

Patients with Lesions to Left Prefrontal Cortex (BA 9 and BA 10) Have Less Entrenched Beliefs and Are More Skeptical Reasoners

Vinod Goel^{1,2}, Miriam Marling¹, Vanessa Raymont³,
Frank Krueger^{4,5}, and Jordan Grafman^{6,7}

Abstract

■ The effect of prior beliefs on reasoning and decision-making is a robust, poorly understood phenomenon, exhibiting considerable individual variation. Neuroimaging studies widely show the involvement of the left pFC in reasoning involving beliefs. However, little patient data exist to speak to the necessity and role of the left pFC in belief-based inference. To address this shortcoming, we tested 102 patients with unilateral focal penetrating traumatic brain injuries and 49 matched controls. Participants provided plausibility ratings (plausible/implausible) to simple inductive arguments and (separately) strength of believability ratings of the conclusion to those same arguments. A voxel-based lesion symptom mapping analysis identified 10 patients, all with lesions to the left pFC (BA 9 and BA 10) as rating significantly fewer arguments with highly believable conclusions as “plausible,” compared with all other patients. Subsequent analyses, incorporating the right hemisphere homologue of these patients ($n = 12$) and normal controls ($n = 24$), revealed

patients with lesions to left pFC found fewer arguments plausible in the high believable than either of these groups, and there was no difference in the behavioral scores of the right pFC patients and normal controls. Further analysis, utilizing the belief ratings as the dependent measure, revealed a Group \times Belief Rating interaction, with left pFC patients having less intense beliefs about the conclusions of moderately believable and highly believable arguments. We interpreted these results to indicate that lesions to left pFC (BA 9, BA 10) increase incredulity and make these patients more skeptical reasoners. The former can partially, but not fully, explain the latter. The other relevant factor may be that unilateral left pFC lesions disrupt hemispheric equilibrium and allow for an increased inhibitory role of the right pFC. We speculate that individual differences in belief bias in reasoning in the normal population may be a function of individual differences in the left and right pFC interactional dynamics. ■

INTRODUCTION

We read in the local newspaper that “pomegranate juice leads to weight loss.” Given that we are trying to lose weight, do we start drinking pomegranate juice? This will depend upon our evaluation of the given evidence and/or argument. Reading further, we find the following argument: “Pomegranate juice is high in antioxidants; antioxidants increase metabolic activity in lab mice; increasing metabolic activity leads to weight loss; therefore drinking pomegranate juice will lead to weight loss.” Like most real-world arguments, this is an inductive argument. Logical form provides no guidelines as to the acceptability of the conclusion. Evaluation of the argument involves evaluating the premises and conclusion (and the relationship between them) in light of our existing belief network.

Are the premises and conclusion believable (i.e., Are they consistent with our existing belief network? Does incorporating the conclusion clarify and expand our belief network or does it result in inconsistencies? If there are inconsistencies, can additional steps be taken to resolve them?)? What is the relationship between the premises and conclusion? Is it a conceptual or causal relationship? Is the truth of the premises sufficient to increase our acceptance of the conclusion (in the context of our existing belief network)? Can we generate counterexamples to the conclusion? For example, Donald drinks pomegranate juice but is overweight.

We know that there are individual differences in how people process new information and undertake belief revision (Stanovich, 1999; Stanovich & West, 1997; Weber & Crocker, 1983). Some people will accept the given information at face value and add it to their set of beliefs (after all, it was in the newspaper!). Others may display some skepticism and give it some limited credence but be open to the possibility of revision. Still, others will be more skeptical and want to know the source and data

¹York University, Toronto, Canada, ²Capital Normal University, Beijing, China, ³Oxford University, ⁴George Mason University, ⁵University of Mannheim, ⁶Shirley Ryan AbilityLab, Chicago, IL, ⁷Northwestern University Medical School

underlying the claim before determining how much credence to assign to it. What was the experimental design of the study? What did the authors of the study actually conclude? It is also possible that one is unable to accept the new information at all because it is inconsistent with some other piece of information in the belief network and the inconsistency cannot be reconciled (e.g., we may know that antioxidants have no effect on metabolic rate). The level of credibility one assigns to the information will affect how deeply it gets entrenched into our belief network and also the inferences we are willing to draw from it.

Very little is known about how belief networks and inductive inferences are supported in the brain (Krawczyk, 2018; Shallice & Cooper, 2011), though there are two sets of relevant studies in the neuroimaging literature.¹ The first set examines the neural correlates of strongly held beliefs, such as political and religious beliefs, versus normal beliefs, like “all apples are nutritious.” The basic findings are that, whereas strongly held political and religious beliefs have strong emotional components that activate frontal orbital and limbic system regions, normal beliefs involve the left pFC (Kaplan, Gimbel, & Harris, 2016; Moutsiana, Charpentier, Garrett, Cohen, & Sharot, 2015; Gozzi, Zamboni, Krueger, & Grafman, 2010; Harris et al., 2009; Kapogiannis et al., 2009). The second set of studies explore the role of beliefs in logical reasoning. The basic findings here are as follows: (1) Inferences based on world knowledge and semantic/conceptual connections, along with simple logical and causal connections, preferentially activate the left pFC (among other areas; Liang, Goel, Jia, & Li, 2014; Green, Fugelsang, Kraemer, Shamosh, & Dunbar, 2006; Goel & Dolan, 2003, 2004; Luo et al., 2003; Goel, Buchel, Frith, & Dolan, 2000; Wharton et al., 2000; Goel, Gold, Kapur, & Houle, 1997). (2) Where there is a conflict between the believability of the premises and/or conclusion and the logic of the inference, the right pFC is engaged in the detection of the conflict and/or inhibition of the belief-based response (Liang et al., 2014; Stollstorff, Vartanian, & Goel, 2012; Goel & Dolan, 2003; Goel et al., 2000).

Apart from the split-brain patient studies (Wolford, Miller, & Gazzaniga, 2000; Gazzaniga, 1989, 1998; Gazzaniga & Smylie, 1984), there is limited patient data supporting these imaging results (Reverberi, Lavaroni, Gigli, Skrap, & Shallice, 2005). To address this shortcoming, we undertook a study with 102 neurological patients with unilateral focal lesions and 49 normal controls matched for age, education, and socioeconomic backgrounds, utilizing simple inductive inferences that varied in both argument plausibility and conclusion believability, as in Arguments A–C (see below). Based on the above-cited imaging studies and split-brain patient data of belief and inference, we expected patients with lesions to left pFC to perform differently on both the belief and inference components of the task than patients with right pFC lesions and normal

controls.² Furthermore, given the role of the left pFC in belief-based inferences, we expected performance differences to be greater in trials with highly believable conclusions than in trials with unbelievable conclusions.

METHODS

Patient Selection

Participants were selected from the Vietnam Head Injury Study (Phase 3; Raymont, Salazar, Krueger, & Grafman, 2011). All patients served in the Vietnam War, sustaining penetrating traumatic brain injuries between 1967 and 1970. They were matched with 49 normal controls who were also Vietnam combat veterans but did not sustain any head injuries during the war (or subsequently). The experimental protocol was approved by the institutional review board at the National Naval Medical Center in Bethesda, MD. All participants understood the procedures and gave informed consent.

One hundred sixty-five neurological patients with focal lesions resulting from penetrating traumatic brain injury and 49 normal matched controls completed an inductive inference task.³ To address the issue of hemispheric asymmetry, we excluded patients with bilateral damage. To avoid outliers, patients were also excluded if their total brain damage was less than 1 cc or greater than 100 cc. This left a total of 102 patients, 43 patients with unilateral left hemisphere lesions and 59 patients with unilateral right hemisphere lesions. The patients and normal controls were matched in terms of age, education, preinjury intelligence, and other neuropsychological measures (see below). A proportion test showed no significant differences in handedness between the groups, $\chi^2 = 0.19$, $df = 2$, $p = .91$ (percentage of right-handed left-damaged patients = 79.7%, right-damaged patients = 79.1%, controls = 83.3%). Overlays of lesion extent of the patient group are displayed in Figure 1.

Neuropsychological Assessment

All participants completed the Armed Forces Qualification Test (AFQT-7A; Eitelberg, Laurence, Waters, & Perelman, 1984) upon induction into the Armed Forces and were readministered it as part of their neuropsychological assessment for this study.⁴ The Wechsler Adult Intelligence Scale (WAIS-III; Wechsler, 1997a), Delis–Kaplan Executive Function System (D-KEFS) Sorting Test (Heled, Hoofien, Margalit, Natovich, & Agranov, 2012), and Wechsler Memory Scale (WMS-III; Wechsler, 1997b) were also administered to assess participants’ cognitive functioning level. Psychological and emotional functioning was assessed by Beck Depression Inventory (BDI; Beck, 1987). The age, education, cognitive profiles, preinjury and postinjury AFQT scores, emotional and psychological

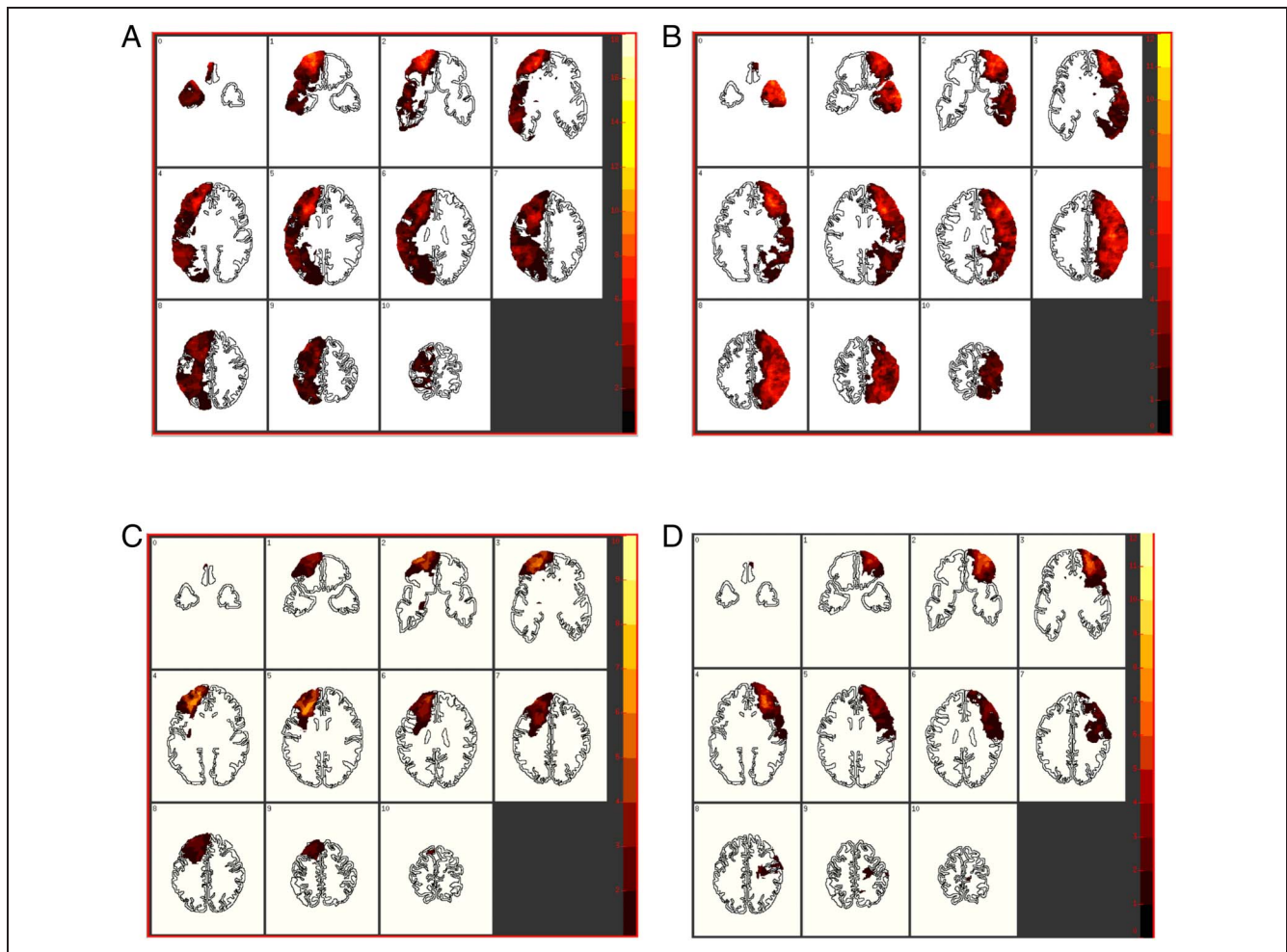


Figure 1. Lesion overlay maps displaying the lesion locations (R = R) of (A) the 43 patients in the left hemisphere lesion group, (B) the 59 patients in the right hemisphere lesion group, (C) the 10 patients in the left pFC patient group selected in VLSM analysis, and (D) the 12 patients selected for the right homologue patient group. These lesions are overlaid onto a Damasio and Damasio (1989) template. Color indicates the lesion overlap.

measurements (BDI), and the size of lesion (see below) are noted in Table 1.

The participants' WMS memory and WAIS IQ averages are all within the normal range. ANOVA was used to examine differences in demographic and neuropsychological scores between the right hemisphere, left hemisphere, and normal control groups. The three groups were well matched for age, education, WAIS-III verbal IQ (and Similarities subtest), D-KEFS Sorting Test, and BDI scores ($p > .05$). There was no significant difference for total brain volume loss between the right and left frontal groups. The right hemisphere lesion patients had a significantly lower full IQ (mean difference = -7.98 , $p = .039$) and preinjury IQ (AFQT-7A; mean difference = -17.68 , $p = .045$) than normal controls. The difference between the normal controls and patients with right hemisphere lesions in the postinjury AFQT scores was reduced to a trend level (mean difference = 13.44 , $p = .058$). The left hemisphere lesion patients had significantly lower WMS general memory (mean difference =

-9.39 , $p = .047$) and working memory than normal controls (mean difference = -8.35 , $p = .034$).

Lesion Location and Extent

Lesion location and extent, as determined from patient CT scans, are specified in summary overlay images in Figure 1A, B. The CT scans were acquired on a GE Light Speed Plus CT scanner in helical mode (150 slices per participant, the field of view covering head only). Images were reconstructed with an in-plane voxel size of 0.4×0.4 mm, overlapping slice thickness of 2.5 mm, and a 1-mm slice interval. Skull and scalp components were removed using the BET algorithm in MEDx (Medical Numerics, Inc.). Patient CT volumes were imported into ABLe (Medical Numerics, Inc.) software (Solomon, Rayment, Braun, Butman, & Grafman, 2007; Makale et al., 2002) and displayed as a series of slices in a light box format. A trained neuropsychiatrist manually traced the lesions on all relevant slices. The tracings were

then reviewed by J. G., who was blind to the results of the neuropsychological testing. Lesion location and volume were determined from the CT images using the Analysis of Brain Lesion software (Solomon et al., 2007; Makale et al., 2002) contained in MEDx v3.44 (Medical Numerics, Inc.) with enhancements to support the Automated Anatomical Labeling atlas (Tzourio-Mazoyer et al., 2002). Total lesion volume (in cubic centimeters) was calculated by voxel count and is summarized in Table 1. The two patient groups did not differ in terms of total volume loss, $t(101) = 0.21, ns$.

The patient volume was then normalized to a reference template volume by a 12-parameter affine linear transformation (allowing for translation, rotation, scaling, and shearing). The lesion voxels were included in the registration process. The ABLe reference volume is an MRI of a 27-year-old normal man transformed to Talairach space with a 12-parameter affine linear transformation. The volume is resliced at 17° relative to the inferior orbitomeatal line, and 11 transverse slices that best match the Damasio and Damasio (1989) templates have been selected by a neuroradiologist and interactively labeled with Brodmann's areas (BA) by reference to the Damasio and Damasio templates. Although the locations of BAs in these templates are approximate, they

are widely accepted in the neuropsychology and neurology communities.

Each registered patient volume was then resliced at a 17° cranial angle, and the 11 sections that matched the ABLe reference volume (hence the Damasio and Damasio templates) were automatically extracted. Because the BAs are premarked on the 11 slices of the ABLe reference volume (see above) and the patient brain volume has been registered and resliced to conform to this template, the intersection of lesion with BAs was calculated by a simple voxel-by-voxel comparison. The two patient groups did not differ in terms of total volume loss, $t(101) = 0.21, ns$.

One concern with patients with penetrating traumatic brain injuries is in regard to contrecoup brain injury associated with significant force and trauma. However, there is no strong evidence that penetrating traumatic brain injury due to shell fragments, which is the cause of damage among most of our patient population, routinely results in contrecoup effects (Raymont et al., 2011; Grafman et al., 1988). Certainly, our CT scan analysis only rarely gave such hints. Although we cannot eliminate the possibility of any microscopic damage (only an autopsy could do that), based on CT scan data, we believe that most of the energy imparted from the penetrating brain wounds occurred at the point of entry and along the missile path.

Table 1. Demographic and Neuropsychological Mean (SD) Scores for the 43 Patients with Left Hemisphere Lesions, 59 Patients with Right Hemisphere Lesions, and Normal Controls

<i>Measures</i>	<i>Normal Controls (n = 24)</i>	<i>Patient Group</i>	
		<i>Left Hemisphere Lesion Patients (n = 43)</i>	<i>Right Hemisphere Lesion Patients (n = 59)</i>
Age (years)	58.21 (2.34)	58.42 (2.81)	58.34 (3.56)
Education	14.95 (2.36)	14.89 (2.48)	14.68 (2.57)
WAIS-III IQ			
Full IQ	112.13 (9.72)*	106.12 (14.84)	104.15 (12.58)*
Verbal IQ	111.38 (11.39)	106.28 (14.95)	105.39 (12.89)
Similarities subtest	11.46 (2.30)	12.42 (3.40)	11.39 (2.57)
Preinjury AFQT-7A	76.50 (19.77)*	67.38 (21.55)	58.82 (26.33)*
Postinjury AFQT-A	72.75 (19.03)	58.40 (22.24)	59.31 (24.15)
WMS			
General memory	107.13 (12.07)*	97.74 (17.47)*	103.07 (14.11)
Working memory	108.04 (13.72)*	99.69 (12.27)*	103.13 (12.6)
D-KEFS Card Sorting	11.54 (2.86)	10.51 (3.07)	10.70 (2.88)
BDI	11.00 (8.68)	7.79 (7.63)	9.80 (9.85)
Total brain volume loss (cc)		26.51 (19.62)	25.78 (19.04)

WAIS-III = Wechsler Adult Intelligence Scale III, scaled scores reported; AFQT-7A = Armed Forces Qualification Test percentile rank; WMS = Wechsler Memory Scale, scaled scores reported; D-KEFS Card Sorting = Delis-Kaplan Executive Function System Card Sorting Task (formerly the California Card Sorting Task), composite scaled scores reported; cc = cubic centimeters.

* $p = .075$ (Hochberg post hoc tests).

Task, Experimental Design, and Stimuli

The task comprised 32 inductive reasoning arguments of the type used in previous studies (Sloman & Lagnado, 2005; Goel & Dolan, 2004; Goel et al., 1997), each containing two premises followed by a conclusion. Arguments utilized were as follows (fully reproduced in Appendix A):

Argument A: Rexdale is a German shepherd; Rexdale lives in Dusseldorf; all German shepherds live in Dusseldorf.

Argument B: Lipstick is moist and glossy; fish scales are moist and glossy; lipstick is made from fish scales.

Argument C: Snakes are reptiles; snakes are cold-blooded; all reptiles are cold-blooded.

Such arguments can vary along a number of dimensions. Two such dimensions are the plausibility of the arguments and the strength of our belief in the conclusions (given our existing belief network). In the above three examples, we all have background knowledge and beliefs about German shepherds, reptiles, and the manufacture of lipstick, but the strength of those beliefs will vary. Most of us have strong beliefs that all reptiles are cold-blooded (though this is actually false) and find the conclusion that “all German shepherds live in Düsseldorf” inconsistent with our belief that German shepherds can be found in many cities around the world. Most of us probably do not have many strong beliefs about the manufacture of lipstick and would be less certain about the use of fish scales in the manufacturing process. In these examples, we would expect most people to find the conclusion of A to be highly unbelievable, the conclusion of B to be uncertain, and the conclusion of C to be highly believable.

Given that inductive inference cannot rely upon logical form, we expected conclusion believability to modulate argument plausibility ratings. However, the two are not identical. In the case of believability judgments, one is evaluating the conclusion/proposition in the context of one’s existing belief network. In the case of argument plausibility, one is evaluating that same conclusion/proposition in light of the premises and one’s existing belief network. In the latter case, the presentation of the premises will add this information to the belief network (if not present), or at least highlight it (if already present), or perhaps correct an existing belief. So these are two distinct judgments, though one might expect a correlation between them.

The participants carried out two tasks with these materials. In one task, argument premises plus conclusion were presented on a computer screen and remained on screen until the participant responded or the time limit expired. Participants were given up to 30 sec and asked to evaluate each argument as either “plausible” or “implausible” with a button press. RTs were also recorded. In a second task, just the argument conclusions alone were presented on a computer screen, and

participants were asked to provide believability ratings on a scale of 1 (*completely unbelievable*) to 5 (*completely believable*) for each conclusion. Again, the material remained on the screen until the participant responded or the time expired (30 sec). The two components of the task were administered on different days, and the order was counterbalanced across participants.

Task Classification into Levels of Conclusion Believability

Although the task materials were developed to vary along the dimension of conclusion believability to ensure that our a priori conclusion believability judgments were consistent with those of our participants, we took the following steps:

1. A subset of normal controls ($n = 25$) rated the believability of each argument conclusion on a scale of 1 (*completely unbelievable*) to 5 (*completely believable*). (These 25 normal controls were not used in subsequent analyses.) Appendix Figure A.1A shows that the stimuli ratings spanned the full range of the scale.⁵
2. A fuzzy c-means clustering algorithm (Bezdek, 1981), implemented in R, was used to sort items into categories based on the pattern of belief ratings provided by the 25 normal controls for each item. As there was a nearly continuous range of ratings from low to high believability, we used the fuzzy c-means algorithm to split the stimuli into the three clusters shown in Appendix Figure A.2B, which were then labeled according to the average belief rating for each cluster as follows: highly believable ($M = 4.15$, $SD = 1.20$, $n = 9$), moderately believable ($M = 2.47$, $SD = 1.27$, $n = 12$), and least believable ($M = 1.35$, $SD = 0.77$, $n = 8$). These three categories were used as levels of task factor in subsequent analysis.

Data Analysis

We began by entering the plausibility scores for all patients in a voxel-based lesion symptom mapping (VLSM) analysis, with automated anatomical labeling, in ABL (Solomon et al., 2007). This technique is an exploratory method for establishing relationships between behavioral deficits and lesion status. It does not require an a priori division of patients into groups.⁶ VLSM analyses associate behavior with lesion site, on a voxel-by-voxel basis, by grouping patients into those who have a lesion in a particular voxel and comparing their behavioral task performance with that of all other patients. *t* Tests are performed at every voxel to determine if the behavioral scores for patients with lesions in that voxel are significantly different from those of all other patients (i.e., patients without lesions in that voxel). For the current study, the inference plausibility scores, for all patients, organized by the three task categories, were entered into

the analysis. A false discovery rate (FDR) correction of 0.05 was applied for multiple comparisons. The minimum cluster size was set to 10 voxels, and only those voxels where at least three patients showed overlapping lesions were used for the analysis.

We followed up with ANOVA and ANCOVA analysis of cognitive baseline scores, argument believability scores, and RTs to further clarify these results.

RESULTS

Voxel-Based Lesion Symptom Mapping Analysis

All patient CT scans and behavioral plausibility scores for the three levels of the task (least believable, moderately believable, and most believable) were entered into whole-brain VLSM analysis. A group of 10 patients, all with only lesions to left middle and superior frontal cortex (BA 8, BA 9, BA 10, BA 11; with all patients having lesions in BA 9 and BA 10; FDR corrected, $p < .001$, total $k = 964$ voxels) were identified as having significantly lower plausibility ratings on the highly believable items than all other patients. No other comparisons were significant.⁷ The details of lesion extent and location are provided in Figure 1C and Table A.1. The neuropsychological and demographic characteristics of these patients are reported in Table 2.

These results suggested that a subset of patients, all with lesions to left pFC (BA 9 and BA 10) were impaired in one component of the task. Given the fact that all selected patients had unilateral lesions to left pFC (all implicating left BA 9 and BA 10), we selected a right pFC homologue of these patients ($n = 12$) (all implicating right BA 9 and BA 10; Figure 1D; Appendix Table A.2) and a normal control group ($n = 24$; excluding the 25 used to categorize the task stimuli in terms of conclusion believability) for additional analyses, to explore the hemispheric asymmetry issue within frontal lobes.

The neuropsychological profile of these three groups is reported in Table 2. The participants' memory and IQ averages are all within the normal range. ANOVA was used to examine differences in demographic and neuropsychological scores between the right frontal, left frontal, and normal control groups. The three groups were well matched for age, education, preinjury intelligence (AFQT-7A), postinjury AFQT-A, WAIS-III full scale IQ, WMS overall general memory and working memory, D-KEFS Sorting Test, and BDI scores ($p > .05$). There was no significant difference for total brain volume loss between the right and left frontal groups. There was a marginal difference between the groups in WAIS-III verbal IQ scores, $F(2, 43) = 3.12$, $p = .054$ (but not within the Similarities subtest). Hochberg post hoc analysis

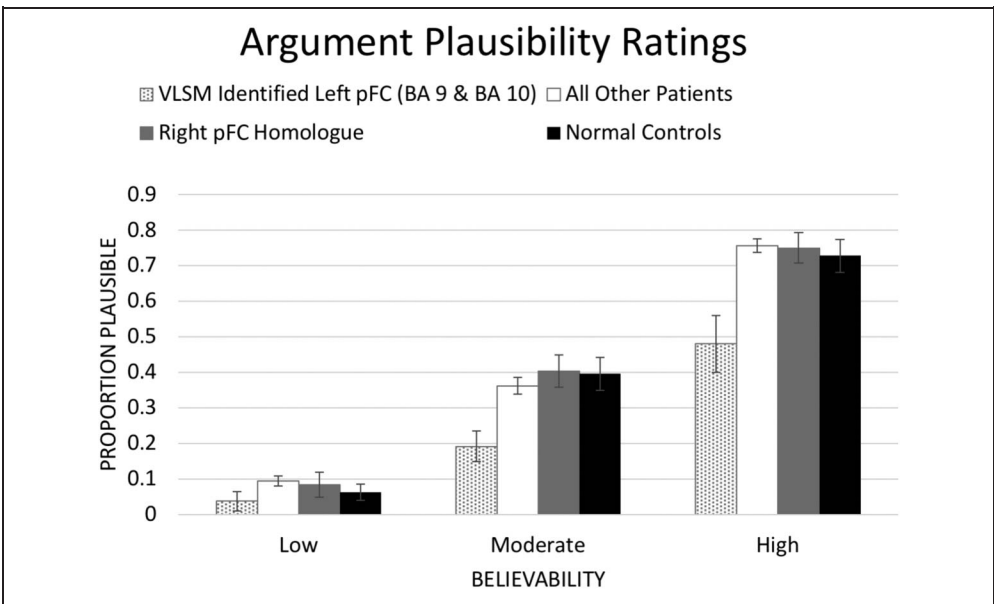
Table 2. Demographic and Neuropsychological Mean (SD) Scores for the 10 Patients with Left pFC Lesions Selected by the VLSM Analysis, 12 Patients with Right Homologue pFC Lesions, and Normal Controls

<i>Measures</i>	<i>Normal Controls (n = 24)</i>	<i>Left Frontal Lesion Patients (n = 10)</i>	<i>Right Frontal Lesion Patients (n = 12)</i>
Age (years)	58.21 (2.34)	57.80 (1.81)	57.73 (6.01)
Education	14.95 (2.36)	15.39 (2.23)	13.86 (2.2)
WAIS-III IQ			
Verbal IQ	112.86 (12.23)*	106.90 (11.73)	101.55 (10.68)*
Full IQ	112.23 (11.36)	105.44 (13.48)	103.36 (12.69)
Similarities subtest	11.38 (2.31)	12.30 (3.59)	10.50 (2.15)
Preinjury AFQT-7A	76.50 (19.77)	66.10 (20.79)	61.55 (24.74)
Postinjury AFQT-A	76.50 (18.88)	58.40 (22.25)	56.18 (26.56)
WMS			
General memory	107.13 (12.07)	102.90 (23.17)	101.91 (17.05)
Working memory	108.04 (13.72)	102.30 (11.67)	105.63 (12.55)
D-KEFS Card Sorting	11.81 (2.75)	10.90 (2.99)	9.60 (2.22)
BDI	11.00 (8.68)	5.00 (3.53)	14.27 (11.55)
Total brain volume loss (cc)		28.67 (19.27)	23.50 (14.18)

WAIS-III = Wechsler Adult Intelligence Scale III, scaled scores reported; AFQT-7A = Armed Forces Qualification Test percentile rank; WMS = Wechsler Memory Scale, scaled scores reported; D-KEFS Card Sorting = Delis-Kaplan Executive Function System Card Sorting Task (formerly the California Card Sorting Task), composite scaled scores reported; cc = cubic centimeters.

* $p = .075$ (Hochberg post hoc tests).

Figure 2. Mean proportion “plausible” responses for “impaired” subset of patients (left pFC; BA 9 and BA 10; $n = 10$) as selected by VLSM analysis in ABLe, compared with the responses of all other patients ($n = 92$), the right pFC homologue patients ($n = 12$), and normal controls ($n = 24$). Error bars show the *SEM*.



indicated that the right frontal group had a lower verbal IQ (but within normal limits) than the normal controls (mean difference = -9.79 , $p = .05$). Thus, verbal IQ was added as a covariate where possible, as noted below. A proportion test showed no significant differences in handedness between the groups, $\chi^2 = 1.39$, $df = 2$, $p = .50$ (percentage of right-handed left-damaged patients = 80.0%, right-damaged patients = 94.1%, controls = 83.3%).

To directly contrast the performance of the right pFC homologue group ($n = 12$) and the normal control group ($n = 24$) with the left pFC group ($n = 10$), we entered the behavioral plausibility rating scores of these groups, for the three task levels, into the VLSM analysis, applying an FDR correction for multiple comparisons ($p = .05$), and found the left pFC group had significantly lower plausibility ratings on the highly believable items than either the right pFC homologue

Figure 3. Strength of beliefs associated with the argument conclusions categorized by three levels of believability for the left pFC group, the right pFC homologue group, and normal controls. Error bars show the *SEM*.

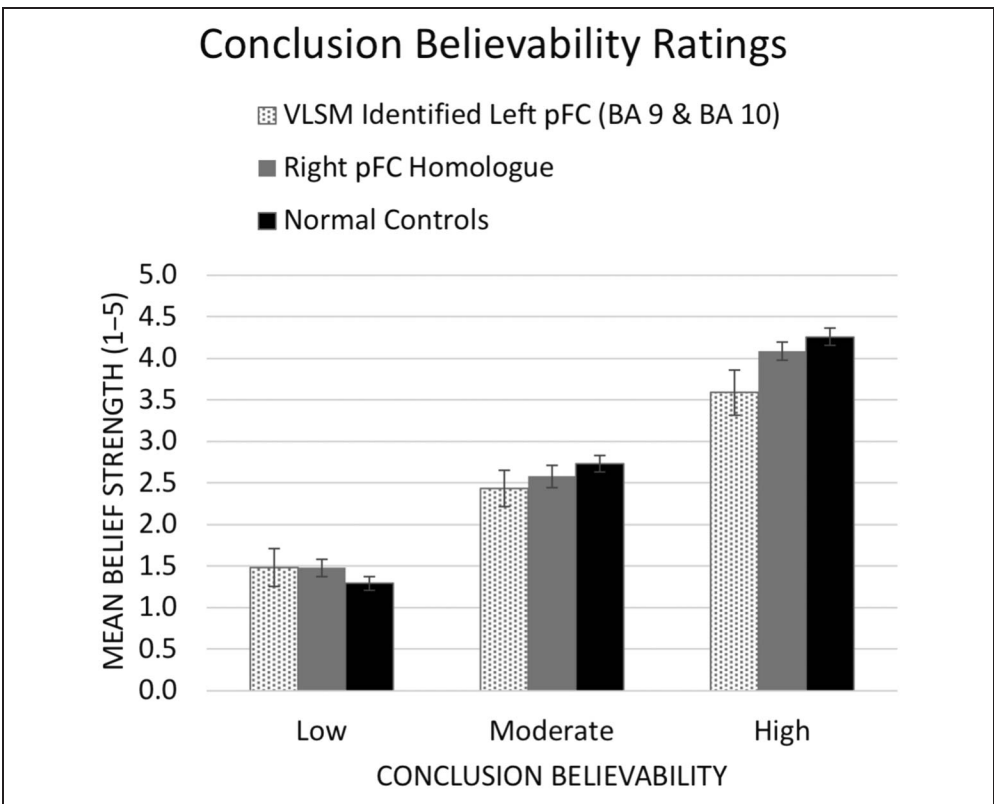
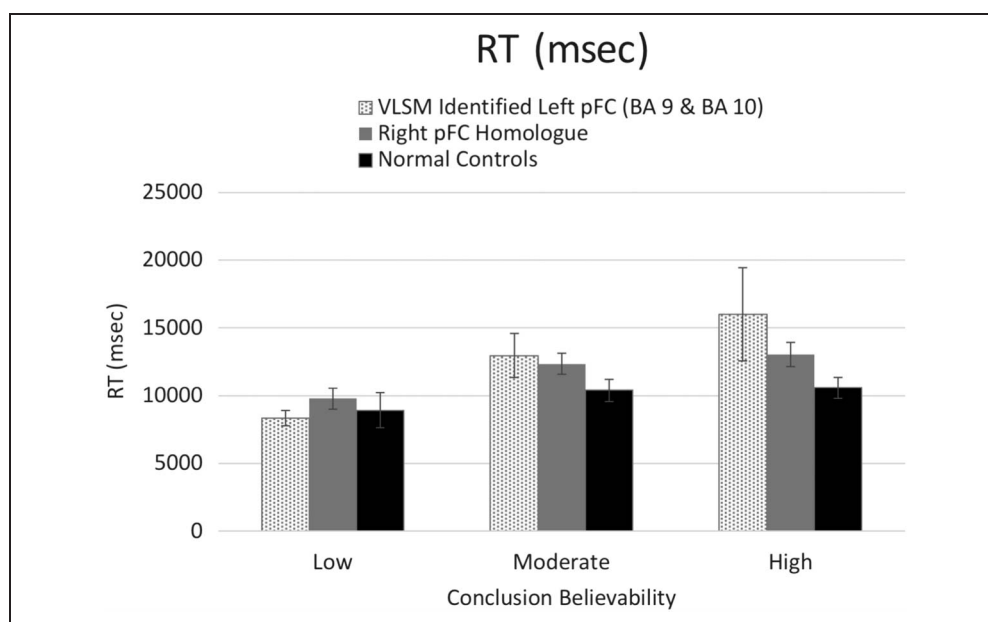


Figure 4. RTs (in msec) in the inductive reasoning task (plausibility judgments) for the left pFC group, the right pFC homologue group, and normal controls. Error bars show the SEM.



group and the normal controls. There were no significant differences between the behavioral scores of the right pFC group and normal controls. The mean responses of the various groups are plotted in Figure 2.

Subsequent Analyses

Given the critical role of beliefs in inductive inference, we also looked at the belief ratings participants provided for the conclusions of all arguments. Figure 3 shows the strength of beliefs associated with the different levels of task for each group. A mixed-models ANCOVA using belief ratings as the dependent variable revealed a significant Group \times Task interaction, $F(4, 74) = 2.7$, $p = .036$, $\eta_p^2 = .128$, driven by the fact that left frontal patients had significantly lower belief ratings than normal controls (mean difference = -0.667 , $p = .005$) and patients with right pFC lesions (mean difference = -0.551 , $p = .039$) on items with high conclusion believability (Figure 3). There were no significant group differences on belief ratings on low or moderate conclusion believability trials.

For reasons discussed above, we expected a degree of correlation between participants' believability ratings of the conclusions and their plausibility judgments of the arguments. To explore this, we carried out correlations between believability ratings and plausibility ratings for each of our three groups of participants. As expected, the Pearson correlation for normal controls yielded a positive relationship between the two variables, $r = .638$, $p = .002$. A regression analysis revealed that the believability ratings could account for 41% of the variance in plausibility ratings ($R^2 = .41$, $F(1, 19) = 13.05$, $p = .002$). Believability ratings significantly predicted plausibility ratings in normal controls ($B = .33$, $p = .002$). However, the same analyses showed nonsignificant relationships between

believability ratings and plausibility ratings for both the right pFC ($r = -.23$, $p = .52$; $R^2 = .053$, $F(1, 8) = .45$, $p = .52$; $B = -.11$, $p = .52$) and left pFC ($r = .184$, $p = .61$; $R^2 = .034$, $F(1, 8) = .285$, $p = .61$; $B = .042$, $p = .61$) patient groups.

Finally, in terms of RTs, there was a significant Group \times Task interaction, $F(4, 84) = 3.88$, $p = .006$, $\eta_p^2 = .156$, whereby the left pFC group was significantly slower than normal controls on arriving at plausibility responses on high conclusion believability trials (mean difference = -5434 msec, $p = .020$). There were no other significant differences (Figure 4).

DISCUSSION

The goal of the study was to examine the role of prior beliefs in inference and how they are modulated by unilateral pFC lesions. There are several findings that further our understanding of this issue. First, of the 102 neurological patients tested, only 10 of them, all with lesions to left pFC (all encompassing BA 9 and BA 10) found fewer arguments with highly believable conclusions "plausible" than their right hemisphere homolog counterparts ($n = 12$) and normal controls (Figure 2). There were no group differences in plausibility ratings in arguments with unbelievable conclusions, probably due to a floor effect (as everyone had very low believability ratings). There was also a significant Group \times Task interaction in terms of RTs, whereby patients with lesions to left BA 9 and BA 10 took longer to consider and respond to arguments, particularly those trials involving highly believable conclusions. These results suggest that patients with lesions to left BA 9 and BA 10 were more skeptical reasoners and are less likely to accept the conclusion.

Given our group composition (Table 2), our results cannot be explained in terms of differences in IQ, working memory, general memory, and lesion size. There were no significant group differences along these dimensions. There was a marginal difference in verbal IQ, between normal controls and patients with right pFC lesions, though both were within normal range. There were no differences in verbal or full-scale IQ between patients with left pFC lesions (i.e., the patients exhibiting the behavior under consideration), right pFC lesions, and normal controls. However, to ensure that verbal IQ would not be a confound in the study, it was used as a covariate of no interest, where possible.

Our findings are consistent with the classic split-brain patient studies (Gazzaniga, 1989, 1998; Gazzaniga & Smylie, 1984) and more recent imaging studies of belief-based inference (Liang et al., 2014; Green et al., 2006; Goel & Dolan, 2003, 2004; Luo et al., 2003; Goel et al., 1997, 2000; Wharton et al., 2000). The former studies include classic experiments such as those in which two words are presented in serial order (e.g., “pan” followed by “water”) to either the left or right hemisphere, and the patient is required to point to a picture that best depicts what happens when the words are causally related (e.g., picture of water boiling in a pan). The left hemisphere finds such inference tasks to be trivial, even to the point of being automatically compelled to make these connections, whereas the right hemisphere cannot do them (Gazzaniga, 1989, 1998; Gazzaniga & Smylie, 1984). Although these studies specifically involved causal relations, subsequent studies have generalized the findings to statistical relations (Wolford et al., 2000) and conceptual and logical relations (Goel & Dolan, 2003, 2004; Goel, Shuren, Sheesley, & Grafman, 2004; Goel et al., 2000), encompassing the variety of relations involved in this study. Our findings clarify these data by suggesting that lesions to the left pFC weaken the strength of inferential connections, resulting in hemispheric disequilibrium, allowing the right pFC to be more assertive in its checking/inhibition function, as discussed below.

Neuroimaging studies of language-based inductive inference generally indicate activation in large areas, including the left frontal and parietal lobes. For example, Goel and Dolan (2004) had participants determine the plausibility of arguments, like in this study, and reported activation in the left dorsolateral pFC (BA 9, BA 8, BA 45) for inductive reasoning. Goel et al. (1997) reported similar results with similar material.

Other studies have examined brain activation associated with judgment of analogous word pairs (Green et al., 2006), such as:

Planet : Sun versus Electron : Nucleus

Green et al. (2006) report enhanced activation of a left-sided network of parietal-frontal regions, most notably the left superior frontal gyrus. Of particular interest was

the fact that BA 9 activation was specific to the analogy condition, but not to the control condition, which involved a contiguity association. In a recent follow-up (Green et al., 2017), they showed improved performance in the analogy condition as a function of transcranial direct current stimulation of the left frontopolar cortex.

Beliefs are known to play a critical role in logical reasoning. In deductive reasoning, the belief bias effect (i.e., the finding that subjects perform more accurately when the conclusion is consistent with their beliefs [Argument D] than when it is inconsistent with their beliefs [Argument E]) is one of the oldest and most robust findings in the reasoning literature (Evans, Barston, & Pollard, 1983; Wilkins, 1928):

Argument D (congruent):

All apples are red fruit; all red fruit are nutritious; all apples are nutritious.

Argument E (incongruent):

All apples are red fruit; all red fruit are poisonous; all apples are poisonous.

In inductive reasoning, where the logical form is always invalid, the role of beliefs is integral to argument evaluation. Without some knowledge and beliefs about a situation, there is no basis for an inductive inference. As the strength of our beliefs about the subject matter increases, inferences should be more biased by these beliefs.

To account for belief bias effects, we also solicited participants' strength of belief ratings in the conclusions of arguments. This analysis yielded two interesting results. First, patients with lesions to left pFC had reduced degrees of belief in the argument conclusions compared with both normal controls and patients with lesions to right pFC. Second, the expected association between degree of conclusion believability and argument plausibility ratings occurred only in the normal controls. There was no such correlation in the patient groups. These findings suggest that the skepticism about argument plausibility displayed by patients with lesions to left pFC stems only partly from reduced strength in the believability of the conclusion. As noted in the introduction, whereas believability ratings involve evaluating the conclusion against our existing belief network, the argument plausibility ratings do introduce/trigger the information contained in the premises, which must then be evaluated in the context of the belief network.

We interpreted these results to suggest that the issue here may not simply be the diminished involvement of the left pFC but that unilateral lesions may also result in hemispheric disequilibrium. This shift in equilibrium may allow patients with lesions to the left hemisphere to be extra cautious in their inferences, perhaps due to more effective fact checking and generation of counterexamples by the right pFC. Both Marinsek, Turner, Gazzaniga, and Miller (2014) and Goel (2010, 2015) have proposed the complementary roles of hypothesis

generation and hypothesis coherency checking for the left and right pFC, respectively. It is this combined effect that may result in more skeptical inference. Indeed, the increased RTs, in arguments with the highly believable conclusions, for patients with left pFC (BA 9, BA 10) lesions are consistent with this account, in that the higher RTs may be indicative of these patients being more careful and conscientious in doing the evaluation. This pattern also appears in studies of real-world planning. Patients with lesions to the left pFC are slower than patients with right pFC lesions in generating plans, but this speed difference is accompanied by greater number of steps in revising and evaluating solutions (Goel et al., 2013).

There are relevant data from imaging, lesion, and repetitive transcranial magnetic stimulation (rTMS) studies that support this interpretation. The finding that lesions to left pFC attenuate belief strength is consistent with a recent imaging study (Kaplan et al., 2016) showing that challenges to our normal everyday beliefs (as opposed to strongly held political beliefs) show increased activity in dorsolateral pFC ($L > R$). Furthermore, activity in left dorsolateral pFC was significantly correlated with the persistence of (nonpolitical) beliefs, like in our study. Similarly, a structural MRI study (Moutsiana et al., 2015) reports the involvement of a left inferior frontal-subcortical network in belief updating.

In terms of patient data, a study of the Wason Card Selection Task (Goel et al., 2004), carried out with the same cohort of patients as this study, showed that left pFC patients, right pFC patients, and normal controls performed equally well on the arbitrary rule ("If a card has an 'A' on one side, then it must have a '4' on the other side") version of the task. However, when real-world content was introduced ("If a person is to drink alcohol, he or she must be at least 21"), the normal controls and patients with right pFC lesions benefited from the content and increased their accuracy of response from approximately 30% to 90%. As predicted by our pattern of results, the patients with the left pFC lesions failed to benefit from the presence of the content, continuing to perform at the arbitrary rule level.

An even more compelling pattern of results is provided by rTMS studies of categorical syllogisms, as in Arguments D and E. Both of these are valid arguments but differ along the very important dimension of congruency. In D, the conclusion is believable, in addition to the argument being valid (congruent). In E, the argument is still valid but the conclusion is unbelievable (incongruent). The accuracy rates of the second argument are always lower than the accuracy rates of the first argument, even though they are logically identical. This is explained as a belief bias effect.

In the congruent case, the responses based on validity and believability point in the same direction. If participants respond correctly, it is not possible to tell whether they are responding on the basis of the belief bias or the formal evaluation of the argument. However, based on

RTs and very high accuracy rates, it is generally assumed that participants are responding on the basis of belief bias. If they were responding based on logical form, we would expect lower accuracy rates (due to performance factors, such as working memory and attention span limitations). In the incongruent case, the belief bias and validity judgments point in opposite directions. Reasoners are often misled by the believability (or unbelievability) of the conclusion. In these trials it is possible to determine that if the correct response is given, then the conflict between the believability of the conclusion and validity of the argument has been detected, the belief bias inhibited, and the argument formally evaluated to generate the correct solution. Our results suggest that any impairment in the left pFC should lead to a reduction in belief bias effects. This would manifest itself as a decrease in accuracy in congruent arguments like A and an increase in accuracy for incongruent arguments like B, because the decrease in belief bias will result in greater reliance on formal logical evaluation (in both cases).

Several rTMS studies have replicated this very effect. Tsujii et al. (Tsujii, Sakatani, Masuda, Akiyama, & Watanabe, 2011; Tsujii, Masuda, Akiyama, & Watanabe, 2010) show that rTMS disruption of the left pFC not only reduces reasoning accuracy in congruent trials (Argument D) but also improves performance in incongruent trials (Argument E). That is, when the left pFC is impaired, participants are less likely to go with the believability of the conclusion and recruit other cortical regions to formally evaluate the argument.

Finally, if we make a connection between "skepticism" and "reduction in certainty," then there are also some relevant results in the literature on ill-structured problem-solving. There is evidence from ill-structured/real-world problem-solving (Goel et al., 2013; Goel, 2010) that patients with lesions to left pFC (intact right pFC) engage in problem-solving strategies that involve more revision, consideration, and double checking (and take longer) than strategies used by patients with lesions to right pFC. But interestingly, this need not be a detriment or deficit. In the planning studies referenced above, patients with left pFC lesions performed better than patients with right pFC lesions.

The last several tasks raise an interesting question regarding the real-world "detrimental" effect of these left pFC lesions. In the Wason Card Selection Task, patients with left pFC lesions are disadvantaged in the familiar rule condition (but not the formal rule condition). In the planning studies referenced above, patients with lesions to left pFC performed better than patients with right pFC lesions. In the belief bias task, left pFC disruption with rTMS was an advantage in the incongruent condition, leading to more accurate performance, but a disadvantage in the congruent condition, leading to less accurate performance.

Generally speaking, we would expect individuals with lesion profiles similar to those of our left pFC patients to

be less credulous and require a greater degree of evidence before accepting new beliefs. For instance, in the pomegranate juice example, with which we began the article, we would expect these patients to be less likely to accept the conclusion and incorporate pomegranate juice in their diets for weight loss. They may require more convincing and evidence. Whether this be detrimental or beneficial depends on the task, circumstances, and severity. If the individual is unable to make any causal and conceptual connections to draw inferences, their real-world functioning will be severely hampered. Our patients were not so severely impaired. Their level of impairment seems to be more of a shifting in equilibrium between left and right pFC. That is, optimal real-world performance requires a judicious balancing act between left and right pFC (Goel, 2015). Lesions to left pFC, such as sustained by our patients, shift this balance from left pFC to right pFC, resulting in reduced belief strength and greater evidentiary requirements. It is possible that individual differences in the equilibrium between left and right pFC result in differences in belief credulity and inference proclivity in the general population.

One puzzling result is the breakdown in correlation between believability and plausibility ratings in the right pFC patient group. A descriptive examination of the data (Figure 5) suggests that it may be a function of a very narrow range of belief ratings in the right pFC patients (Figure 5B). All the right pFC patients fall within the range of 2.6–3.0. In fact, six of the patients are at 3.0. However, the corresponding plausibility scores of these patients occupy a much wider range (0.2–0.6; Figure 5A). Twenty of the 24 normal controls fall in this range. Thus, the breakdown in the correlation may be an artifact of limited significance.

In terms of limitations of the study, four factors are worth emphasizing. First, all of our participants were male. Although there are no reasons to expect gender differences in inductive reasoning tasks, the possibility cannot be precluded. Second, as in all human lesion studies, the distribution of lesions does not give us uniform brain coverage. There may be other brain regions involved that our data do not address. Third, because the tests were carried out many decades after the penetrating traumatic brain injury, the effects of neural plasticity are unknown. Finally, the participants were in their late fifties or early sixties when tested. Again, we have no reason to think the results are not representative of all adults, but cannot rule out the possibility.

In conclusion, unilateral lesions to left pFC (BA 9, BA 10) resulted in patients being reluctant to accept common

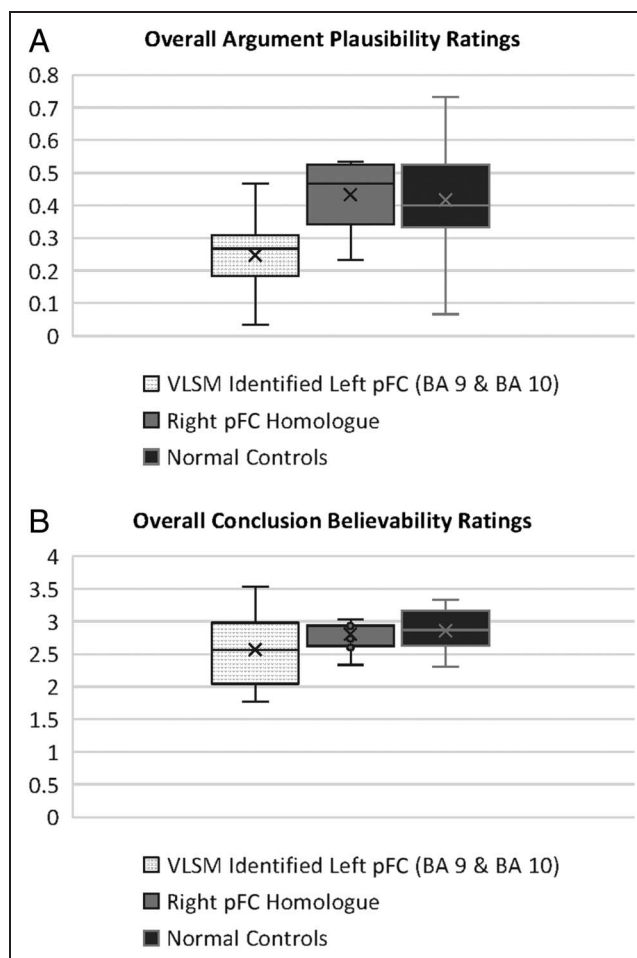


Figure 5. Box and whisker plots showing median and range of (A) conclusion believability ratings and (B) argument plausibility ratings for the left pFC group, the right pFC homologue group, and normal controls.

inductive inferences as “plausible,” at least in part due to increased incredulity about argument conclusions. We have proposed an explanation whereby unilateral lesions to left pFC result in a reduction in strength of beliefs, further augmented by the resulting disequilibrium of unilateral lesions, the latter allowing the right pFC to play a greater inhibitory role in terms of fact checking and generating counterexamples. Finally, we suggest that individual differences in belief formation and revision and inference in the general population may be underwritten by individual differences in the left and right pFC interactional dynamics.

APPENDIX A

<i>Highly Believable Conclusions</i>	<i>Moderately Believable Conclusions</i>	<i>Least Believable Conclusions</i>
All planets in our solar system rotate about their axis	Humans are omnivores	Cars have 4 wheels
Our sun rotates about its axis	Chimps are omnivores	Buses have 4 wheels
All celestial bodies rotate about their axis	All primates are omnivores	All road vehicles have 4 wheels
Salamanders are cold-blooded	Life is carbon-based	Mozart was a great musician
Salamanders are reptiles	Carbon is widespread in our galaxy	Mozart died of bleeding
All reptiles are cold-blooded	Life is widespread in our galaxy	All great musicians die of bleeding
Marathon runners have muscles	Minerals are crystalline	Latex paint is water soluble
Marathon runners are young women	Minerals are insoluble	Acrylic paint is water soluble
All young women have muscles	All insoluble matter is crystalline	All paint is water soluble
Nicotine is known to cause cancer	Lipstick is moist and glossy	Beethoven was a composer of symphonies
Chewing tobacco contains nicotine	Fish scales are moist and glossy	Beethoven became deaf
Chewing tobacco causes cancer	Fish scales are used to make lipstick	All composers of symphonies become deaf
Humans are mortal	Some strains of mice are lab mice	Boeing makes military airplanes
Humans are living beings	Lab mice can run mazes	Boeing makes passenger airplanes
All living beings are mortal	All strains of mice can run mazes	All airplanes are made by Boeing
Lightning is electrical discharge	T-Rex were dinosaurs	Fido Barkwell is a pit bull
Electrical discharge generates heat	T-Rex had scaly skin	Fido Barkwell is overweight
Expansion of heated air causes thunder	All dinosaurs had scaly skin	All pit bulls are overweight
Mushrooms have finite existence	T-Rex were dinosaurs	The clock at St. Mary is very accurate
Mushrooms are life forms	T-Rex buried their eggs	The clock at St Peter is very accurate
All life forms have a finite existence	All dinosaurs buried their eggs	Clocks at all churches are very accurate
Smoking causes cancer in women	Coffee is a stimulant	Rexdale is a German Shepherd
Smoking causes cancer in men	Drinking coffee reduces suicide rates in women	Rexdale lives in Dusseldorf
Smoking causes cancer in all primate	Drinking coffee will reduce suicide rates in men	All German Shepherds live in Dusseldorf

(continued)

Highly Believable Conclusions	Moderately Believable Conclusions	Least Believable Conclusions
Felix Prowlton is a Burmese tiger	Diamonds were formed under great heat	
Felix Prowlton consumes meat	Sapphires were formed under great heat	
All Burmese tigers consume meat	All gemstones were formed under great heat	
	The Pinto tribe lives in Uganda	
	Pintos are pygmies	
	All Pygmy tribes are from Uganda	
	Cheetahs cannot acquire HIV	
	Cheetahs are felines	
	No felines can acquire HIV	
	Swedes carry coliform bacteria	
	Finns carry the coliform bacteria	
	All humans carry the coliform bacteria	

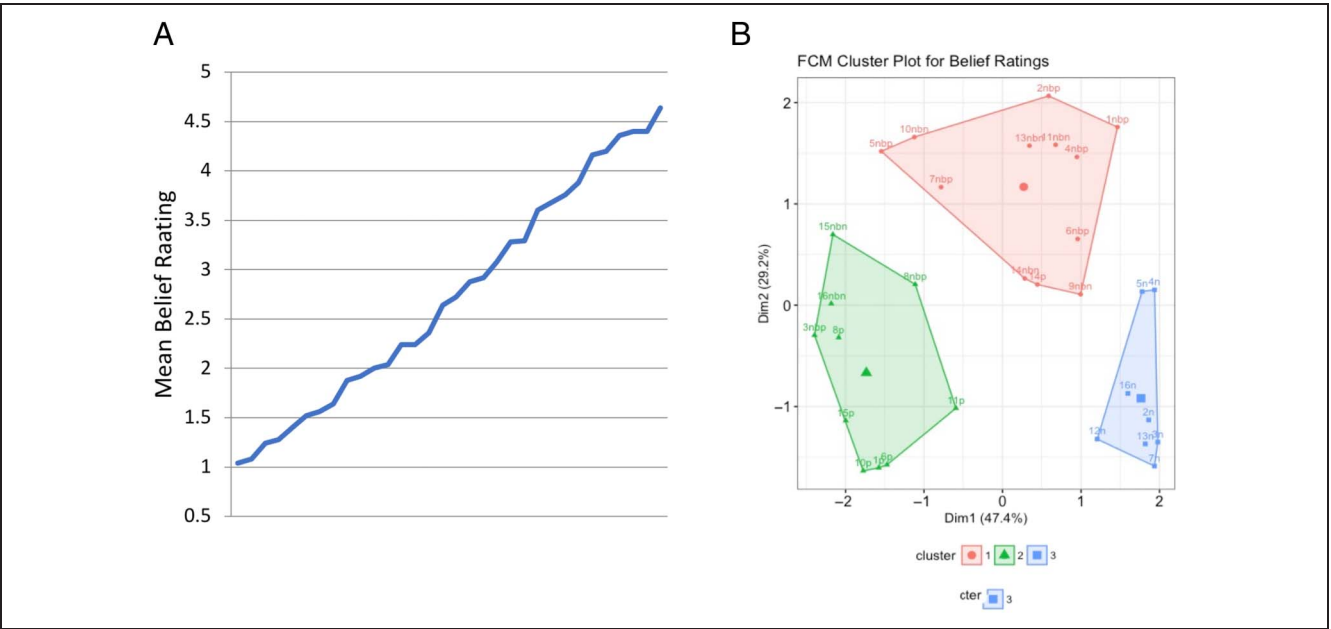


Figure A.1. (A) Mean belief ratings of believability of the conclusions of the 30 induction items arranged in ascending order along the x-axis. These ratings were provided by 25 normal controls not used in subsequent analyses. (B) Plot showing the three clusters generated by the fuzzy c-means algorithm for the 30 induction items. Cluster 1, depicted as circles, represents moderately believable items. Cluster 2, depicted as triangles, represents highly believable items. Cluster 3, depicted as squares, represents low believable items (i.e., unbelievable). The clusters are graphed along the two principal components that account for most of the variability along the five dimensions of belief ratings that were imputed for each item. These two components account for 76.6% of the variability.

Table A.1. Percentage of Brodmann's Areas Implicated in Patients with Left pFC Lesions

Left pFC Patients (N = 10)		Brodmann's Areas with Percent Damage																		
		6	8	9	10	11	13	22	24	25	28	31	32	35	36	38	44	45	46	47
182		0	8.03	33.49	70.46	63.47	2.67	1.24	0	0.08	0	0	4.04	0	0	0	3.46	44.25	46.87	26.02
309		0	0	0.98	20.32	0	0	0	0	0	0	0	0.37	0	0	0	0	0	9.08	0
318		0	0	2.18	46.13	32.99	0.29	0	0	15.93	0	0	5.12	0	0	0	0	0	0	10.64
524		17.59	45.07	8.65	0.05	0	0	0	6.11	0	0	0	16.25	0	0	0	0	0	0	0
1216		0	0	0.53	54.22	36.39	0	0	0	0	0	0	0.71	0	0	0	0	0	3.03	0
1273		0.5	47.98	36.15	7.5	0	0	0	0.36	0	0	0	3.02	0	0	0.06	0	0	5.97	0.2
1347		24.3	42.57	6.24	0.06	0	0.52	0	4.2	0	4.57	0.04	9.37	28.96	0.59	0	0	0	0	0
1364		0	0	2.8	26.12	0.9	0	0	0	0	0	0	5.2	0	0	0	0	0	0	0
2221		0	5.08	9.91	0.4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2354		0	0	2.26	34.72	3.35	0	0	0.12	0	0	0	8.93	0	0	0	0	0	0	1.27

Table A.2. Percentage of Brodmann's Areas Implicated in Patients with Right pFC Lesions

Right Frontal Patients (N = 12)		Brodmann's Areas with Percent Damage																											
Pot. ID	1	2	3	4	5	6	7	8	9	10	11	13	21	22	24	25	27	30	31	32	40	42	43	44	45	46	47		
430	0	0	0	0	0	0.25	0	7.18	15.9	5.62	0	0	0	0	0	0	0	0	0	1.84	0	0	0	0	0	0	0	0	
529	0	0	0	0	0	0.01	0	15.54	16.25	0.37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.35	0	0	
1124	0	0	0.14	2.23	0	3.02	0	0	1.73	20.31	5.39	13.72	0	0	0	0	0	0	0	6.21	0	0	0	3.77	16.01	15.59	18.89	0	
1741	0	0	0	0	0	0	0	32.93	26.52	4.28	0	0	0	0	0.02	0	0	0	0	1.63	0	0	0	0	0	8.87	0	0	
2034	0	0	0	0	0	0	0	0	1.5	54.51	31.04	0.08	0	0	0	0.31	0	0	0	2.51	0	0	0	0	0.24	0.85	1.94	0	
2101	0	0	0	0	0	0	0	0	10.09	16.84	0	0	0	0	0	0	1.25	0.49	0	1.76	0	0	0	0	0	0	0	0	
2196	12.34	2.13	4.87	0.38	0	0.86	0	0	2.82	26.14	9.68	0	0	0	0	0	0	0	0	3.99	0	0	0	0	0	0.3	0	0	
2218	0	0	0	0	0	0.19	0	9.02	9.6	46.55	21.97	0	0	0	0	0	0	0	0	0.46	0	0	0	1.95	19.9	26.75	16.71	0	
2455	6.33	0.41	9.49	21.11	0	20.02	0	12.22	23.32	14.84	0.15	5.73	0.27	17.91	0.34	0	0	0	0.04	0	0.07	7.32	51.3	78.89	35.34	17.36	3.41	0	
3013	0	0	0	0	0	0	0	0.46	3.42	46.18	21.4	0	0	0	0	1.95	0	0	0	2.64	0	0	0	0	0	2.2	0	0	
3068	0	0	0	0	0	0	0	9.95	17.95	24.73	0	0	0	0	0	0	0	0	0	0.69	0	0	0	0	0	0	0	0	
3125	0	0	0	0	1.05	0.05	1.94	11.53	25.33	12.19	0	0	0	0	1.31	0	0	0	0.25	0.15	0	0	0	0	0	9.24	0	0	

Acknowledgments

This work was funded, in part, by an National Sciences and Engineering Council of Canada grant to V. G. This research was also partly supported by the Therapeutic Cognitive Neuroscience Fund to J. G.

Reprint requests should be sent to Vinod Goel, Department of Psychology, York University, 4700 Keele St., Toronto, Canada, M3J1P3, or via e-mail: vgoel@yorku.ca.

Notes

1. There is actually a very large third literature on delusional beliefs in schizophrenic patients (Krueger & Grafman, 2012; Chapman & Chapman, 1988). However, this literature deals with abnormal beliefs. We are concerned with normal, everyday beliefs. We believe that an understanding of normal beliefs is a prerequisite for understanding abnormal beliefs.
2. As there are no wrong and right answers in belief evaluations and inductive inference, the issue here is one of difference in performance rather than deficit.
3. Twenty-five of the 49 normal controls were utilized to obtain believability strength ratings of the conclusions of the arguments. They were excluded from further analysis, leaving a total of 24 normal controls.
4. The AFQT is a standardized test that measures a candidate's abilities in the areas of paragraph comprehension, word knowledge, arithmetic reasoning, and mathematics knowledge. It is administered to all enlisted members (not officers) of the U.S.

Armed Forces. The scores (while reported as percentiles) correlate highly with WAIS-III IQ scores (Grafman et al., 1988).

5. Two items were removed from the analysis since there was a general lack of consensus in the ratings for these items.

6. Despite our a priori hypothesis implicating left pFC patients in this task, we chose not to manually select patients to see if any other patients might also be impaired or if it was indeed only left pFC patients.

7. It should be noted that part of the power of detecting differences in various parts of the brain is contingent of having enough patients with overlapping lesions in those areas, so it is possible this effect could be found elsewhere if there were more overlapping patients.

REFERENCES

- Beck, A. T. (1987). *Beck depression inventory*. San Antonio, TX: The Psychological Corporation.
- Bezdek, J. C. (1981). *Pattern recognition with fuzzy objective function algorithms*. New York: Plenum Press.
- Chapman, L. J., & Chapman, J. P. (1988). The genesis of delusions. In T. F. Oltmanns & B. A. Maher (Eds.), *Delusional beliefs*. New York: Wiley.
- Damasio, H., & Damasio, A. R. (1989). *Lesion analysis in neuropsychology*. Oxford, United Kingdom: Oxford University Press.
- Eitelberg, M. J., Laurence, J. H., Waters, B. K., & Perelman, L. S. (1984). *Screening for service: Aptitude and education*

- criteria for military entry*. Revised (No. FR-PRD-83-24). Alexandria, Virginia: Human Resources Research Organization.
- Evans, J., Barston, J., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory & Cognition*, 11, 295–306.
- Gazzaniga, M. S. (1989). Organization of the human brain. *Science*, 245, 947–952.
- Gazzaniga, M. S. (1998). *The mind's past*. Berkeley, CA: University of California Press.
- Gazzaniga, M. S., & Smylie, C. S. (1984). Dissociation of language and cognition. A psychological profile of two disconnected right hemispheres. *Brain*, 107, 145–153.
- Goel, V. (2010). Neural basis of thinking: Lab problems versus real-world problems. *Wiley Interdisciplinary Reviews. Cognitive Science*, 1, 613–621.
- Goel, V. (2015). Indeterminacy tolerance as a basis of hemispheric asymmetry within prefrontal cortex. *Frontiers in Human Neuroscience*, 9, 326.
- Goel, V., Buchel, C., Frith, C., & Dolan, R. J. (2000). Dissociation of mechanisms underlying syllogistic reasoning. *Neuroimage*, 12, 504–514.
- Goel, V., & Dolan, R. J. (2003). Explaining modulation of reasoning by belief. *Cognition*, 87, B11–B22.
- Goel, V., & Dolan, R. J. (2004). Differential involvement of left prefrontal cortex in inductive and deductive reasoning. *Cognition*, 93, B109–B121.
- Goel, V., Gold, B., Kapur, S., & Houle, S. (1997). The seats of reason? An imaging study of deductive and inductive reasoning. *NeuroReport*, 8, 1305–1310.
- Goel, V., Shuren, J., Sheesley, L., & Grafman, J. (2004). Asymmetrical involvement of frontal lobes in social reasoning. *Brain*, 127, 783–790.
- Goel, V., Vartanian, O., Bartolo, A., Hakim, L., Ferraro, A. M., Isella, V., et al. (2013). Lesions to right prefrontal cortex impair real-world planning through premature commitments. *Neuropsychologia*, 51, 713–724.
- Gozzi, M., Zamboni, G., Krueger, F., & Grafman, J. (2010). Interest in politics modulates neural activity in the amygdala and ventral striatum. *Human Brain Mapping*, 31, 1763–1771.
- Grafman, J., Jonas, B. S., Martin, A., Salazar, A. M., Weingartner, H., Ludlow, C., et al. (1988). Intellectual function following penetrating head injury in Vietnam veterans. *Brain*, 111, 169–184.
- Green, A. E., Fugelsang, J. A., Kraemer, D. J., Shamos, N. A., & Dunbar, K. N. (2006). Frontopolar cortex mediates abstract integration in analogy. *Brain Research*, 109, 125–137.
- Green, A. E., Spiegel, K. A., Giangrande, E. J., Weinberger, A. B., Gallagher, N. M., & Turkeltaub, P. E. (2017). Thinking cap plus thinking zap: tDCS of frontopolar cortex improves creative analogical reasoning and facilitates conscious augmentation of state creativity in verb generation. *Cerebral Cortex*, 27, 2628–2639.
- Harris, S., Kaplan, J. T., Curiel, A., Bookheimer, S. Y., Iacoboni, M., & Cohen, M. S. (2009). The neural correlates of religious and nonreligious belief. *PLOS ONE*, 4, e7272.
- Heled, E., Hoofien, D., Margalit, D., Natovitch, R., & Agranov, E. (2012). The Delis–Kaplan Executive Function System Sorting Test as an evaluative tool for executive functions after severe traumatic brain injury: A comparative study. *Journal of Clinical and Experimental Neuropsychology*, 34, 151–159.
- Kaplan, J. T., Gimbil, S. I., & Harris, S. (2016). Neural correlates of maintaining one's political beliefs in the face of counterevidence. *Scientific Reports*, 6, 39589.
- Kapogiannis, D., Barbey, A. K., Su, M., Zamboni, G., Krueger, F., & Grafman, J. (2009). Cognitive and neural foundations of religious belief. *Proceedings of the National Academy of Sciences, U.S.A.*, 106, 4876–4881.
- Krawczyk, D. (2018). *Reasoning: The neuroscience of how we think*. New York: Academic Press.
- Krueger, F., & Grafman, J. (2012). *The neural basis of human belief systems*. London: Psychology Press.
- Liang, P., Goel, V., Jia, X., & Li, K. (2014). Different neural systems contribute to semantic bias and conflict detection in the inclusion fallacy task. *Frontiers in Human Neuroscience*, 8, 797.
- Luo, Q., Perry, C., Peng, D., Jin, Z., Xu, D., Ding, G., et al. (2003). The neural substrate of analogical reasoning: An fMRI study. *Cognitive Brain Research*, 17, 527–534.
- Makale, M., Solomon, J., Patronas, N. J., Danek, A., Butman, J. A., & Grafman, J. (2002). Quantification of brain lesions using interactive automated software. *Behavior Research Methods, Instruments, & Computers*, 34, 6–18.
- Marinsek, N., Turner, B. O., Gazzaniga, M., & Miller, M. B. (2014). Divergent hemispheric reasoning strategies: Reducing uncertainty versus resolving inconsistency. *Frontiers in Human Neuroscience*, 8, 839.
- Moutsiana, C., Charpentier, C. J., Garrett, N., Cohen, M. X., & Sharot, T. (2015). Human frontal-subcortical circuit and asymmetric belief updating. *Journal of Neuroscience*, 35, 14077–14085.
- Raymont, V., Salazar, A., Krueger, F., & Grafman, J. (2011). “Studying injured minds”—The Vietnam head injury study and 40 years of brain injury research. *Frontiers in Neurology*, 2, 15.
- Reverberi, C., Lavaroni, A., Gigli, G. L., Skrap, M., & Shallice, T. (2005). Specific impairments of rule induction in different frontal lobe subgroups. *Neuropsychologia*, 43, 460–472.
- Shallice, T., & Cooper, R. (2011). *The organization of mind*. Oxford, United Kingdom: Oxford University Press.
- Sloman, S., & Lagnado, D. A. (2005). The problem of induction. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 95–116). Cambridge, United Kingdom: Cambridge University Press.
- Solomon, J., Raymont, V., Braun, A., Butman, J. A., & Grafman, J. (2007). User-friendly software for the analysis of brain lesions (ABLE). *Computer Methods and Programs in Biomed*, 86, 245–254.
- Stanovich, K. E. (1999). *Who is rational?: Studies of individual differences in reasoning*. New York, NY: Psychology Press.
- Stanovich, K. E., & West, R. F. (1997). Reasoning independently of prior belief and individual differences in actively open-minded thinking. *Journal of Educational Psychology*, 89, 342–357.
- Stollstorff, M., Vartanian, O., & Goel, V. (2012). Levels of conflict in reasoning modulate right lateral prefrontal cortex. *Brain Research*, 1428, 24–32.
- Tsujii, T., Masuda, S., Akiyama, T., & Watanabe, S. (2010). The role of inferior frontal cortex in belief-bias reasoning: An rTMS study. *Neuropsychologia*, 48, 2005–2008.
- Tsujii, T., Sakatani, K., Masuda, S., Akiyama, T., & Watanabe, S. (2011). Evaluating the roles of the inferior frontal gyrus and superior parietal lobule in deductive reasoning: An rTMS study. *Neuroimage*, 58, 640–646.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*, 15, 273–289.
- Weber, R., & Crocker, J. (1983). Cognitive processes in the revision of stereotypic beliefs. *Journal of Personality and Social Psychology*, 45, 961–977.
- Wechsler, D. (1997a). *WAIS-III*. San Antonio, TX: Psychological Corporation.
- Wechsler, D. (1997b). *Wechsler Memory Scale (WMS-III)*. San Antonio, TX: Psychological Corporation.
- Wharton, C. M., Grafman, J., Flitman, S. S., Hansen, E. K., Brauner, J., Marks, A., et al. (2000). Toward neuroanatomical models of analogy: A positron emission tomography study of analogical mapping. *Cognitive Psychology*, 40, 173–197.
- Wilkins, M. C. (1928). The effect of changed material on the ability to do formal syllogistic reasoning. *Archives of Psychology*, 16, 5–83.
- Wolford, G., Miller, M. B., & Gazzaniga, M. (2000). The left hemisphere's role in hypothesis formation. *Journal of Neuroscience*, 20, RC64.