) with a mean age of 54 years (ranging from 22 to 60 years) without previous history of neurological illness or voice disorder were recruited for the functional scans in this prospective study. Data from the lab DTI repository on 20 subjects (14 females), ranging in age from 23 to 52 years of age, who did not participate in the functional scans, were used to perform DTI analysis. All subjects were strongly right-handed on the Edinburgh Handedness Inventory (Oldfield, 1971) and were native English speakers. All subjects provided written informed consent, approved by Columbia University's Institutional Review Board.

4.1.2 Imaging

Functional T2*-weighted images were acquired on a 1.5T GE Twin-Speed scanner with a gradient echo pulse sequence (echo time=39 ms, repetition time=2 s, flip angle=90°) and were

obtained with matched anatomic high resolution T1-weighted scans. During each of 3 runs, 234 whole brain scans were made, each consisting of 27 contiguous slices (4 mm thick), parallel to the AC/PC line (192x192 mm field of view, 64x64 matrix size, 3x3 mm in plane resolution).

DTI images were acquired on a 1.5T GE Twin-Speed scanner using an 8 channel sense head coil with a single-shot sequence of 55 unique diffusion directions at a b-value = 900 with TE=7.8ms and TR=17000ms. A single volume (b-value = 0) was acquired and used as a reference to correct for eddy currents and head motion (Jenkinson and Smith, 2001). Isotropic (2.5 mm^3 voxels) diffusion-weighted data were acquired for all subjects. Array size was 128x128 in a FOV of 320x320mm. 58 slices were acquired and the total scan time was 16 minutes and 32 seconds.

4.1.3 Functional Tasks

A stimulus-induced volitional movement paradigm was employed. In addition, an event-related paradigm was used to minimize motion-related artifact inherent to studies requiring overt productions and head and neck movements (Birn, Bandettini, Cox &, Shaker, 1999; Birn, Cox, & Bandettini, 2004). The experimental tasks consisted of sustained vocalization at a comfortable pitch, sustained exhalation through the nose, and rest as a baseline condition. Specifically, subjects were presented with a 3s visual stimulus that instructed them to either produce an “uh” sound with closed lips (production of “uh” using a neutral lingual position and without concomitant labial movement), sustain exhalation through the nose, or rest. The mouth was closed at all times, thus ensuring breathing through the nose. These activities were followed by a fixation point, with an Interstimulus Interval (ISI) that varied randomly from 2.5-5.0s. The cycle then repeated until 36 stimuli (12 per condition) were presented in pseudorandom order. MATLAB® software (2007a, The MathWorks, New York , NY, USA) was used to control and

present experimental stimuli to the subjects, control scanner initialization sequences, and coordinate timing of stimulus presentation with scanner operations. Subjects were instructed to perform the tasks, which were displayed on the screen visible with custom-designed glasses. A microphone was positioned close to the subject's nose. An observer monitored the subject's task performance (including pitch levels, start/end timing for vocalization) by listening for vocalizations through the speaker attached to scanner equipment. Respiratory cycles were monitored via a nasal canula using a PowerLab/16SP data acquisition system and Chart software (ADInstruments, www.adinstruments.com/). The measured respiratory signals were not included in the analysis (i.e. as covariates to factor out effects of respiration) due to technical issues. However, subjects were explicitly instructed to perform the task at whichever point in the respiratory cycle they were at, thus precluding the possibility that differences in expiration and inspiration between our conditions of interest would confound results.

4.1.4 Preprocessing

Pre-processing was carried out using SPM5 (Wellcome Department of Imaging Neuroscience, London, UK; see http://www.fil.ion.ucl.ac.uk/spm/). The first 4 volumes of each run were discarded. Functional data were then slice-time corrected, spatially realigned to the first volume of the first run, and spatially normalized to the Montreal Neurological Institute (MNI) template brain (re-sampled voxel size: 2 mm^3). These normalized functional images were spatially smoothed with an 8 mm^3 kernel.

4.1.5 First-level Analysis

Functional data were analyzed with SPM5 (Wellcome Department of Imaging Neuroscience, London, UK; see http://www.fil.ion.ucl.ac.uk/spm/). 1st-level regressors were created by convolving the onset of each condition with the canonical HRF (duration of 3

seconds) combined with time and dispersion derivatives, and motion regressors were included as regressors of no interest. The primary contrast of interest (vocalization greater than controlled expiration: M > O) was created from the resulting estimated parameters.

4.1.6 Functional Connectivity

For the psychophysiologic interaction (PPI) analyses (Friston, et al., 1997) the deconvolved time course from a 5 mm radius sphere around the group peak activation voxel from the M > O contrast in the PAG (2 -26 -4) was extracted. Activity throughout the whole-brain was then regressed on a voxel-wise basis against the product of this time course and the vector of the psychological variable of interest (+1 for M conditions, -1 for O condition), with the physiological (PAG time course) and the psychological variables (M and O conditions) serving as regressors of no interest.

4.1.7 Second-level Analysis

Random-effects group analysis was done using SPM5 (Wellcome Department of Imaging Neuroscience, London, UK; see http://www.fil.ion.ucl.ac.uk/spm/). A one sample t-test was conducted on the contrast images (M > O) and PPI maps generated from the 1st level analysis and were masked a standardized grey matter mask. To identify the clusters that survived multiple comparison correction ($p < 0.05$, corrected) we employed 1000 Monte Carlo simulations of whole-brain fMRI data with respective data parameters of the present study using the 3DClustSim program as implement in AFNI (version 2015, May 19th) (Cox, 1996). Note we used an updated version of AFNI 3dClustSim that corrects for the bug identified by Eklund et al., 2016. **See Supplement for further discussion on how our analysis avoids recently identified pitfalls of cluster-extent multiple comparisons correction.** For uncorrected threshold of $p=0.05$, uncorrected, cluster sizes greater than 721 were deemed significant at $p<0.05$ cluster-

extent corrected (marked with an asterisk in Table 1), and for an uncorrected threshold of $p<0.005$, uncorrected, cluster sizes greater than 154 were deemed significant at $p<0.05$ corrected (marked with asterisks in Table 2). Regions that survived this threshold are denoted in Tables 1 and 2 with an asterisk. Given that inherent motion during the task likely decreased our sensitivity in detecting voicing-related activation and that we were interested in functionally identifying a cluster in the midbrain in the vicinity of the PAG, we used a looser exploratory threshold ($p<0.05$ unc, $k>30$) for the activation results, and a more stringent, but still exploratory, threshold ($p=0.005$ unc, $k>30$) for the functional connnectivity map. For ROI analysis we used Marsbar v0.42.

4.1.8 Tractography Analysis

DTI probabilistic tractography analysis was completed using the FMRIB's software library diffusion toolbox (Smith et al., 2004). As described previously(Behrens et al., 2003), a probability of connectivity map was generated for regions of interest. Briefly, in native diffusion space, the principle diffusion direction (PDD) of non-isotropic water movement was modeled as a tensor for each voxel in the brain (Behrens et al., 2003). Complex fiber structure (i.e. crossing or diverging fibers) increases the uncertainty of the PDD estimate. Bayesian statistics were used to generate probability density functions (pdfs) of PDD uncertainty allowing for the detection of non-dominant fiber pathways (Behrens, Berg, Jbabdi, Rushworth, & Woolrich, 2007). From these pdfs, 5000 tract-following samples were taken with a maximum curvature threshold of +/- 80 degrees and the exclusion of pathways that returned onto themselves. As hypothesized, the largest number of streamlines followed a direct path from the seed (i.e., lvLMA) to the waypoint (i.e., PAG) mask in all subjects. These paths were threshold (> 10 streamlines/voxel) and then

binarized and added across all subjects, generating a group representation of individual pathways.

ACKNOWLEDGMENTS

Funding was provided in part by NIH K01MH108721 (SPP). We would like to thank the radiographer, Steve Dashnaw for his expert assistance. We would also like to thank Brianna Rogers and Kristina Sergeev for their editorial support during the preparation of this manuscript. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

An., X., Bandler, R., Ongur, D., & Price, J.L. (1998). Prefrontal cortical projections to longitudinal columns in the midbrain periaqueductal gray in macaque monkeys. *Journal of Compartive Neurology, 401*, 455-479.

Adamec, R. (1997). Transmitter systems involved in neural plasticity underlying increased anxiety and defense—Implications for understanding anxiety following traumatic stress. *Neuroscience & Biobehavioral Reviews, 21*(6), 755-765.

Augustine, J. R. (1996). Circuitry and functional aspects of the insular lobe in primates including humans. *Brain Research Reviews, 22*(3), 229-244.

Baker, J. (2003). Psychogenic voice disorders and traumatic stress experience: a discussion paper with two case reports. *Journal of Voice, 17*(3), 308-318.

Baker, J. (2008). The role of psychogenic and psychosocial factors in the development of functional voice disorders. *International journal of speech-language pathology, 10*(4), 210-230.

Bamiou, D. E., Musiek, F. E., & Luxon, L. M. (2003). The insula (Island of Reil) and its role in auditory processing: literature review. *Brain Research Reviews, 42*(2), 143-154.

Bandler, R., & Carrive, P. (1988). Integrated defence reaction elicited by excitatory amino acid microinjection in the midbrain periaqueductal grey region of the unrestrained cat. *Brain Research, 439*(1-2), 95-106.

Bandler, R., & Shipley, M. T. (1994). Columnar organization in the midbrain periaqueductal gray: modules for emotional expression?. *Trends in Neurosciences, 17*(9), 379-389.

Behrens, T. E. J., Berg, H. J., Jbabdi, S., Rushworth, M. F. S., & Woolrich, M. W. (2007). Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *Neuroimage, 34*(1), 144-155.

Behrens, T. E., Woolrich, M. W., Jenkinson, M., Johansen-Berg, H., Nunes, R. G., Clare, S., ... & Smith, S. M. (2003). Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magnetic Resonance in Medicine, 50*(5), 1077-1088.

Birn, R. M., Bandettini, P. A., Cox, R. W., & Shaker, R. (1999). Event-related fMRI of tasks involving brief motion. *Human Brain Mapping, 7*(2), 106-114.

Birn, R. M., Cox, R. W., & Bandettini, P. A. (2004). Experimental designs and processing strategies for fMRI studies involving overt verbal responses. *Neuroimage, 23*(3), 1046-1058.

Brown, S., Ngan, E., & Liotti, M. (2007). A larynx area in the human motor cortex. *Cerebral Cortex, 18*(4), 837-845.

Blank, S.C., Scott, S.K., Murphy, K., Warburton, E., & Wise, R.J.S. (2002). Speech production: Wernicke, Broca, and beyond. *Brain, 125,* 1829-1838.

Botez, M.I., & Barbeau, A., (1971). Role of subcoritcal structures, and particularly of the thalmus, in the mechanisms of speech and language. A review. *International Journal of Neuroscience*, 1-10.

Burnett, T. A., Freedland, M. B., Larson, C. R., & Hain, T. C. (1998). Voice F0 responses to manipulations in pitch feedback. *The Journal of the Acoustical Society of America, 103*(6), 3153-3161.

Cannito, M. P. (1991). Emotional considerations in spasmodic dysphonia: psychometric quantification. *Journal of Communication Disorders, 24*(5), 313-329.

Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research, 29*(3), 162-173.

Davis, P. (1998). Emotional influences on singing. *Australian Voice, 4*, 13-18.

Davis, P. J., Zhang, S. P., Winkworth, A., & Bandler, R. (1996). Neural control of vocalization: Respiratory and emotional influences. *Journal of Voice, 1*0(1), 23-38.

Dietrich, M., Andreatta, R. D., Jiang, Y., Joshi, A., & Stemple, J. C. (2012). Preliminary findings on the relation between the personality trait of stress reaction and the central neural control of human vocalization. *International Journal of Speech-Language Pathology, 14*(4), 377-389.

Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences, 113*(28), 7900-7905.

Esposito, A., Demeurisse, G., Alberti, B., & Fabbro, F. (1999). Complete mutism after midbrain periaquaeductal gray lesion. *NeuroReport, 10*, 681-685.

Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E., & Dolan, R. J. (1997). Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage, 6*(3), 218-229.

Fu, C. H., Vythelingum, G. N., Brammer, M. J., Williams, S. C., Amaro, E., Andrew, C. M., ... & McGuire, P. K. (2006). An fMRI study of verbal self-monitoring: neural correlates of auditory verbal feedback. *Cerebral Cortex, 16*(7), 969-977.

Graeff, F. G., Silveira, M. C. L., Nogueira, R. L., Audi, E. A., & Oliveira, R. M. W. (1993). Role of the amygdala and periaqueductal gray in anxiety and panic. *Behavioural Brain Research, 58*(1), 123-131.

Hadjipavlou, G., Dunckley, P., Behrens, T. E., & Tracey, I. (2006). Determining anatomical connectivities between cortical and brainstem pain processing regions in humans: A diffusion tensor imaging study in healthy controls. *Pain, 123*(1), 169-178.

Holstege, G. (1987a). Anatomical evidence of an ipsilateral rubrospinal pathway and for direct rubrospinal projections to motorneurons in the cat. *Neuroscience Letters, 74*, 269-274.

Holstege, G. (1989). Anatomical study of the final common pathway for vocalization in the cat. *Journal of Comparative Neurology, 284*(2), 242-252.

Holstege, G. (1992). The emotional motor system. *European Journal of Morphology, 30,* 67-79.

Holstege, G. (2014). The periaqueductal controls brainstem emotional motor systems including respiration. *Progress in Brain Research, 209*, 379-405.

Holstege, G., Bandler, R., & Saper, C. B. (1996). The emotional motor system. *Progress in Brain Research, 107*, 3-6.

Holstege, G., Kerstens, L., Moes, M. C., & Vanderhorst, V. G. J. M. (1997). Evidence for a periaqueductal gray–nucleus retroambiguus–spinal cord pathway in the rat. *Neuroscience, 80*(2), 587-598.

Holstege, G., Meiners, L., & Tan, K., (1985). Projections of the bed nucleus of the stria terminalis to the mesencephalon, pons, and medulla oblongata in the cat. *Experimental Brain Research,* 58, 379-391.

Holstege, G., & Subramanian, H. H. (2015). Two different motor systems are needed to generate human speech. *Journal of Comparative Neurology, 524*, 1558–1577.

Hopkins, D.A., & Holstege, G. (1978). Amygdaloid projections to the mesencephalon, pons, and medulla oblongata in the cat. *Experimental Brain Research, 32,* 529-547.

Jenkinson, M., & Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis, 5*(2), 143-156.

Johnstone, T., van Reekum, C. M., Hird, K., Kirsner, K., & Scherer, K. R. (2005). Affective speech elicited with a computer game. *Emotion, 5*(4), 513.

Johnstone, T., Van Reekum, C. M., Oakes, T. R., & Davidson, R. J. (2006). The voice of emotion: an FMRI study of neural responses to angry and happy vocal expressions. *Social Cognitive and Affective Neuroscience, 1*, 242-249.

Jürgens, U. (1998). Neuronal control of mammalian vocalization, with special reference to the squirrel monkey. *Naturwissenschaften, 85*(8), 376-388.

Jürgens, U. (2002). Neural pathways underlying vocal control. *Neuroscience & Biobehavioral Reviews, 26*(2), 235-258.

Jürgens, U. (2009). The neural control of vocalization in mammals: a review. *Journal of Voice, 23*(1), 1-10.

Jürgens, U., & Lu, C. L. (1993a). The effects of periaqueductally injected transmitter antagonists on forebrain-elicited vocalization in the squirrel monkey. *European Journal of Neuroscience, 5*(6), 735-741.

Jürgens, U., & Lu, C. L. (1993b). Interactions between glutamate, GABA, acetylcholine and histamine in the periaqueductal gray's control of vocalization in the squirrel monkey. *Neuroscience letters, 152*(1), 5-8.

Jürgens, U., & Zwirner, P. (1996). The role of the periaqueductal grey in limbic and neocortical vocal fold control. *Neuroreport, 7*(18), 2921-2924.

Kanai, T., Wang, S.C. (1962). Localization of the central vocalization mechanism in the brainstem of the cat. *Experimental Neurology, 6,* 426-434. 113, 7900–7905.

Kober, H., Barrett, L. F., Joseph, J., Bliss-Moreau, E., Lindquist, K., & Wager, T. D. (2008). Functional grouping and cortical–subcortical interactions in emotion: a meta-analysis of neuroimaging studies. *Neuroimage, 42*(2), 998-1031.

Kuipers, R., Mensinga, G.M., Boers, J., Klop, E.M., & Holstege, G. (2006). Infralimbic cortext projects to all parts of the pontine and medullary lateral tegmental field in cat. *European Journal of Neuroscience, 23,* 3014-3024.

Kreifelts, B., Ethofer, T., Grodd, W., Erb, M., & Wildgruber, D. (2007). Audiovisual integration of emotional signals in voice and face: an event-related fMRI study. *Neuroimage, 37*(4), 1445-1456.

Larson, C. R., & Kistler, M. K. (1986). The relationship of periaqueductal gray neurons to vocalization and laryngeal EMG in the behaving monkey. *Experimental Brain Research, 63*(3), 596-606.

Lattner, S., Meyer, M. E., & Friederici, A. D. (2005). Voice perception: Sex, pitch, and the right hemisphere. *Human Brain Mapping, 24*(1), 11-20.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia, 9*(1), 97-113.

Penfield, W. & Roberts, L. (1959). *Speech and brain-mechanisms*. Princeton, NJ: Princeton University Press.

Penfield, W. & Rasmussen, T. (1950). *The cerebral cortex of man: A clinical study of localization of function*. New York: Macmillan.

Rämä, P., & Courtney, S. M. (2005). Functional topography of working memory for face or voice identity. *Neuroimage, 24*(1), 224-234.

Roy, N. & Bless, D. M. (2000). Personality traits and psychological factors in voice pathology: A foundation for future research. *Journal of Speech, Language, and Hearing Research, 43*(3), 737.

Saleem, K.S., Kondo, H., & Price, J.L. (2008). Complementary circuits connecting the orbital and medial prefrontal networks with the temporal, insular, and opercular cortex in the macaque monkey. *Journal of Comparative Neurology, 506,* 659-692.

Sapir, S., & Aronson, A. E. (1987). Coexisting psychogenic and neurogenic dysphonia: a source of diagnostic confusion. *British Journal of Disorders of Communication, 22*(1), 73-80.

Schulz, G. M., Varga, M., Jeffires, K., Ludlow, C. L., & Braun, A. R. (2005). Functional neuroanatomy of human vocalization: an H215O PET study. *Cerebral cortex, 15*(12), 1835-1847.

Seifert, E., & Kollbrunner, J. (2005). Stress and distress in non-organic voice disorders. *Swiss Medical Weekly, 135*(27/28), 387.

Simonyan, K., & Horwitz, B. (2011). Laryngeal motor cortex and control of speech in humans. *Neuroscientist, 17(2),* 197-208.

Simonyan, K., Ostuni, J., Ludlow, C. L., & Horwitz, B. (2009). Functional but not structural networks of the human laryngeal motor cortex show left hemispheric lateralization during syllable but not breathing production. *The Journal of Neuroscience, 29*(47), 14912-14923.

Steriade, M., Botex, M.I., & Petrovici, I., (1961). On certain dissociations of consciousness levels with the sypmtoms of akinetic mutism. *European Neurology, 141,* 49-58.

Subramanian, HH., Arun, M., Silburn, PA., & Holstege, G. (2008). The midbrain periaqueductal gray control of expiration. *Journal of Neuroscience, 28,* 12274-12283

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., ... & Niazy, R. K. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage, 23,* S208-S219.

Vanderhorst, V. G., Terasawa, E., Ralston, H. J., & Holstege, G. (2000). Monosynaptic projections from the lateral periaqueductal gray to the nucleus retroambiguus in the rhesus monkey: Implications for vocalization and reproductive behavior. *Journal of Comparative Neurology, 424*(2), 251-268.

Wattendorf, E., Westermann, B., Fiedler, K., Kaza, E., Lotze, M., & Celio, M. R. (2013). Exploration of the neural correlates of ticklish laughter by functional magnetic resonance imaging. *Cerebral Cortex, 23*(6), 1280-1289.

Wild, B., Rodden, F. A., Grodd, W., & Ruch, W. (2003). Neural correlates of laughter and humour. *Brain, 126*(10), 2121-2138.

Zhang, S. P., Bandler, R., & Davis, P. J. (1995). Brain stem integration of vocalization: role of the nucleus retroambigualis. *Journal of Neurophysiology, 74*(6), 2500-2512.

Zhang, S. P., Davis, P. J., Bandler, R., & Carrive, P. (1994). Brain stem integration of vocalization: Role of the midbrain periaqueductal gray. *Journal of Neurophysiology, 72*(3), 1337-1356.

Journal Section: Behavioral/Systems/Cognitive Neuroscience

Supplement for

Functional connectivity of PAG with core limbic system and laryngeal cortico-motor structures during human phonation

Jessica Galgano[a,b], Spiro Pantazatos[c,d], Kachina Allen[e,f], Ted Yanagihara[g,h], Joy Hirsch[i,j]

[a]Program for Imaging & Cognitive Sciences (PICS), Columbia University, New York, NY, USA [b]Department of Rehabilitation, New York University Langone School of Medicine, New York, NY, USA [c]Department of Psychiatry, Columbia University, New York, NY, USA [d]Molecular Biology and Neuropathology Division, New York Psychiatric Institute, New York, NY, USA [e]Department of Psychology, Princeton University, Princeton, NJ, USA [f]Department of Psychology, Rutgers University, Newark, NJ, USA [g]Department of Radiation Oncology, Weill Cornell Medical College, New York, NY USA [h]New York Presbyterian – Brooklyn Methodist Hospital, Brooklyn, NY, USA [i]Departments of Psychiatry, Neuroscience, and Comparative Medicine, Yale School of Medicine, New Haven, CT, USA [j]Department of Medical Physics and Biomedical Engineering, University College London, UK

Corresponding author to whom reprint requests should be sent:
Jessica Galgano, Ph.D.
NYU Langone School of Medicine
Department of Rehabilitation
Open Lines Speech and Communication
252 West 76th Street #1A
New York, NY 10023
Phone: 212.430.6800; Fax: 212.430.6550
jessica.galgano@nyumc.org
jgphd@openlinesny.com

This file contains:
2 total pages.

Key words: human phonation (movement: motion, motor activity), PAG, fMRI, DTI, connectivity, vocal emotion

Discussion

We applied cluster-extent thresholding using AFNI's 3dClustSim, a standard procedure to correct for multiple comparisons across the brain. The key advantage behind this correction is its greater sensitivity to weaker, spatially distributed signals relative to voxel-wise FDR/FWE correction, which is too strict for fMRI data given the thousands of voxels that are tested. Based upon the combination of individual voxel probability thresholding (or cluster-determining thresholding, CDT) and minimum cluster size, the probability that a cluster of a given size is a false positive is estimated using Monte Carl simulation of the underlying noise distribution.

In 2016 Eklund et al. found most cluster-based parametric methods for multiple comparisons correction that use an uncorrected CDT, implemented in standard processing packages (i.e. SPM, FSL and AFNI), had inflated familywise Type I error rates, and that this inflation was worse for lower cluster-defining thresholds. Although we did not use parametric cluster-extent thresholding, the Eklund paper also raised the following two concerns about AFNI's 3dClustSim tool: 1) a 15 year old bug in 3dClustSim led to systematically lower p-values and 2) group smoothness as estimated by AFNI may be too low. Note that our study applied the most recent version of 3dClustSim (compiled May 19th, 2015) which does not include the bug mentioned in point 1. In addition, we used SPM, not AFNI, to estimate smoothness. Whereas AFNI estimates overall smoothness by averaging together estimates from the first-level analyses, SPM (and FSL) estimate smoothness using group residuals from the general linear model (Kiebel et al., 1999).

Eklund et al. also raised an additional concern regarding the assumption of constant smoothness across the brain. While it is true that our cluster correction assumed a constant spatial smoothness over the brain, it is not clear whether this contributes to an increased likelihood of false positives. In fact, the authors of the Eklund paper calculated FWE rates for stationary (constant smoothness) vs. non-stationary assumptions and did not find significant differences in false positive rates. However, they state "This inconclusive performance can be attributed to additional assumptions and approximations introduced by the nonstationary cluster size test that can degrade its performance…we still cannot rule out heterogeneous smoothness as contributor to the standard cluster size methods' invalid performance." (Eklund et al., 2016).

In addition, it is worth noting that since the 2016 Eklund paper several articles have argued that conventional cluster-wise extent thresholding methods do in fact have acceptable false-positive rates (Slotnick, 2017a)***. Among criticisms of the Eklund et. al. is the fact that the 'null-data' used in the study reflects resting-state default mode network activity*** (Slotnick, 2017b) ***and thus may not***

accurately reflect the probability distribution of the test-statistic when the null hypothesis of no activation is true during an event-related task fMRI.

REFERENCES

Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences*, *113*(28), 7900-7905.

Kiebel, S.J., Poline, J.B., Friston, K.J., Holmes, A.P., Worsley, K.J., (1999). Robust smoothness estimation in statistical parametric maps using standardized residuals from the general linear model. *NeuroImage 10*, 756–766.

Slotnick, S.D., (2017a). Cluster success: fMRI inferences for spatial extent have acceptable false-positive rates. *Cognitive Neuroscience. 8*, 150–155.

Slotnick, S.D., (2017b). Resting-state fMRI data reflects default network activity rather than null data: A defense of commonly employed methods to correct for multiple comparisons. *Cognitive Neuroscience. 8*, 141–143.

FIGURE CAPTIONS

Figure 1. Group statistical activation map of the contrast (M > O). Whole-brain ($p < 0.05$ unc, k=30): bilateral ventral laryngeal motor areas (LMA), dorsomedial prefrontal cortex (top panel), bilateral temporal gyrus and midbrain/PAG (middle panel) and right amygdala cluster.

Figure 2. Group functional connectivity statistical map: regions showing increased functional coupling with PAG during voicing compared to controlled expiration. Whole-brain ($p < 0.005$ unc, k=30): **(A)** bilateral visual cortex (BA18) and superior temporal gyrus, **(B)** vLMA and insula and **(C)** right dLMA. vLMA cluster is significant at $p<0.05$, corrected.

Figure 3. White matter paths between left ventral laryngeal motor area (lvLMA) and midbrain in the vicinity of periaqueductal grey matter (PAG) for 30 subjects that did not participate in the task. The lvLMA cluster mask was defined from the group psychophysiologic interaction (PPI) t-map (areas with greater correlation with PAG during voicing vs. controlled expiration) and the PAG cluster mask was defined by the group contrast map (M > O) (see methods and results section). Intensity values at each voxel for the Diffusion Tensor Imaging (DTI) analysis group image corresponded to the number of subjects with a streamline passing through that voxel. Note that the group image does not correspond to a map of probabilistic connectivity from the seed to the waypoints mask as presented for individual subjects, but instead represents the importance of each voxel to this pathway with respect to all subjects. Paths were thresholded to show voxels that included waypoints from 15 or more subjects. Numbers in lower left indicate the Y value of the coronal section in MNI space.

Figure 1.

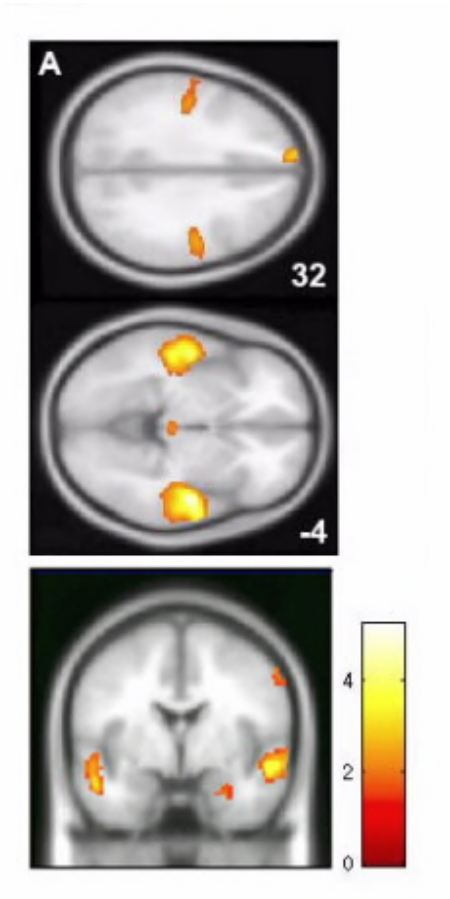

Figure 2.

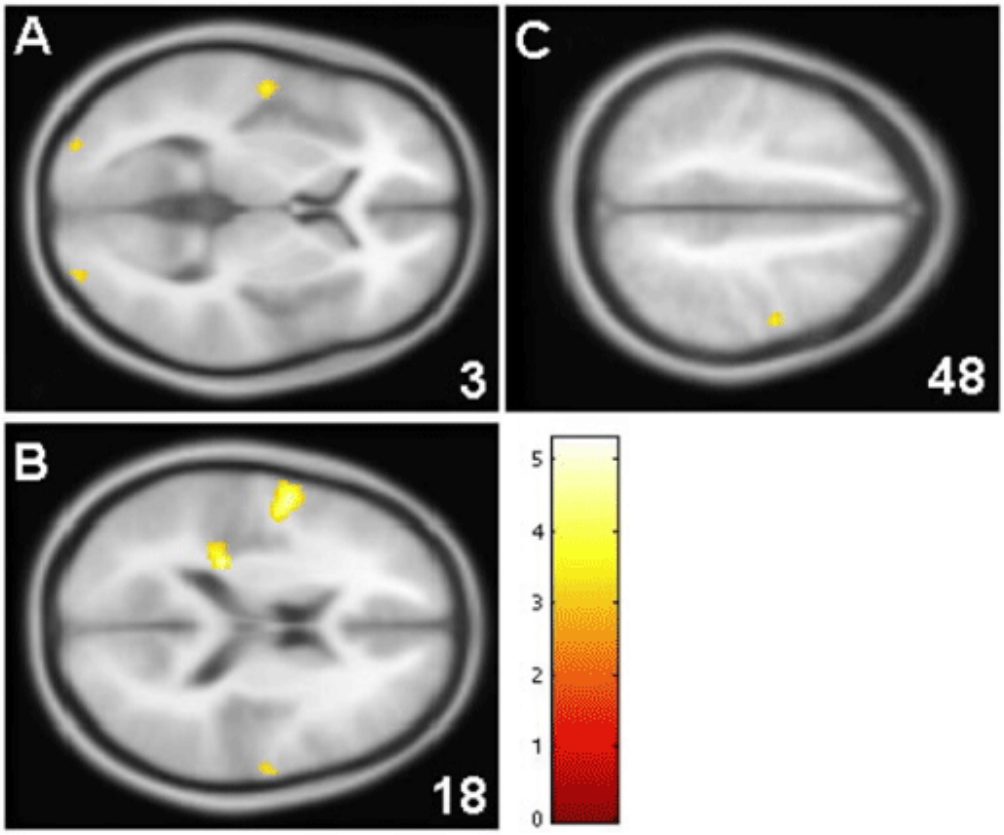

Figure 3.