



Lesion network localization of criminal behavior

R. Ryan Darby^{a,b,c,d,e,1}, Andreas Horn^{b,c,f}, Fiery Cushman^g, and Michael D. Fox^{b,c,d,h,1}

^aDepartment of Neurology, Vanderbilt University Medical Center, Nashville, TN 37232; ^bBerenson-Allen Center for Non-Invasive Brain Stimulation, Department of Neurology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA 02215; ^cDivision of Cognitive Neurology, Department of Neurology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA 02115; ^dDepartment of Neurology, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114; ^eDepartment of Neurology, McLean Psychiatric Hospital, Harvard Medical School, Belmont, MA 02478; ^fDepartment of Neurology, Movement Disorder and Neuromodulation Unit, Charité-University Medicine, 10117 Berlin, Germany; ^gDepartment of Psychology, Harvard University, Cambridge, MA 02138; and ^hAthinoula A. Martinos Centre for Biomedical Imaging, Charlestown, MA 02129

Edited by Kent A. Kiehl, The Mind Research Network, Albuquerque, NM, and accepted by Editorial Board Member Michael S. Gazzaniga November 20, 2017 (received for review April 25, 2017)

Following brain lesions, previously normal patients sometimes exhibit criminal behavior. Although rare, these cases can lend unique insight into the neurobiological substrate of criminality. Here we present a systematic mapping of lesions with known temporal association to criminal behavior, identifying 17 lesion cases. The lesion sites were spatially heterogeneous, including the medial prefrontal cortex, orbitofrontal cortex, and different locations within the bilateral temporal lobes. No single brain region was damaged in all cases. Because lesion-induced symptoms can come from sites connected to the lesion location and not just the lesion location itself, we also identified brain regions functionally connected to each lesion location. This technique, termed lesion network mapping, has recently identified regions involved in symptom generation across a variety of lesion-induced disorders. All lesions were functionally connected to the same network of brain regions. This criminality-associated connectivity pattern was unique compared with lesions causing four other neuropsychiatric syndromes. This network includes regions involved in morality, value-based decision making, and theory of mind, but not regions involved in cognitive control or empathy. Finally, we replicated our results in a separate cohort of 23 cases in which a temporal relationship between brain lesions and criminal behavior was implied but not definitive. Our results suggest that lesions in criminals occur in different brain locations but localize to a unique resting state network, providing insight into the neurobiology of criminal behavior.

morality | functional connectivity | criminality | sociopathy | lesion

Criminal behavior comes at an enormous cost to society, upwards of \$1 trillion dollars by some estimates (1). Despite this burden, the neurobiological substrate underlying criminal behavior remains unclear (2). To address this, some previous studies have used brain imaging, which have identified a variety of abnormalities in criminals (3–5). In most cases, however, it is unclear whether these abnormalities are a cause, compensation, or incidental correlate of criminality (6). This distinction is important, both for neuroscience and the legal field, given the increasing use of neuroimaging in criminal court cases (7).

Another approach is to study patients who develop criminal behavior following focal brain lesions, referred to as “pseudopsychopathy” or “acquired sociopathy” (8, 9). Although many factors besides brain lesions contribute to criminality (2, 10–14), the lesion is often interpreted as causal in the sense that it contributes or predisposes to the behavior (2, 6). This causal inference is strengthened when there is a clear temporal relationship between the lesion and the criminal behavior (6), although often this temporal association is unclear.

Famous examples of acquired sociopathy include Phineas Gage, who developed antisocial personality changes after an iron rod blast damaged his medial frontal lobes, and Charles Whitman, who murdered 16 people following growth of a brain tumor in his right temporal lobe (15). This pair of cases also illustrates a problem with lesion-based localization: different cases often implicate different brain regions. Patients such as Gage have damage to the ventromedial prefrontal cortex (vmPFC) and orbital prefrontal cortex (16), while other patients like Whitman

have damage outside this area (15). Despite several reviews (5, 17–19), there is currently no widely accepted and parsimonious explanation for this heterogeneity.

Our approach is motivated by the observation that neurologic symptoms can come from dysfunction in remote brain regions connected to the lesion location rather than from the lesion location itself, a phenomenon referred to as diaschisis (20). A new technique termed lesion network mapping can potentially account for these remote effects (21–25). This technique uses a database of resting state functional connectivity from normal subjects to identify regions functionally connected to each lesion location (*Materials and Methods*). Because lesions causing the same symptom tend to share functional connectivity to the same brain regions, this approach has provided insight into complex neurologic symptoms, including hallucinations (21), involuntary movements (25), delusions (23), and loss of consciousness (24).

Here we report a systematic mapping of lesion locations temporally associated with criminal behavior. We used lesion network mapping to test (i) whether lesions temporally associated with criminal behavior are part of a common brain network, and (ii) whether this network overlaps regions activated by neuropsychological processes hypothesized to be abnormal in criminals (19, 26, 27). Finally, we tested whether our results can be replicated in a second cohort in which the temporal relationship between lesion onset and criminal behavior is uncertain.

Significance

Cases like that of Charles Whitman, who murdered 16 people after growth of a brain tumor, have sparked debate about why some brain lesions, but not others, might lead to criminal behavior. Here we systematically characterize such lesions and compare them with lesions that cause other symptoms. We find that lesions in multiple different brain areas are associated with criminal behavior. However, these lesions all fall within a unique functionally connected brain network involved in moral decision making. Furthermore, connectivity to competing brain networks predicts the abnormal moral decisions observed in these patients. These results provide insight into why some brain lesions, but not others, might predispose to criminal behavior, with potential neuroscience, medical, and legal implications.

Author contributions: R.R.D., F.C., and M.D.F. designed research; R.R.D. performed research; A.H. contributed new reagents/analytic tools; R.R.D. analyzed data; and R.R.D. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. K.A.K. is a guest editor invited by the Editorial Board.

This open access article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](https://creativecommons.org/licenses/by-nc-nd/4.0/).

See Commentary on page 451.

¹To whom correspondence may be addressed. Email: darby.ryan@gmail.com or foxmdphd@gmail.com.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1706587115/-DCSupplemental.

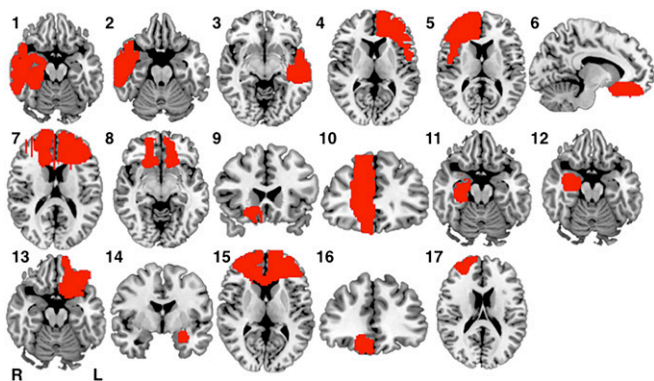


Fig. 1. Lesions temporally associated with criminal behavior. Lesions from 17 patients with acquired criminal behavior, manually traced onto a common brain atlas (MNI).

Results

Lesions Temporally Associated with Criminal Behavior Are Spatially Heterogeneous. Seventeen patients with a documented temporal relationship between a brain lesion and criminal behavior were identified through a systematic literature search (*SI Appendix, Fig. S1 and Table S1*). Criminal behaviors included “white collar” crimes, such as fraud or theft; however, most of the patients (12 of 17) had committed violent crimes, such as assault, rape, and murder. Cases included documentation of no criminal behavior before the lesion (15 cases) or resolution of criminal behavior following treatment of the lesion (two cases). The 17 lesions were spatially diverse, including nine in the medial frontal or orbitofrontal structures, three in the medial temporal lobe/amygdala, three in the anterior lateral temporal lobe, one in the dorsomedial prefrontal cortex, and one involving the ventral striatum and parts of the orbitofrontal cortex (Fig. 1). Although the most common lesion location was the vmPFC/orbitofrontal cortex, at least seven lesions were documented to not extend into this area.

Network Localization of Lesions Temporally Associated with Criminal Behavior. The set of regions functionally connected to each lesion location was determined using a database of resting state functional connectivity from normal subjects (21, 28, 29). Regions with activity positively or negatively correlated with activity at each lesion site were identified (Fig. 2*B*). Finally, lesion network maps from each patient were overlapped to determine brain regions significantly connected to all or most lesions causing criminal behavior (Fig. 2*C*).

All 17 lesions temporally associated with criminal behavior were functionally connected (i.e., positively correlated) to the inferior orbitofrontal cortex and anterior temporal lobes, and most (16 of 17) were connected to the vmPFC and nucleus accumbens (Fig. 3*A* and *SI Appendix, Table S2*). In addition, all 17 lesions were functionally connected (i.e., negatively correlated) with the intraparietal sulcus, and 15 of 17 were functionally connected with the dorsolateral prefrontal cortex (Fig. 3*A* and *SI Appendix, Table S2*). Statistical analysis showed that this pattern of connectivity was specific to lesions temporally associated with criminal behavior compared with lesions causing four other neurologic syndromes (21) (Fig. 3*B* and *SI Appendix, Table S3*). There was no difference in connectivity between early-onset (age <18 years; $n = 7$) and later-onset (age >18 years; $n = 10$) lesions. These results suggest that lesions in different locations temporally associated with criminal behavior are characterized by a unique pattern of brain connectivity. Thus, while these lesions are spatially diverse, they are part of a common functional network.

Network Localization to Moral Decision-Making Regions. Criminality is presumed to arise in part from abnormalities in moral decision making (17, 19). We therefore hypothesized that lesions temporally associated with criminal behavior would be functionally

connected to regions involved in moral decision making in normal subjects. We identified moral decision-making regions in two ways: (i) activation likelihood estimation (ALE) of regions activated by moral vs. nonmoral stimuli (Fig. 4, light green) (30, 31) and (ii) activations associated with the keyword “moral” on the website Neurosynth (www.neurosynth.org), an automated engine for performing neuroimaging meta-analyses that avoids potential bias associated with study selection (Fig. 4, dark green) (32).

We then quantitatively tested whether lesion locations causing criminal behavior were functionally connected to regions involved in moral decision making. Specifically, we used our resting-state functional connectivity dataset to compute the temporal correlation between spontaneous brain activity recorded from each lesion location with activity recorded from binarized maps of our morality meta-analyses. As a group, lesion locations temporally associated with criminal behavior were functionally connected to regions activated by moral tasks ($P < 0.0001$). This connectivity was significantly greater than for lesion locations causing other neurologic syndromes ($P < 0.0001$; Fig. 4*B* and *C*). Results were independent of the meta-analysis method used to identify moral decision-making regions.

Network Localization to Regions Associated with Subcomponents of Moral Decision Making. Our results indicate that lesions temporally associated with criminal behavior are uniquely functionally connected to regions involved in moral decision making. However, moral decision making incorporates several different neuropsychological processes including cognitive control (33), value or reward-based decision making (34), theory of mind (35), and empathy (27). The potential contributions of each process to criminal behavior are contested (36).

We approach this problem at a network level of analysis: how does the functional network that we identify in lesion-induced criminal behavior align with brain regions implicated in these subcomponents of moral decision making? We used Neurosynth to perform unbiased meta-analyses of “value,” “theory of mind,” “empathy,” and “cognitive control.” As expected, brain regions identified in each of these meta-analyses partially overlapped portions of our original morality meta-analysis (*SI Appendix, Fig. S2*). Lesions temporally associated with criminal behavior were functionally connected to regions involved in value-based decision making ($P_c < 0.0001$; Fig. 5*A*) and theory of mind ($P_c < 0.0001$; Fig. 5*B*) but not with empathy (Fig. 5*C*) or cognitive control (Fig. 5*D*). Lesions were significantly more connected to value-based

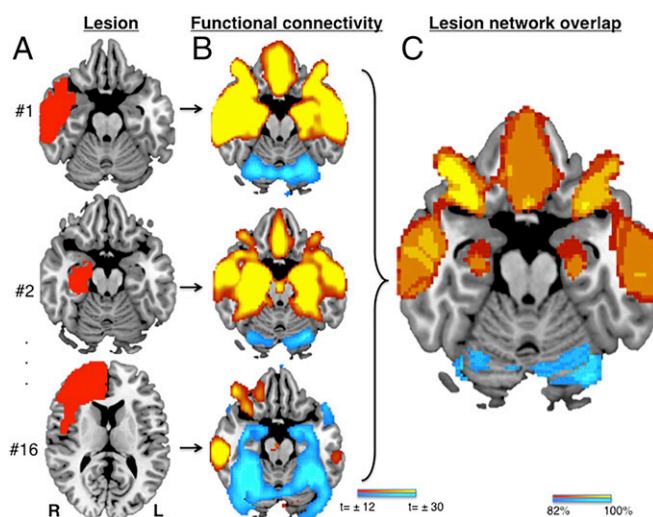


Fig. 2. Lesion network mapping technique. (A) Lesions traced onto a standardized MNI brain template. (B) Brain regions functionally connected to each lesion location based on a large resting-state functional connectivity database. (C) Overlap in the functional connectivity maps from each lesion identifies brain regions functionally connected to the greatest number of lesion locations.

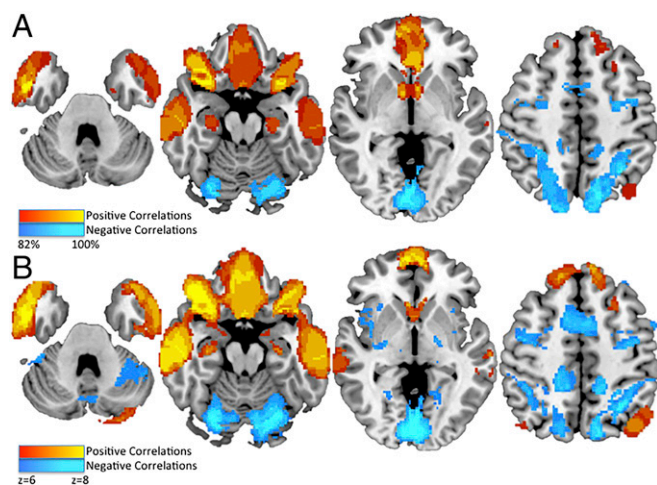


Fig. 3. Lesions temporally associated with criminality are part of a unique connected brain network. (A) Regions functionally connected to 17 lesion locations temporally associated with criminal behavior. (B) Nonparametric voxel-wise lesion-symptom mapping results showing regions in which 17 lesions temporally associated with criminal behavior were significantly more connected (positive in red, negative in blue) compared with the connectivity of 63 lesions causing other syndromes.

decision making and theory of mind regions than empathy or cognitive control regions (all comparisons $P < 0.01$). Finally, this connectivity was specific to lesions temporally associated with criminal behavior compared with lesions causing other neurologic syndromes ($P < 0.0001$; Fig. 5).

Network Localization to Competing Brain Networks Matches Abnormal Moral Choices. The foregoing analyses focus on brain regions with positive correlations to lesion locations; however, a crucial theme in both functional network analysis (37) and moral decision making (33, 38) is the role of opponent networks—those exhibiting systematic negative correlations. We next explored whether known behavioral abnormalities in individuals with criminal behavior are consistent with lesion connectivity to opponent networks.

We focused on the widely studied conflict between competing networks involved in resolving ambiguous moral dilemmas, such as the “trolley problem” (33, 38). One network is associated with a strong aversion to directly harming others, while the second network is involved in overcoming this aversion to make more utilitarian decisions (e.g., harm one person to save the lives of five persons). Patient populations with increased risk of criminal behavior [e.g., those with vmPFC lesions, frontotemporal dementia, psychopathy] are more likely than normal subjects to endorse harming others to make more utilitarian decisions (39–41). We hypothesized that lesions temporally associated with criminal behavior would show differential connectivity to these competing brain regions.

Indeed, activity from our lesion locations was positively correlated with brain regions activated when deciding to avoid personal harm ($P < 0.0001$; Fig. 6A) and negatively correlated with brain regions activated by utilitarian decisions ($P < 0.0001$; Fig. 6B). In both cases, results were specific to lesion locations temporally associated with criminal behavior vs. lesions causing other syndromes ($P < 0.0001$). Crucially, the combination of connectivity to these two competing sets of brain regions was a better predictor of lesion-induced criminal behavior than was connectivity to either set alone ($P < 0.01$).

Patient populations with increased criminality also tend to reject unfair offers more than normal subjects during the ultimatum game (42–44). As above, competing brain regions are associated with the decision to reject an unfair offer made by another player vs. a decision to accept a fair offer (45). Lesion locations temporally associated with criminal behavior were negatively correlated with the brain region most activated when rejecting unfair offers ($P < 0.05$) and positively correlated with

the brain region most activated by accepting fair offers ($P < 0.05$). This pattern of connectivity was again specific for lesion-induced criminal behavior compared with lesions causing other neurologic syndromes ($P < 0.0001$; *SI Appendix*, Fig. S3).

Replication in Criminals with Lesions of Uncertain Temporal Association.

In many cases, the temporal association between a brain lesion and criminal behavior is uncertain. We identified 23 cases of uncertain temporal association from our initial search (*SI Appendix*, Fig. S1 and Table S1). These cases included brain lesions found incidentally in incarcerated criminals (20 cases), brain lesions present since birth (two cases), or lesions being treated when the criminal behavior occurred (one case). As in our initial cohort, these lesions were spatially heterogeneous (*SI Appendix*, Fig. S4).

Lesion network mapping showed nearly identical results to those for our initial cohort (Fig. 7). All 23 lesions were functionally connected to the orbitofrontal cortex, and most were connected to the anterior temporal lobe, vmPFC, mesial temporal lobe/amygdala, and nucleus accumbens (Fig. 7B). Connectivity to these regions was specific compared with lesions causing other syndromes (Fig. 7B). This second cohort showed the same pattern of connectivity to regions involved in morality ($P < 0.0001$; Fig. 7B), subcomponents of morality ($P < 0.0001$ for value and theory of mind), harm aversion vs. utilitarian decisions in moral dilemmas ($P < 0.0001$), and fair vs. unfair offers in the ultimatum game ($P < 0.0001$). There was no significant difference in connectivity between lesions with and those without a documented temporal association with criminal behavior.

Discussion

There are four important findings in the present study. First, lesions temporally associated with criminal behavior occur in different brain regions but lie within a single connected brain network that is distinct from lesions not associated with criminal behavior. Second, lesions temporally associated with criminal behavior are functionally connected to regions activated by moral decision making, value-based decision making, and theory of mind tasks, but not with regions activated by empathy or cognitive control. Third, lesions temporally associated with criminal behavior show opposite connectivity to brain regions activated by competing moral choices, predicting the biases seen in these patients. Finally, all results are

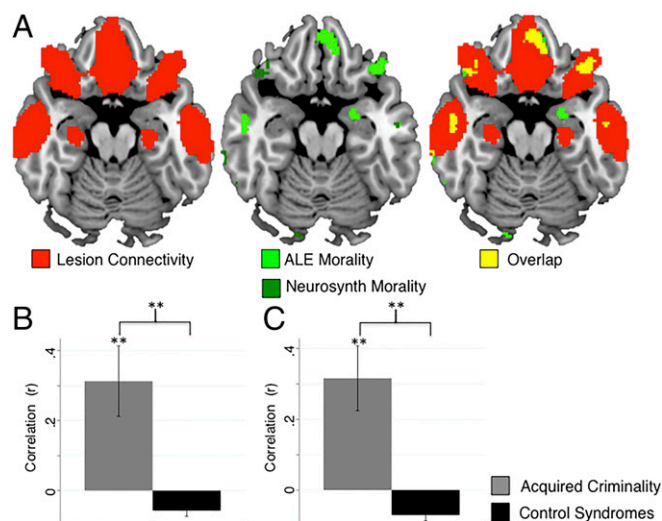


Fig. 4. Lesions temporally associated with criminality are functionally connected to morality regions. (A) Regions functionally connected to lesion locations temporally associated with criminal behavior (red) are similar to regions activated by moral decision-making tasks (green; overlap in yellow). (B and C) Lesion locations temporally associated with criminal behavior (gray) are functionally connected to regions activated by moral decision-making tasks identified in meta-analyses using Neurosynth (B) or ALE (C), unlike lesions causing other neurologic syndromes (control syndromes; black). ** $P < 0.0001$.

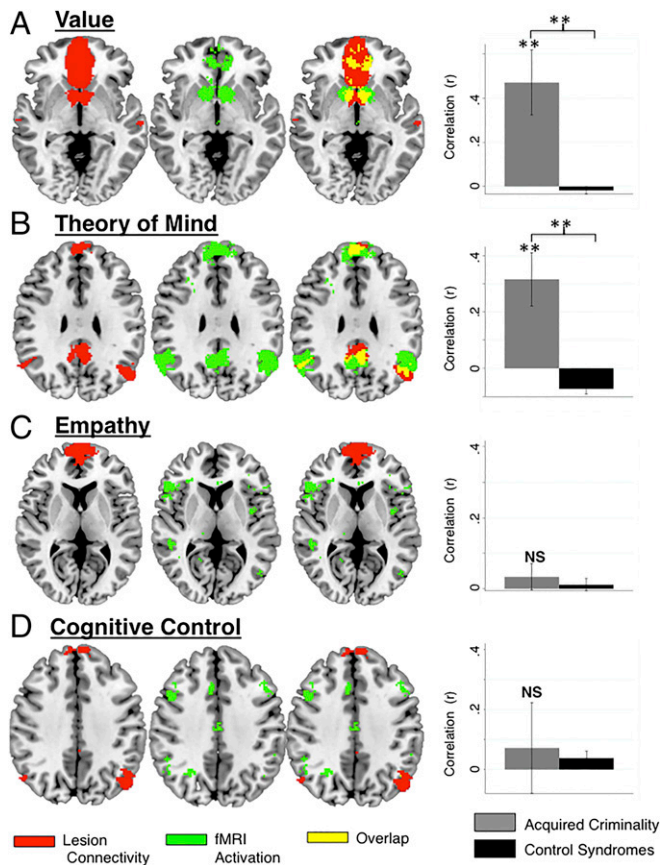


Fig. 5. Lesions temporally associated with criminality are functionally connected to some, but not all, subcomponents of moral decision making. Connectivity with lesion locations temporally associated with criminal behavior (red) is similar to that in regions activated by some but not other components of moral decision making (green; overlap in yellow). Quantitatively, lesion locations temporally associated with criminal behavior are functionally connected to regions activated by value-based decision making (A) and theory of mind tasks (B), but not with empathy (C) or cognitive control tasks (D). These results were specific to lesions temporally associated with criminal behavior (gray) compared with lesions causing other neurologic syndromes (control syndromes; black). $^{***}P < 0.0001$.

reproducible in a second cohort of patients with uncertain temporal associations between lesions and criminal behavior.

Lesions Associated with Criminal Behavior Are Part of a Single Connected Brain Network. The heterogeneity of brain lesions associated with criminal behavior has been noted in previous reviews (5, 15, 19), raising the question of what features these lesions share that could result in a similar behavioral phenotype. Our results show that these heterogeneous lesion locations are part of a single connected brain network that includes the orbitofrontal cortex, vmPFC, and anterior temporal lobes. Moreover, this network is specific to lesions associated with criminal behavior compared with lesions causing other syndromes.

Abnormalities in these same brain regions have been identified in other patient populations prone to criminal behavior (2–5). For example, up to 57% of patients with frontotemporal dementia have committed a crime and exhibit pathological changes in regions overlapping those identified in the present study (46–48). Similarly, psychopaths display atrophy distributed across the vmPFC, orbitofrontal cortex, and anterior temporal lobes (49, 50), as well as white matter and functional connectivity changes in these regions (51).

Lesion Connectivity to Regions Involved in Moral Decision Making. Lesions temporally associated with criminal behavior showed not only a similar pattern of brain connectivity, but also connectivity

specifically to regions implicated in moral decision making. These results link lesions resulting in criminal behavior to regions whose activity is correlated with moral decision making, value-based decision making, and theory of mind in normal subjects. Exactly which psychological processes are abnormal or disrupted in criminals remains a matter of ongoing research and debate (36). It is also unknown whether deficits in moral decision making contribute to lesion-associated criminal behavior. Nonetheless, our present results provide testable hypotheses for future work—namely, that patients whose lesions are connected to brain regions involved in specific aspects of moral decision making will demonstrate behavioral deficits in those tasks.

Connectivity to Brain Networks Involved in Competing Moral Decisions and Behavioral Bias. Previous work has demonstrated that the relative balance in activation between two different networks of brain regions is associated with different choices in moral dilemmas (33, 38). This has led to the hypothesis that there is “competition” between opponent networks during moral decision making (33), which is supported by evidence that these networks are normally anticorrelated at rest (37). We found that lesions temporally associated with criminal behavior were connected to both sets of brain regions, but in opposite directions. Thus, lesions may bias patients toward utilitarian moral decisions (i.e., push someone off the bridge to save others) (39) and rejection of unfair offers in the ultimatum game (42) (*SI Appendix, Fig. S5*).

Replication and Causal Inferences. Our main analysis included 17 cases with a strong temporal relationship between the lesion and criminal behavior. Our results were nearly identical using lesions of uncertain temporal association with criminal behavior. This replication increases our confidence that lesions within a well-defined network increase the relative probability of criminality. In a judicial setting, it is often desirable to know whether a lesion contributed to criminality or was an incidental finding. In this respect, the implications of our findings are unclear. The published cases included in our second cohort were often believed to be contributing to criminality (and not incidental), even if a clear temporal association between the lesion and criminality was not documented. Whether lesion network mapping can help differentiate incidental lesions from

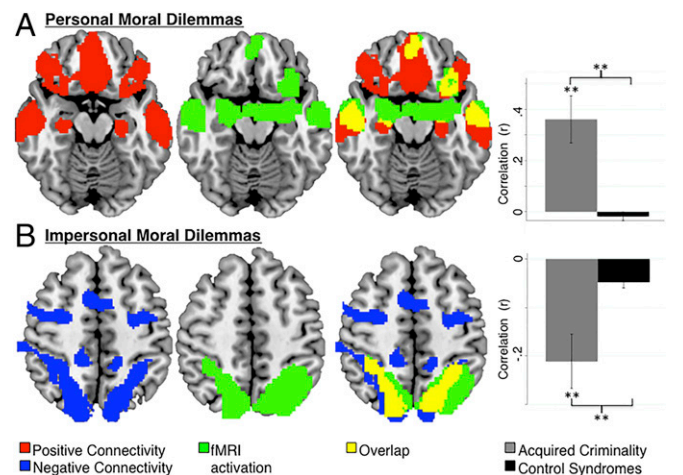


Fig. 6. Lesions temporally associated with criminality are connected in opposite directions to opponent networks involved in moral decision making. (A) Lesion locations are positively connected (red) with regions activated by decisions to avoid harm in personal moral dilemmas (green; overlap in yellow). (B) Lesion locations are negatively connected (blue) with regions activated by utilitarian decisions in impersonal moral dilemmas (green; overlap in yellow). Results are specific to lesions temporally associated with criminal behavior (gray) compared with lesions causing other neurologic syndromes (black). $^{***}P < 0.0001$.

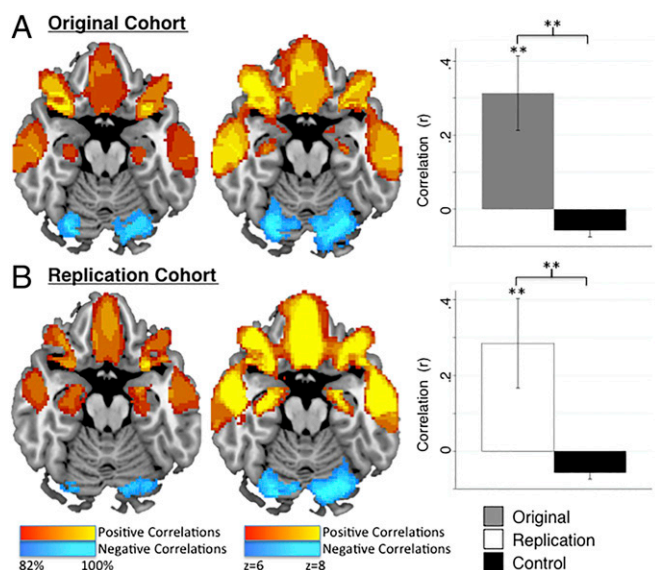


Fig. 7. Replication in lesions of uncertain temporal association with criminal behavior. Lesions with a clear temporal association (original cohort) (*A*) and uncertain temporal association (replication cohort) (*B*) are connected to the same regions (lesion network overlap, *Right*; voxel-wise nonparametric comparison with control syndromes, *Middle*). Lesions in both cohorts were significantly connected to regions involved in morality compared with control lesions (*Right*). *** $P < 0.0001$.

lesions contributing to a given behavior remains unknown, but is beginning to be tested in this context (29).

It is important not to overinterpret our results with respect to the prediction of criminal behavior. Factors including genetics (13), age at lesion onset (10, 12, 14), lesion etiology (11), environment, social support, and premorbid personality traits (2) may contribute to criminal behavior, either independently or through interaction with the lesion location. Violence or crime occurs in only ~9% of patients with traumatic brain injury (52, 53), 14% of patients with frontal lobe injury (54), and up to 57% of patients with frontal temporal dementia (46–48). Specifically, criminal behavior has been reported in 37–57% of patients with behavioral variant frontotemporal dementia and in 27–56% of those with semantic variant primary progressive aphasia, although these percentages include nonviolent crimes, such as shoplifting and traffic violations (46–48). These findings suggest that many patients with lesions lying within our network will not develop criminal behavior. Thus, lesions within our identified network may increase the risk of criminal behavior, but should not be interpreted as an inevitable or sole cause of criminal behavior. Prospective studies addressing the likelihood that a lesion with connectivity to our identified regions will result in criminal behavior, as well as other contributing factors, are needed.

Limitations. Several potential confounders associated with lesion network mapping have been identified and addressed in previous studies (21, 23–25, 28, 29). First, the accuracy of lesion tracing is limited by the quality of the published images, as well as the use of a 2D lesion to approximate a 3D lesion. While we did not have any 3D lesions available for the present study, previous studies have shown that lesion network mapping of 2D approximations of 3D lesions are highly similar (21, 23). Second, results might vary based on the specific connectome dataset or resting-state functional magnetic resonance imaging (fMRI) processing strategy used. We have previously demonstrated that results do not change when using an age-matched or disease-specific connectome, or using alternative processing strategies (21). Of note, the present study used a larger ($n = 1,000$) publicly available connectome (55), rather than the smaller ($n = 98$) dataset used in our previous

lesion network mapping studies (21, 23–25). The results were nearly identical across these two datasets (*SI Appendix, Fig. S6*).

A limitation of the present study is that no functional imaging or behavioral testing was actually performed in patients with lesion-induced criminal behavior. Rather, we used previous behavioral testing, previously published lesion locations, neuroimaging meta-analyses, and large connectome datasets to generate our findings. This approach avoids confounders often associated with fMRI in patients, including movement, task compliance, low subject numbers (especially in rare syndromes such as acquired criminal behavior), and ambiguity regarding whether imaging abnormalities are caused by the lesion, compensating for the lesion, or secondary to altered task performance (56). However, without functional imaging in patients, it is unknown whether remote brain regions functionally connected to a lesion location (and thus predicted to be dysfunctional in the present study) are actually dysfunctional in these patients. Similarly, while we show that brain lesions associated with criminal behavior are functionally connected to regions involved in moral decision making, it remains unknown whether moral decision making is actually abnormal in these patients.

A second limitation is that the study focused on retrospective cases of lesion-induced criminal behavior. Thus, our sample is susceptible to bias due to our search and inclusion criteria, and might not represent the true distribution of patients with lesion-induced criminality. In addition, our determination of criminal behavior, and the temporal relationship between the lesion and change in behavior, are limited by the accuracy and variable assessments in each case. As discussed previously, prospective studies are needed to determine the likelihood that a lesion with connectivity to the regions that we identified will develop criminal behavior, as well as whether patients with lesion-induced criminal behavior have abnormalities in moral decision making.

Conclusion. Brain lesions temporally associated with criminal behavior are characterized by a unique pattern of brain connectivity. These results may prove useful for ongoing efforts to understand, predict, and assign responsibility to criminal acts.

Materials and Methods

This study was approved by the Massachusetts General Hospital Institutional Review Board. Informed consent was not required.

Patient Cases from the Literature. We identified patients via a PubMed search. Inclusion criteria included case description of criminal behavior in a patient, an intrinsic focal brain lesion, and a published image of the brain lesion of sufficiently high quality to trace the lesion's location onto a standardized brain atlas. Among the 40 cases meeting these criteria, 17 had a clear temporal relationship between the lesion and criminal behavior and were included in the primary analyses, while the remaining 23 cases had an uncertain temporal association between lesion onset and criminal behavior and were used in a replication cohort (*SI Appendix, Fig. S1 and Table S1*). Lesions were traced by hand onto a standardized brain atlas ($2 \times 2 \times 2$ MNI space) using FSL as in previous work (21, 23–25) (*Fig. 1 and SI Appendix, Fig. S4*).

Lesion Network Mapping. Traced lesions were used as individual seeds in a resting-state connectivity analysis using data obtained from 1,000 healthy subjects (55). Functional connectivity to each lesion was determined by calculating the correlated time course between each lesion location and every other brain voxel using the resting-state data from each individual normal control (28, 29). These correlations for all 1,000 subjects were then combined to calculate a T-score value for each individual voxel. Positive and negative correlations were thresholded at $T > \pm 12$ to create a binarized map of functionally connected regions to each patient's lesion site ($P < 10^{-17}$ uncorrected). Finally, maps from all patients were combined to form a group map showing the number of patients with lesions functionally connected with each individual voxel (*Fig. 2C*).

Comparison with Lesions Causing Other Neurologic Syndromes. Lesion network mapping results from lesions associated with criminal behavior were compared with lesion network mapping results from 63 lesions causing other syndromes (21). Voxel-wise nonparametric testing was used to identify voxels significantly more likely to be connected with lesions causing acquired criminal behavior

compared with control lesions using a standard Bonferroni-corrected family-wise error rate of $P < 0.05$ (Fig. 3).

Neuroimaging Meta-Analyses of Morality. To define regions involved in morality, we performed meta-analyses of task-fMRI activation studies using the terms "morality," "value," "theory of mind," "empathy," and "cognitive control" in Neurosynth (www.neurosynth.org), a data-driven, automated engine for performing neuroimaging meta-analyses (32). To ensure that results were not dependent on our choice of meta-analysis technique, we used the results of a recent, manually performed meta-analysis of fMRI studies of morality (31) using ALE (30).

Connectivity to Regions Involved in Morality. fMRI time courses were extracted from each lesion location and the binarized map of significant regions from each meta-analysis. The Pearson correlation coefficient between time courses was computed for each subject in our normative 1,000-subject dataset. Resulting r values were converted to a normal distribution using the Fisher r -to- z transform and compared statistically using a two-tailed t test. Lesion locations causing acquired criminal behavior were compared with lesion locations causing other neurologic syndromes identified in our previous

work (21). All statistics were computed using Stata version 14.0 (StataCorp, College Station, TX).

Connectivity to Regions Involved in Moral Dilemmas. We obtained activation maps for personal moral dilemmas and impersonal moral dilemmas from Greene et al. (33). For the ultimatum game, we created 8-mm-radius regions of interest at the peak reported coordinates from a recent meta-analysis for accepting fair offers and rejecting unfair offers (45). Time course correlations and statistical comparisons were performed as above. We tested whether the combination of connectivity to regions activated by personal and impersonal dilemmas was superior to connectivity to either set of regions alone using a likelihood ratio test.

ACKNOWLEDGMENTS. We thank Joshua Greene and Danilo Bzdok for sharing data. This work was supported by funding from the Sidney R. Baer, Jr. Foundation (R.R.D. and M.D.F.), the National Institutes of Health Grant K23NS083741 (to M.D.F.), the Dystonia Foundation (M.D.F.), the Nancy Lurie Marks Foundation (M.D.F.), the Alzheimer's Association (R.R.D.), and the BrightFocus Foundation (R.R.D.).

1. Anderson DA (1999) The aggregate burden of crime. *J Law Econ* 42:611–642.
2. Glenn AL, Raine A (2014) Neurocriminology: Implications for the punishment, prediction and prevention of criminal behaviour. *Nat Rev Neurosci* 15:54–63.
3. Blake PY, Pincus JH, Buckner C (1995) Neurologic abnormalities in murderers. *Neurology* 45:1641–1647.
4. Anderson NE, Kiehl KA (2012) The psychopath magnetized: Insights from brain imaging. *Trends Cogn Sci* 16:52–60.
5. Bufkin JL, Luttrell VR (2005) Neuroimaging studies of aggressive and violent behavior: Current findings and implications for criminology and criminal justice. *Trauma Violence Abuse* 6:176–191.
6. Adolphs R (2016) Human lesion studies in the 21st century. *Neuron* 90:1151–1153.
7. Denno DW (2015) The myth of the double-edged sword: An empirical study of neuroscience evidence in criminal cases. *Boston Coll Law Rev* 56:493–552.
8. Saver JL, Damasio AR (1991) Preserved access and processing of social knowledge in a patient with acquired sociopathy due to ventromedial frontal damage. *Neuropsychologia* 29:1241–1249.
9. Blumer D, Benson DF (1975) Personality changes with frontal lobe lesions. *Psychiatric Aspects of Neurological Disease*, eds Benson DF, Blumer D (Grune & Stratton, New York), pp 151–170.
10. Anderson SW, Bechara A, Damasio H, Tranel D, Damasio AR (1999) Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Nat Neurosci* 2:1032–1037.
11. Anderson SW, Damasio H, Tranel D (1990) Neuropsychological impairments associated with lesions caused by tumor or stroke. *Arch Neurol* 47:397–405.
12. Taber-Thomas BC, et al. (2014) Arrested development: Early prefrontal lesions impair the maturation of moral judgement. *Brain* 137:1254–1261.
13. Pardini M, et al. (2011) Prefrontal cortex lesions and MAO-A modulate aggression in penetrating traumatic brain injury. *Neurology* 76:1038–1045.
14. Anderson SW, Barrash J, Bechara A, Tranel D (2006) Impairments of emotion and real-world complex behavior following childhood- or adult-onset damage to ventromedial prefrontal cortex. *J Int Neuropsychol Soc* 12:224–235.
15. Batts S (2009) Brain lesions and their implications in criminal responsibility. *Behav Sci Law* 27:261–272.
16. Damasio H, Grabowski T, Frank R, Galaburda AM, Damasio AR (1994) The return of Phineas Gage: Clues about the brain from the skull of a famous patient. *Science* 264:1102–1105.
17. Mendez MF (2009) The neurobiology of moral behavior: Review and neuropsychiatric implications. *CNS Spectr* 14:608–620.
18. Braun CMJ, Léveillé C, Guimond A (2008) An orbitofrontostriatopallidal pathway for morality: Evidence from postlesion antisocial and obsessive-compulsive disorder. *Cogn Neuropsychiatry* 13:296–337.
19. Fumagalli M, Priori A (2012) Functional and clinical neuroanatomy of morality. *Brain* 135:2006–2021.
20. Monakow C (1914) *Die Lokalisation im Grosshirn: und der Abbau der Funktion durch kortikale Herde* (von J.F. Bergmann, Wiesbaden, Germany).
21. Boes AD, et al. (2015) Network localization of neurological symptoms from focal brain lesions. *Brain* 138:3061–3075.
22. Sutterer MJ, et al. (2016) Canceled connections: Lesion-derived network mapping helps explain differences in performance on a complex decision-making task. *Cortex* 78:31–43.
23. Darby RR, Laganieri S, Pascual-Leone A, Prasad S, Fox MD (2017) Finding the imposter: Brain connectivity of lesions causing delusional misidentifications. *Brain* 140:497–507.
24. Fischer DB, et al. (2016) A human brain network derived from coma-causing brainstem lesions. *Neurology* 87:2427–2434.
25. Laganieri S, Boes AD, Fox MD (2016) Network localization of hemichorea-hemiballismus. *Neurology* 86:2187–2195.
26. Buckholtz JW (2015) Social norms, self-control, and the value of antisocial behavior. *Curr Opin Behav Sci* 3:122–129.
27. Baron-Cohen S (2013) *The Science of Evil: On Empathy and the Origins of Cruelty* (Basic Books, New York).
28. Horn A, et al. (2017) Connectivity predicts deep brain stimulation outcome in Parkinson disease. *Ann Neurol* 82:67–78.
29. Darby RR, Fox MD (2017) Reply. Capgras syndrome: Neuroanatomical assessment of brain MRI findings in an adolescent patient. *Brain* 140:e44.
30. Eickhoff SB, et al. (2009) Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: A random-effects approach based on empirical estimates of spatial uncertainty. *Hum Brain Mapp* 30:2907–2926.
31. Bzdok D, et al. (2012) Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. *Brain Struct Funct* 217:783–796.
32. Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD (2011) Large-scale automated synthesis of human functional neuroimaging data. *Nat Methods* 8:665–670.
33. Greene JD, Nystrom LE, Engell AD, Darley JM, Cohen JD (2004) The neural bases of cognitive conflict and control in moral judgment. *Neuron* 44:389–400.
34. Cushman F (2013) Action, outcome, and value: A dual-system framework for morality. *Pers Soc Psychol Rev* 17:273–292.
35. Young L, Cushman F, Hauser M, Saxe R (2007) The neural basis of the interaction between theory of mind and moral judgment. *Proc Natl Acad Sci USA* 104:8235–8240.
36. Darby RR, Edersheim J, Price BH (2016) What patients with behavioral-variant frontotemporal dementia can teach us about moral responsibility. *AJOB Neurosci* 7:193–201.
37. Fox MD, et al. (2005) The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc Natl Acad Sci USA* 102:9673–9678.
38. Greene JD, Sommerville RB, Nystrom LE, Darley JM, Cohen JD (2001) An fMRI investigation of emotional engagement in moral judgment. *Science* 293:2105–2108.
39. Koenigs M, et al. (2007) Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature* 446:908–911.
40. Koenigs M, Kruepke M, Zeier J, Newman JP (2012) Utilitarian moral judgment in psychopathy. *Soc Cogn Affect Neurosci* 7:708–714.
41. Mendez MF, Anderson E, Shapira JS (2005) An investigation of moral judgement in frontotemporal dementia. *Cogn Behav Neurol* 18:193–197.
42. Koenigs M, Tranel D (2007) Irrational economic decision-making after ventromedial prefrontal damage: Evidence from the ultimatum game. *J Neurosci* 27:951–956.
43. Koenigs M, Kruepke M, Newman JP (2010) Economic decision-making in psychopathy: A comparison with ventromedial prefrontal lesion patients. *Neuropsychologia* 48:2198–2204.
44. O'Callaghan C, et al. (2016) Fair play: Social norm compliance failures in behavioural variant frontotemporal dementia. *Brain* 139:204–216.
45. Feng C, Luo Y-J, Krueger F (2015) Neural signatures of fairness-related normative decision making in the ultimatum game: A coordinate-based meta-analysis. *Hum Brain Mapp* 36:591–602.
46. Mendez MF, Chen AK, Shapira JS, Miller BL (2005) Acquired sociopathy and frontotemporal dementia. *Dement Geriatr Cogn Disord* 20:99–104.
47. Liljehgren M, et al. (2015) Criminal behavior in frontotemporal dementia and Alzheimer disease. *JAMA Neurol* 72:295–300.
48. Diehl-Schmid J, Pernecky R, Koch J, Nedopil N, Kurz A (2013) Guilty by suspicion? Criminal behavior in frontotemporal lobar degeneration. *Cogn Behav Neurol* 26:73–77.
49. Ermer E, Cope LM, Nyalakanti PK, Calhoun VD, Kiehl KA (2012) Aberrant paralimbic gray matter in criminal psychopathy. *J Abnorm Psychol* 121:649–658.
50. de Oliveira-Souza R, et al. (2008) Psychopathy as a disorder of the moral brain: Frontotemporo-limbic grey matter reductions demonstrated by voxel-based morphometry. *Neuroimage* 40:1202–1213.
51. Motzkin JC, Newman JP, Kiehl KA, Koenigs M (2011) Reduced prefrontal connectivity in psychopathy. *J Neurosci* 31:17348–17357.
52. Fazel S, Lichtenstein P, Grann M, Långström N (2011) Risk of violent crime in individuals with epilepsy and traumatic brain injury: A 35-year Swedish population study. *PLoS Med* 8:e1001150.
53. Timonen M, et al. (2002) The association of preceding traumatic brain injury with mental disorders, alcoholism and criminality: The northern Finland 1966 birth cohort study. *Psychiatry Res* 113:217–226.
54. Grafman J, et al. (1996) Frontal lobe injuries, violence, and aggression: A report of the Vietnam head injury study. *Neurology* 46:1231–1238.
55. Yeo BT, et al. (2011) The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol* 106:1125–1165.
56. Grefkes C, Fink GR (2014) Connectivity-based approaches in stroke and recovery of function. *Lancet Neurol* 13:206–216.