Amygdala lesions do not compromise the cortical network for false-belief reasoning

Robert P. Spunt[a,1], Jed T. Elison[a,b], Nicholas Dufour[c], René Hurlemann[d], Rebecca Saxe[c], and Ralph Adolphs[a]

[a]Division of Humanities and Social Sciences, California Institute of Technology, Pasadena, CA 91125; [b]Institute of Child Development, University of Minnesota, Minneapolis, MN 55455; [c]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139; and [d]Department of Psychiatry, University of Bonn, 53113 Bonn, Germany

Edited by Joseph E. LeDoux, New York University, New York, NY, and approved February 27, 2015 (received for review December 2, 2014)

The amygdala plays an integral role in human social cognition and behavior, with clear links to emotion recognition, trust judgments, anthropomorphization, and psychiatric disorders ranging from social phobia to autism. A central feature of human social cognition is a theory-of-mind (ToM) that enables the representation other people's mental states as distinct from one's own. Numerous neuroimaging studies of the best studied use of ToM—false-belief reasoning—suggest that it relies on a specific cortical network; moreover, the amygdala is structurally and functionally connected with many components of this cortical network. It remains unknown whether the cortical implementation of any form of ToM depends on amygdala function. Here we investigated this question directly by conducting functional MRI on two patients with rare bilateral amygdala lesions while they performed a neuroimaging protocol standardized for measuring cortical activity associated with false-belief reasoning. We compared patient responses with those of two healthy comparison groups that included 480 adults. Based on both univariate and multivariate comparisons, neither patient showed any evidence of atypical cortical activity or any evidence of atypical behavioral performance; moreover, this pattern of typical cortical and behavioral response was replicated for both patients in a follow-up session. These findings argue that the amygdala is not necessary for the cortical implementation of ToM in adulthood and suggest a reevaluation of the role of the amygdala and its cortical interactions in human social cognition.

theory-of-mind | amygdala | lesions | false-belief | fMRI

The amygdala is considered a critical node of the "social brain" that contributes to myriad social behaviors exhibited by primates (1–4). Neurons in both the monkey (5) and human amygdala (6) respond prominently to faces, and lesions of the monkey amygdala result in complex impairments in social behavior (7, 8). Rare bilateral lesions of the amygdala in human patients impair the ability to infer emotions from facial expressions (9, 10), to make more complex social judgments from faces (11), and to guide appropriate social behaviors (12).

A core social ability of humans that emerges early in childhood has been long studied under the name of "theory-of-mind" (ToM), an ability to impute mental states to other people. Amygdala lesions can impair the ability to impute such mental states spontaneously to animated geometric shapes (13, 14) as well as other complex expressions of ToM (15). These impairments in social cognition following amygdala lesions also have been compared with the intensively studied impairments in mental-state understanding observed in autism spectrum disorder (16, 17). Indeed, the amygdala has been implicated in emotional and social dysfunction in a number of psychiatric disorders (18).

Neuroimaging studies of ToM-related abilities, on the other hand, have focused largely on cortical networks (19, 20). One of these networks, based on using a localizer requiring subjects to infer false beliefs from written stories (the "False-Belief Localizer") (21, 22) has become so well established that it is commonly referred to as the "ToM network" and prominently includes the temporoparietal junction as well as medial frontoparietal and anterior temporal cortices (23–28).

If the amygdala plays a critical role in social cognition, why is it not regularly identified in neuroimaging studies of ToM? One answer may be that these studies have been focused more on cortical networks, and possible amygdala activations are either underreported or underdiscussed. A second answer may be that the blood oxygenation level-dependent (BOLD) response is more difficult to evoke in the amygdala than in cortex (29, 30). However, the amygdala's vast connectivity with most of the neocortex (31), prominently including some of the key nodes of the false-belief network such as the medial prefrontal cortex (32, 33), together with its role in social cognition reviewed above, justifies a strong hypothesis. That hypothesis is that the cortical false-belief network should include or be modulated by the amygdala. The clear prediction from this hypothesis is that lesions of the amygdala should alter the functional response of cortical regions critical to ToM.

To test this prediction in the most direct way, we used functional MRI (fMRI) in two rare patients with bilateral amygdala lesions and closely interrogated BOLD responses within the amygdala in a large group of neurologically healthy controls. The patients with amygdala lesions had developmental-onset calcifications of the amygdala resulting from Urbach–Wiethe disease (34) (raising interesting further questions about the possible developmental contributions of the amygdala to the false-belief reasoning network, issues we take up in *Discussion*). To evoke false-belief network activation, each patient performed the well-established False-Belief Localizer twice in separate MRI sessions. The False-Belief Localizer (often called simply the "ToM

Significance

Humans use a so-called "theory-of-mind" to reason about the beliefs of others. Neuroimaging studies of belief reasoning suggest it activates a specific cortical network. The amygdala is interconnected with this network and plays a fundamental role in social behavior. For the first time, to our knowledge, we test whether amygdala lesions compromise the cortical implementation of theory-of-mind. Two patients with bilateral amygdala lesions performed a belief reasoning test while undergoing functional MRI. Remarkably, both patients showed typical test performance and cortical activity when compared with nearly 500 healthy controls. This result shows that the amygdala is not a necessary part of theory-of-mind function in adulthood and forces a reevaluation of the amygdala's role in social cognition.

Author contributions: R.P.S., J.T.E., and R.A. designed research; R.P.S., J.T.E., and R.H. performed research; R.H. and R.S. contributed new reagents/analytic tools; R.P.S. and N.D. analyzed data; and R.P.S., J.T.E., and R.A. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. Email: spunt@caltech.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1422679112/-/DCSupplemental.

Localizer") developed by Rebecca Saxe and colleagues (21, 22) uses brief verbal narratives to manipulate the demand to represent another person's false belief about reality.

At the outset, we clarify that the False-Belief Localizer does not exhaustively represent the range and complexity of the human capacity to reason about mental states (35). In fact, many different behavioral tasks have been used to manipulate mental-state reasoning in previous neuroimaging studies (23, 26), and recent evidence has demonstrated convincingly that these various tasks are not interchangeable manipulations of a single ToM capacity but rather modulate dissociable cortical networks (28, 36). Nonetheless, several reasons justify our decision to focus here on the False-Belief Localizer. First, given that false-belief representation historically has been considered the most unequivocal expression of ToM (37), theory and research on ToM has long maintained a central focus on the capacity to represent false beliefs (38, 39). Second, the focus of ToM research on false-belief reasoning has remained strong in neuroimaging studies of social cognition, in large part because of the efforts of Saxe and colleagues (21, 22) to optimize and make publicly available an efficient protocol for this purpose. Because this same basic protocol has been used in numerous neuroimaging studies of neurologically healthy adults, it is now possible to generate large empirical distributions against which new data points can be compared (40). Therefore, the present study tests the hypothesis that cortical function during false-belief reasoning would show abnormalities in the absence of the amygdala, using this same false-belief neuroimaging task.

Results

Patient Behavioral Performance. We compared the performance in the patient group's first session with the bootstrapped California Institute of Technology (Caltech) control group distribution of performance in both Belief and Photo trials. The results of this comparison are represented in Fig. S1. When examining the percentage of correct responses, we observed no evidence for atypical performance on false-belief trials (patient = 75.33%; healthy control = 75.99%; $P = 0.940$) or false-photo trials (patient = 65.00%; healthy control = 81.05%; $P = 0.229$). Similarly, we observed no evidence for atypical response times on false-belief trials (patient = 16.22 s; healthy control = 15.38 s; $P = 0.694$) or on false-photo trials (patient = 15.71 s; healthy control = 14.33 s; $P = 0.541$). Finally, both patients showed no evidence for atypical performance in their second session of performing the task (Fig. S1).

Amygdala Responses to False-Belief Reasoning in the Reference Groups. We first describe the proportion of voxels available for analysis in the amygdala regions of interest (ROIs) in the large MIT reference group ($n = 462$ subjects). Usable voxels were defined as those with a value exceeding 12.5% of the mean global signal and for every time point in the time-series [this corresponds to the default criterion for voxel inclusion in analyses conducted using the software Statistical Parametric Mapping (SPM8)]. On average, the percentage of valid voxels present in each ROI for a given participant was high in both hemispheres but was highly variable, in part because of variable signal dropout from well-known susceptibility artifacts in this region of the brain (left: mean = 90.20%, SD = 14.97%; right: mean = 94.71%, SD = 11.49%). We took this approach to prevent SPM's standard group analysis from masking out brain regions where even a single subject might have no useable voxels. In the anatomical amygdala ROIs, a one-sample t test on usable voxels demonstrated activation to the Belief > Photo contrast of parameter estimates in both the left [$t(459) = 5.035$, $P < 0.000001$, 95% CI_{boot} (0.109, 0.247)] and right [$t(459) = 3.325$, $P < 0.001$, 95% CI_{boot} (0.043, 0.167)] amygdala. Corroborating this ROI analysis, a voxelwise whole-brain analysis including voxels with data in at least 100 subjects also revealed a response to the Belief > Photo contrast in both the left (voxel extent = 71; peak: $x = -20$, $y = -6$, z = −14, $t = 6.419$) and right (voxel extent = 39; peak: $x = 22$, $y = -2$, z = −16, $t = 6.331$) amygdala (Fig. 1*C*).

We then used the estimated amygdala response in the MIT reference group to calculate the statistical power for observing an effect in each ROI in an independently conducted study. This analysis suggested that to achieve a detection power of 80%, a study would need to acquire 270 subjects for the left and 470 subjects for the right amygdala. At the typical sample size of 20 used in neuroimaging studies to date, detection power for the left and right amygdala was estimated to be 16.10% and 12.52%, respectively. Unsurprisingly, therefore, we did not observe reliable contrast in either ROI in the Caltech reference group ($n = 18$; Ps > 0.50). However, we did find that individual differences in amygdala activation in the Belief > Photo contrast were significantly associated with activation in several cortical regions of the false-belief network, namely, the superior temporal sulcus and temporoparietal junction bilaterally and the precuneus (Table 1). Although not statistically reliable when taken individually, the correlations of amygdala activation with the remaining cortical ROIs were all positive (minimum $r = 0.32$). Taken together, these findings support the idea that the amygdala contributes to the functioning of the false-belief network, even though its activation is not generally reported.

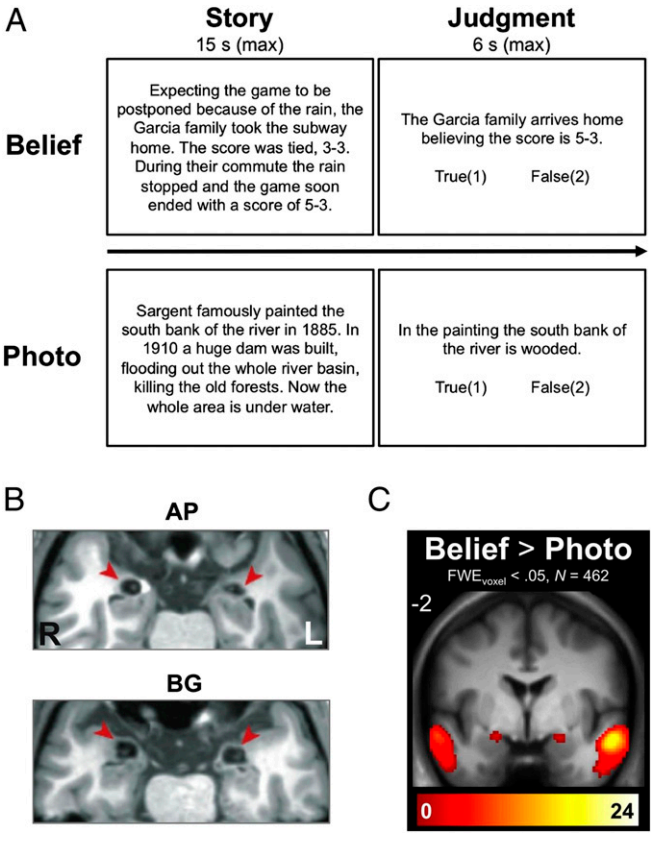

Fig. 1. Study design and rationale. (*A*) Schematic showing the design of the False-Belief Localizer task. The rows show the Story and Judgment screens for an actual trial in the False-Belief and False-Photo conditions. (*B*) Structural MRIs showing each patient's amygdala lesions. Displayed are 1-mm isotropic T1-weighted MRI transverse sections of the patients' anterior medial temporal lobes. Red arrows highlight focal calcification damage in the amygdalas of patients AP and BG. (*C*) Evidence that the Belief > Photo contrast activates bilateral amygdala in the typically developing brain. L, left; R, right.

Table 1. Correlation of individual differences in the Caltech reference group between activation to the Belief > Photo contrast in amygdala and cortical ROIs and percent correct during performance of the False-Belief Localizer task

	Amygdala (AAL)		Percent correct	
Region name	Left	Right	Belief	Photo
L Amygdala (AAL)	—	0.77**	0.32	0.48*
R Amygdala (AAL)	0.77**	—	0.06	0.08
L TPJ	0.48*	0.42	0.62**	0.23
R TPJ	0.55*	0.50*	0.40	0.18
Precuneus	0.60**	0.50*	0.41	0.23
DMPFC	0.35	0.32	0.40	-0.02
MMPFC	0.43	0.43	0.32	-0.00
VMPFC	0.33	0.41	0.25	-0.03
R STS	0.72***	0.76***	0.35	0.08

Amygdala ROIs are from the Automated Anatomical Labeling atlas (AAL). DM, dorsomedial; L, left; MM, mid-medial; PC, precuneus; PFC, prefrontal cortex; R, right; STS, superior temporal sulcus; TPJ, temporoparietal junction; VM, ventromedial. Probability values (uncorrected): *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

Cortical Responses to False-Belief Reasoning in the Patient and Reference Groups.

Whole-brain responses. Fig. 2 displays whole-brain renderings of the thresholded Belief > Photo contrast estimated for the two reference groups, in patient AP, and in patient BG. Table S1 lists the cortical regions surviving correction in each whole-brain analysis. In terms of gross visual comparison, both patients show largely typical cortical responses to false-belief reasoning. The analyses that follow aim to determine if the patient cortical response shows any sign of abnormality.

Comparison with Caltech reference group. We first compared the patient responses with those of the Caltech reference group ($n = 18$), whose data were collected using the same scanner and task used with the patients (although the task was translated into German for patient BG). Given the relatively small size of the Caltech reference group, we used a bootstrapping procedure to create a distribution of the average response for every possible combination of two individuals. This procedure yielded a bootstrapped population estimate based on 153 groups of two, which we used as a reference to evaluate the typicality of the average response on every outcome observed in the two patients.

Using the MIT group-level unthreshholded and gray matter-masked Belief > Photo contrast map as a benchmark ($n = 462$), we first determined if the overall spatial response pattern observed in the Caltech group was more typical than that in the patient group. The result of this comparison is shown in Fig. 3. Compared with the average correlation of the bootstrapped Caltech distribution ($r_{mean} = 0.50$), the patients showed no evidence of atypical response patterns in session 1 ($r_{mean} = 0.50$; $P_{typical} = 0.985$), and this typical response pattern was reproduced in the data collected during the patients' second session ($r_{mean} = 0.54$; $P_{typical} = 0.506$).

We next examined the pattern of response in a mask containing all a priori functional ROIs that were defined on the basis of the Belief > Photo contrast in the MIT reference group (Fig. S2). As before, we used the spatial pattern observed in the MIT reference group as a benchmark. Compared with the average correlation of the bootstrapped Caltech distribution ($r_{mean} = 0.49$), the patients again showed no evidence of atypical response patterns in session 1 ($r_{mean} = 0.48$; $P_{typical} = 0.971$), and once again this typical response pattern was reproduced in session 2 ($r_{mean} = 0.54$; $P_{typical} = 0.425$).

Finally, we examined the magnitude (mean and peak) and peak location (x-, y-, and z-coordinates) of the patient response in each of the seven functional ROIs. Response magnitude results are shown in Table 2. Mirroring the response pattern analyses reported above, the patients did not demonstrate a response that was reliably atypical across the two sessions. In fact, fewer than 3% of the comparisons performed within each session showed evidence of an abnormality, reflecting a false-positive rate that would be expected by chance alone.

Comparison with the MIT reference group. We capitalized on the large MIT reference group to perform a comparison focused on the individual patient response data. We compared the whole-brain spatial pattern of the Belief > Photo contrast for each patient with that of each individual in the MIT reference group ($n = 462$). To create a leave-one-out reference distribution, we took each individual in the MIT reference group and computed the mean correlation of their whole-brain response with the remaining members of the MIT reference group. This procedure yielded a distribution of 462 correlation values (mean = 0.14, SD = 0.07) that we used to test the null hypothesis that each patient's correlation with the MIT Reference group was abnormal.

For patient AP, we observed no evidence for an atypical response pattern when examining the whole-brain contrast from both session 1 ($r_{mean} = 0.21$; $P_{typical} = 0.306$) and session 2 ($r_{mean} = 0.22$; $P_{typical} = 0.256$). For patient BG, we similarly failed to observe any evidence for atypical responses in both session 1 ($r_{mean} = 0.22$; $P_{typical} = 0.237$) and session 2 ($r_{mean} = 0.26$; $P_{typical} = 0.091$). For both patients and across both sessions, we also observed no evidence for atypical response patterns when restricting the space to the functionally defined false-belief network (all Ps > 0.140).

Discussion

We used fMRI to examine cortical function during false-belief reasoning in two patients with rare bilateral amygdala lesions. When comparing the patients with two neurologically healthy reference groups, we found remarkably clear evidence for typical behavioral performance and cortical responses in the patient group. Moreover, this finding was replicated in a second session. These results indicate that the amygdala is not necessary for either the behavioral or neural expression of ToM. However, this

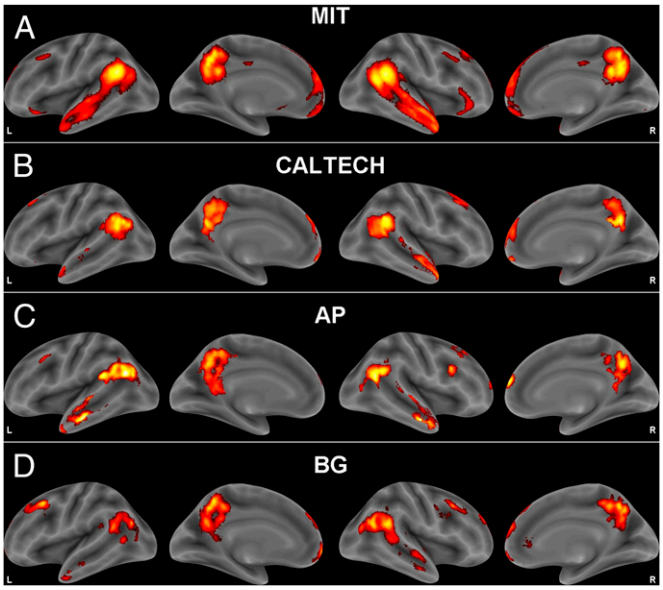

Fig. 2. Whole-brain renderings of the Belief > Photo contrast in the MIT reference group ($n = 462$; corrected at a voxel-level familywise error of 0.05) (*A*), the Caltech reference group ($n = 18$; corrected at a cluster-level familywise error of 0.05) (*B*), and the amygdala-lesion patients AP (*C*) and BG (*D*) (both estimated using combined data from their two independent sessions and corrected at a cluster-level familywise error of 0.05). L, left; R, right.

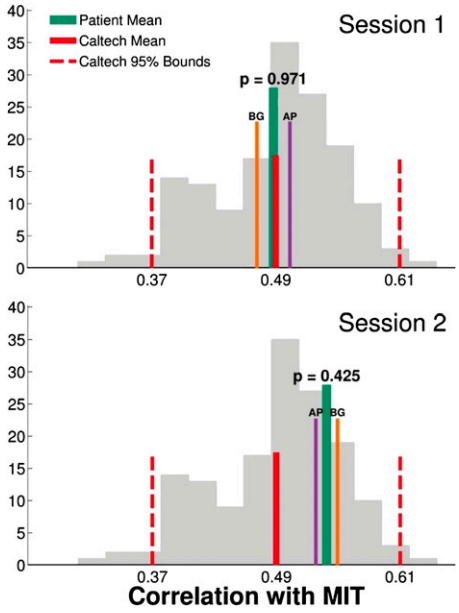

Fig. 3. Comparing global contrast typicality in the patient and Caltech reference groups (using the MIT group's unthreshholded Belief > Photo contrast map as a benchmark). The bootstrapped distribution of mean correlation in the Caltech reference group is shown in light gray, and the individual patient observations are shown in distinct colors with the patient ID indicated above the bars.

conclusion is restricted to the specific task and amygdala lesions we tested: explicit online false-belief reasoning and amygdala lesions of primarily the basolateral amygdala tested in adults. We take up these qualifications further below.

Implications. Our finding corroborates evidence showing typical behavioral performance on ToM tasks in individuals with adult-onset amygdala damage (41) and extends these findings by demonstrating that this typical performance likely does not result from the deployment of compensatory strategies, because such alternative strategies would be expected to produce abnormal cortical responses to the task (42).

Hampton and colleagues (33) used fMRI to test for abnormalities in brain function in patients with amygdala lesions. At first glance, that study's observation of abnormal ventromedial prefrontal cortex function may seem at odds with those of the present study. However, that study specifically examined reward processing in a reversal learning task and therefore only underscores the need for caution when generalizing the present study findings to other behavioral and cognitive domains in which cortical interactions with the amygdala are perhaps more important.

The direct implications of our study are clear: The amygdala is not a necessary component of the cortical network for false-belief reasoning. The amygdala may not be required because false-belief reasoning draws principally on the cortical components or because the network as a whole sustains ToM abilities so that lesions to any single component, cortical or subcortical, would be insufficient to affect these abilities. There is some evidence that certain components of the ToM network may be essential for ToM abilities, but others are not: Lesion and transcranial magnetic stimulation studies implicate the temporoparietal junction as a necessary component (43, 44) but suggest that, like the amygdala in our study, the medial prefrontal cortex may be inessential (45).

Caveats and Future Directions. Several caveats that suggest important avenues for future research on the amygdala's role in higher-order social cognition should be mentioned. First, it is important to note that the lesions in both our patients are incomplete, with likely structural sparing of the central nucleus of the amygdala, as has been reported for other patients with Urbach–Wiethe disease (46). Intriguingly, this potentially spared area of the amygdala is consistent with the region that was activated in our whole-brain analysis of the MIT reference group (Fig. 1*C*). Recent evidence suggests that differential subnuclei connectivity may subserve separable, albeit complementary, cognitive/behavioral functions (47). Although there is no evidence showing functional activity in the spared portions of the amygdala in our two amygdala patients, it remains possible that the typical responses observed in the present study can be attributed to portions of the amygdala that are functionally spared. However, an exploratory analysis reported in *SI Results* provided no evidence that either patient showed a functional response to the Belief > Photo contrast in spared voxels in the vicinity of the amygdala. Future studies in additional patients with more complete amygdala lesions, such as the well-studied patient SM (9, 11), could help shed light on this issue.

Second, these findings cannot speak directly to accumulating evidence suggesting that the role of the amygdala in the performance of various ToM tasks may change over the course of development (41, 48, 49). Indeed, this evidence may account, in part, for less consistently observed amygdala activation in fMRI

Table 2. Comparison of the average patient response to the Belief > Photo contrast in the false-belief ROIs with the bootstrapped distribution of such responses estimated from the Caltech reference group

	Mean *t* value					Peak *t* value				
		Session 1		Session 2			Session 1		Session 2	
Region	Caltech	M	*P*	M	*P*	Caltech	M	*P*	M	*P*
Left TPJ	2.06	1.32	0.264	3.30	0.066	6.96	5.52	0.206	10.44	0.002
Right TPJ	2.39	1.66	0.364	3.03	0.428	8.29	6.53	0.335	9.02	0.690
Precuneus	2.58	2.00	0.615	3.87	0.260	8.43	6.85	0.404	9.32	0.637
DMPFC	1.69	0.79	0.405	2.43	0.495	6.37	4.43	0.130	8.29	0.134
MMPFC	1.65	0.76	0.503	1.24	0.758	5.80	3.83	0.289	7.64	0.319
VMPFC	1.28	0.22	0.126	1.57	0.673	5.15	3.69	0.155	6.47	0.201
Right STS	1.54	0.56	0.140	1.93	0.545	6.83	4.71	0.098	8.49	0.197

DM, dorsomedial; M, patient mean; MM, mid-medial; *P*, two-tailed probably value (uncorrected) for the null hypothesis that the patient mean is not different from the Caltech reference group mean; PFC, prefrontal cortex; STS superior temporal sulcus; TPJ, temporoparietal junction; VM, ventromedial.

studies of ToM in adulthood (23, 25, 26, 28). Developmentally transient amygdala function could account for the findings observed in the present study: The amygdala may well be necessary early in development to acquire normal ToM abilities but become inessential once this function has been offloaded to the mature cortical network for false-belief reasoning. The view that amygdala function may be most important for ToM early in development is supported by evidence suggesting that it plays a critical role in the early expression of joint attention (50, 51), which is thought to be a developmental precursor to ToM (52). Unfortunately, we do not know the age of onset of amygdala lesions in our patients, although we have surmised that their diseases calcified the amygdala around age 10 y (53). Other patients with amygdala lesions, some of them children and adolescents, are available, so in future studies it could be possible to probe ToM abilities across development in such a group (46).

Finally, it should be emphasized that the False-Belief Localizer engages ToM under the demands of a specific experimental task and depends strongly on language. When explicit cues are absent, as is the case in most natural social environments, evidence suggests that patients with amygdala lesions fail to exhibit the spontaneous use of ToM (14). Furthermore, there are a variety of ToM tasks that do not depend on language. Thus it would be important to test both performance and brain activation patterns in patients who have amygdala lesions on such a larger battery of ToM tasks. It remains possible that, even in adulthood, the amygdala plays a key role in the bottom-up control of cortical networks for ToM use, but this role may be revealed only on tasks that are relatively implicit in their cognitive demands, such as nonverbal tasks. This suggestion highlights the more general theme that ToM is quite heterogeneous in its behavioral expression, operational definition, and neural correlates (28, 35, 36). A more comprehensive investigation, such as the one in the present paper but over a larger battery of ToM tasks, could help parse that heterogeneity into types that do not depend on the amygdala and types that may.

Conclusion

We have shown that the amygdala is not a necessary component or modulator of the cortical network for false-belief reasoning assessed with the False-Belief Localizer. Conditional on the caveats we enumerated above, this conclusion was quite robust in our data: It held clearly for whole-brain and ROI-based analyses, and it was replicated across two different patients and across two experimental sessions in each patient. We also documented that the amygdala is indeed activated in healthy participants in the False-Belief Localizer, but that statistical power for detecting its activation requires unusually large sample sizes. Our study provides previously unidentified evidence concerning the amygdala's role in ToM processes and more generally demonstrates the power of combining lesion and fMRI studies in the same individuals.

Materials and Methods

Participants.

Patient group. The patient group originally included three females (referred to herein as "AP," "AM," and "BG") who had focal bilateral amygdala lesions caused by Urbach–Wiethe disease (34). AP is an English-speaking American, was 27 y of age at testing, has worked since she obtained her Bachelor's degree, and is fully right-handed. AM and BG are identical twin sisters from rural southern Germany. They were 36 y of age at testing, are married with children, have been in full-time employment since they completed 13 y of education in Germany. Although BG is fully right-handed, her sister AM is fully left-handed. Given that our control groups were entirely right-handed, and that the False-Belief Localizer task features strong language demands and produces hemispherically asymmetric cortical responses, we chose to exclude AM's data from the present study. Hence, our final patient group consisted of AP and BG, who both have IQs in the average range [BG: Hamburg-Wechsler Intelligence Test for Adults-Revised (HAWIE-R) score: 96; AP: Wechsler Abbreviated Scale of Intelligence (WASI) score: 98] (54). Their lesions are similarly symmetric and confined to the amygdala (BG, 1.15 cm^3; AP, 0.71 cm^3). The damage includes complete ablation of the basolateral amygdala with minor damage to other amygdaloid regions, including anterior and ventral regions at the rostral level and lateral and medial parts of the central nucleus and amygdalo–hippocampal area at the caudal level (Fig. 1*A*). Each patient participated in two separate sessions, both of which involved performing the False-Belief Localizer while undergoing fMRI at the Caltech Brain Imaging Center (CBIC).

The two patients with amygdala lesions were compared with two healthy comparison groups. The first group, the Caltech reference group, provided the closest comparison, because participants were scanned on the same scanner and task as the amygdala patients; the second group, the MIT reference group, provided a larger and more generalizable independent reference group against which our data could be compared. Given that published data on a large sample has documented that there are no apparent age and sex differences in responses to the False-Belief Localizer (40), we included participants regardless of age and sex to maximize the size of our reference groups.

Caltech reference group. The first reference group consisted of 18 neurologically healthy adults (13 males and 5 females; mean age, 28.44 y; age range, 21–46 y), all of whom performed the most recent version of the False-Belief Localizer while undergoing fMRI at the CBIC. Each participant was neurologically and psychiatrically healthy, had normal or corrected-to-normal vision, spoke English fluently, had IQ in the normal range (as assessed using the WAIS), and was not pregnant or taking any psychotropic medications.

MIT reference group. The second reference group consisted of 462 neurologically healthy adults (223 males, 239 females; mean age, 24.9 y; age range, 18–69 y), all of whom performed some version of the False-Belief Localizer while undergoing fMRI at the Martinos Imaging Center for Brain Research at MIT between 2006 and 2013. Complete details about this reference group can be found in Dufour et al. (40).

All participants in the three groups provided written informed consent according to protocols approved by the Institutional Review Boards of the California Institute of Technology or MIT and were compensated monetarily for their time.

False-Belief Localizer Task. The patient and Caltech reference groups performed the most recent version of the publicly available False-Belief Localizer (Fig. 1*B*) (22) (downloaded from saxelab.mit.edu/tomloc.zip, version September 7, 2011). The MIT reference group performed either this most recent (English) version of the task or one of several earlier versions that featured the same conceptual contrast, namely, False-Belief versus False-Photo verbal scenarios, but which differed in one or more minor methodological details (for further details, see ref. 40). Additional information about the task and the analysis of behavioral outcomes are provided in *SI Materials and Methods*.

Image Acquisition. Imaging data for the patient group and the Caltech reference group was acquired using a Siemens Trio 3.0-Tesla MRI scanner outfitted with a 32-channel phased-array head-coil. We acquired 242 T2*-weighted echoplanar image (EPI) volumes (slice thickness = 3 mm, 47 slices, TR = 2,500 ms, TE = 30 ms, flip angle = 85°, matrix = 64 × 64, FOV = 192 mm). We also acquired a high-resolution anatomical T1-weighted image (1 mm isotropic) and field maps for each participant. Imaging data for the MIT control group was acquired using a Siemens 3.0-Tesla MRI scanner outfitted with a 32-channel (n = 74) or 12-channel (n = 388) head-coil (variable slice thickness; in-plane resolution of 3.125 × 3.125 mm; TR = 2,000 ms; TE = 30 ms; flip = 90°).

Image Analysis. Image preprocessing and analysis was conducted using Statistical Parametric Mapping (SPM8; Wellcome Department of Cognitive Neurology, London). Details regarding the preprocessing pipeline and single-subject model estimation are provided in *SI Materials and Methods*. Following model estimation, we computed the Belief > Photo contrast image for each participant, along with a statistical t-image indexing the reliability of the Belief > Photo contrast across the whole brain. Our analyses are focused on this latter contrast and were aimed at answering the question: Is this image atypical in our patient group compared with either the Caltech or MIT reference groups?

To empirically estimate the typical distribution of activity from the smaller Caltech reference group (n = 18), we used a bootstrapping procedure to construct a distribution of the average response for every possible combination of two individuals [in MATLAB: nchoosek(1:18, 2)]. This procedure yielded a bootstrapped population estimate based on 153 groups of two,

which we used as a reference to evaluate the typicality of the average response of patient AP and BG.

Using the MIT group-level unthreshholded and gray matter-masked Belief > Photo contrast map as a benchmark (n = 462), we first determined if the overall spatial response pattern observed in the Caltech group was more typical than that in the patient group. We next examined the pattern of response in a mask containing all a priori functional ROIs that were defined on the basis of the Belief > Photo contrast in the MIT reference group. As before, we used the spatial pattern observed in the MIT reference group as a benchmark. Finally, we examined the magnitude (mean and peak) and peak location (x-, y-, and z-coordinates) of the patient response in seven cortical ROIs. These ROIs were defined from the group-level contrast observed in the MIT reference group in a manner consistent with previous literature (21, 22): the right and left temporoparietal junction, the precuneus, the dorsal, middle, and ventral components of the medial prefrontal cortex, and the right superior temporal sulcus. These ROIs are displayed in Fig. S2.

We capitalized on the large MIT reference group to perform a comparison focused on the individual patient response data. We compared the whole-brain (gray matter-masked) spatial pattern of the Belief > Photo contrast for each patient with each individual in the MIT reference group (n = 462). To create a leave-one-out reference distribution, we took each individual in the MIT reference group and computed the mean Pearson correlation of their whole-brain response with each remaining member of the MIT reference group. For both AP and BG and for each session separately, we computed the Pearson correlation of their whole-brain response with every member of the MIT reference group. We then compared the mean of the resulting correlation distribution with the actual typical distribution of such correlation means estimated from the MIT group.

ACKNOWLEDGMENTS. We thank Mike Tyszka, the California Institute of Technology Brain Imaging Center, and the Martinos Imaging Center at MIT for help with the neuroimaging. Funding support was supplied by the Della Martin Foundation, the National Institute of Mental Health, the Packard Foundation, and the Simons Foundation.

1. Adolphs R (2010) What does the amygdala contribute to social cognition? *Ann N Y Acad Sci* 1191:42–61.
2. Brothers L (1990) The social brain: A project for integrating primate behavior and neurophysiology in a new domain. *Concepts Neurosci* 1:27–51.
3. Frith CD, Frith U (2007) Social cognition in humans. *Curr Biol* 17(16):R724–R732.
4. Whalen PJ, Phelps EA (2009) *The Human Amygdala* (Guilford Press, New York).
5. Gothard KM, Battaglia FP, Erickson CA, Spitler KM, Amaral DG (2007) Neural responses to facial expression and face identity in the monkey amygdala. *J Neurophysiol* 97(2):1671–1683.
6. Rutishauser U, et al. (2011) Single-unit responses selective for whole faces in the human amygdala. *Curr Biol* 21(19):1654–1660.
7. Emery NJ, et al. (2001) The effects of bilateral lesions of the amygdala on dyadic social interactions in rhesus monkeys (Macaca mulatta). *Behav Neurosci* 115(3):515–544.
8. Machado CJ, Bachevalier J (2006) The impact of selective amygdala, orbital frontal cortex, or hippocampal formation lesions on established social relationships in rhesus monkeys (Macaca mulatta). *Behav Neurosci* 120(4):761–786.
9. Adolphs R, Tranel D, Damasio H, Damasio A (1994) Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature* 372(6507):669–672.
10. Adolphs R, et al. (1999) Recognition of facial emotion in nine individuals with bilateral amygdala damage. *Neuropsychologia* 37(10):1111–1117.
11. Adolphs R, Tranel D, Damasio AR (1998) The human amygdala in social judgment. *Nature* 393(6684):470–474.
12. Kennedy DP, Gläscher J, Tyszka JM, Adolphs R (2009) Personal space regulation by the human amygdala. *Nat Neurosci* 12(10):1226–1227.
13. Schultz RT, et al. (2003) The role of the fusiform face area in social cognition: Implications for the pathobiology of autism. *Philos Trans R Soc Lond B Biol Sci* 358(1430): 415–427.
14. Heberlein AS, Adolphs R (2004) Impaired spontaneous anthropomorphizing despite intact perception and social knowledge. *Proc Natl Acad Sci USA* 101(19):7487–7491.
15. Stone VE, Baron-Cohen S, Calder A, Keane J, Young A (2003) Acquired theory of mind impairments in individuals with bilateral amygdala lesions. *Neuropsychologia* 41(2): 209–220.
16. Baron-Cohen S (2004) The cognitive neuroscience of autism. *J Neurol Neurosurg Psychiatry* 75(7):945–948.
17. Pelphrey K, Adolphs R, Morris JP (2004) Neuroanatomical substrates of social cognition dysfunction in autism. *Ment Retard Dev Disabil Res Rev* 10(4):259–271.
18. Kennedy DP, Adolphs R (2012) The social brain in psychiatric and neurological disorders. *Trends Cogn Sci* 16(11):559–572.
19. Gallagher HL, Frith CD (2003) Functional imaging of 'theory of mind'. *Trends Cogn Sci* 7(2):77–83.
20. Amodio DM, Frith CD (2006) Meeting of minds: The medial frontal cortex and social cognition. *Nat Rev Neurosci* 7(4):268–277.
21. Saxe R, Kanwisher N (2003) People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind". *Neuroimage* 19(4):1835–1842.
22. Dodell-Feder D, Koster-Hale J, Bedny M, Saxe R (2011) fMRI item analysis in a theory of mind task. *Neuroimage* 55(2):705–712.
23. Carrington SJ, Bailey AJ (2009) Are there theory of mind regions in the brain? A review of the neuroimaging literature. *Hum Brain Mapp* 30(8):2313–2335.
24. Spreng RN, Mar RA, Kim AS (2009) The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: A quantitative meta-analysis. *J Cogn Neurosci* 21(3):489–510.
25. Van Overwalle F, Baetens K (2009) Understanding others' actions and goals by mirror and mentalizing systems: A meta-analysis. *Neuroimage* 48(3):564–584.
26. Lieberman M (2010) *Handbook of Social Psychology*, eds Fiske ST, Gilbert DT, Lindzey G (McGraw-Hill, New York), pp 143–193.
27. Mar RA (2011) The neural bases of social cognition and story comprehension. *Annu Rev Psychol* 62:103–134.
28. Schurz M, Radua J, Aichhorn M, Richlan F, Perner J (2014) Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neurosci Biobehav Rev* 42: 9–34.
29. Merboldt KD, Fransson P, Bruhn H, Frahm J (2001) Functional MRI of the human amygdala? *Neuroimage* 14(2):253–257.
30. Carr VA, Rissman J, Wagner AD (2010) Imaging the human medial temporal lobe with high-resolution fMRI. *Neuron* 65(3):298–308.
31. Amaral DG, Price JL, Pitkanen A, Carmichael ST (1992) *The Amygdala: Neurobiological Aspects of Emotion, Memory and Mental Dysfunction*, ed Aggleton JP (Wiley-Liss, New York), pp 1–66.
32. Ongür D, Price JL (2000) The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cereb Cortex* 10(3):206–219.
33. Hampton AN, Adolphs R, Tyszka MJ, O'Doherty JP (2007) Contributions of the amygdala to reward expectancy and choice signals in human prefrontal cortex. *Neuron* 55(4): 545–555.
34. Hofer P-A (1973) Urbach-Wiethe disease (lipoglycoproteinosis; lipoid proteinosis; hyalinosis cutis et mucosae). A review. *Acta Derm Venereol Suppl (Stockh)* 53:1–52.
35. Schaafsma SM, Pfaff DW, Spunt RP, Adolphs R (2015) Deconstructing and reconstructing theory of mind. *Trends Cogn Sci* 19(2):65–72.
36. Spunt RP, Adolphs R (2014) Validating the Why/How contrast for functional MRI studies of Theory of Mind. *Neuroimage* 99:301–311.
37. Dennett DC (1978) Beliefs about beliefs. *Behav Brain Sci* 1:568–570.
38. Wellman HM, Cross D, Watson J (2001) Meta-analysis of theory-of-mind development: The truth about false belief. *Child Dev* 72(3):655–684.
39. Bloom P, German TP (2000) Two reasons to abandon the false belief task as a test of theory of mind. *Cognition* 77(1):B25–B31.
40. Dufour N, et al. (2013) Similar brain activation during false belief tasks in a large sample of adults with and without autism. *PLoS ONE* 8(9):e75468.
41. Shaw P, et al. (2004) The impact of early and late damage to the human amygdala on 'theory of mind' reasoning. *Brain* 127(Pt 7):1535–1548.
42. Mihov Y, et al. (2013) Mirroring fear in the absence of a functional amygdala. *Biol Psychiatry* 73(7):e9–e11.
43. Samson D, Apperly IA, Chiavarino C, Humphreys GW (2004) Left temporoparietal junction is necessary for representing someone else's belief. *Nat Neurosci* 7(5):499–500.
44. Young L, Camprodon JA, Hauser M, Pascual-Leone A, Saxe R (2010) Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. *Proc Natl Acad Sci USA* 107(15):6753–6758.
45. Bird CM, Castelli F, Malik O, Frith U, Husain M (2004) The impact of extensive medial frontal lobe damage on 'Theory of Mind' and cognition. *Brain* 127(Pt 4):914–928.
46. Van Honk J *Living Without an Amygdala*, eds Amaral DG, Bauman MD, Adolphs R (Guilford Press, New York).
47. Senn V, et al. (2014) Long-range connectivity defines behavioral specificity of amygdala neurons. *Neuron* 81(2):428–437.
48. Fine C, Lumsden J, Blair RJ (2001) Dissociation between 'theory of mind' and executive functions in a patient with early left amygdala damage. *Brain* 124(Pt 2):287–298.
49. Rice K, Viscomi B, Riggins T, Redcay E (2014) Amygdala volume linked to individual differences in mental state inference in early childhood and adulthood. *Dev Cogn Neurosci* 8:153–163.
50. Elison JT, et al.; IBIS Network (2013) Frontolimbic neural circuitry at 6 months predicts individual differences in joint attention at 9 months. *Dev Sci* 16(2):186–197.
51. Mosconi M, et al. (2010) Neurobehavioral abnormalities in first-degree relatives of individuals with autism. *Arch Gen Psychiatry* 67(8):830–840.
52. Tomasello M, Carpenter M, Call J, Behne T, Moll H (2005) Understanding and sharing intentions: The origins of cultural cognition. *Behav Brain Sci* 28(5):675–691, discussion 691–735.
53. Feinstein JS, Adolphs R, Damasio A, Tranel D (2011) The human amygdala and the induction and experience of fear. *Curr Biol* 21(1):34–38.
54. Becker B, et al. (2012) Fear processing and social networking in the absence of a functional amygdala. *Biol Psychiatry* 72(1):70–77.
55. Tzourio-Mazoyer N, et al. (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the M mri single-subject brain. *NeuroImage* 15(1):273–289.