Social inference deficits in temporal lobe epilepsy and lobectomy: risk factors and neural substrates

Melanie Cohn,1,2 Marie St-Laurent,3 Alexander Barnett,2 and Mary Pat McAndrews1,2

1Krembil Neuroscience Centre at Toronto Western Hospital – UHN, Toronto, ON, Canada, 2Department of Psychology, University of Toronto, Toronto, ON, Canada, and 3Rotman Research Institute at Baycrest, Toronto, ON, Canada

In temporal lobe epilepsy and lobectomy, deficits in emotion identification have been found consistently, but there is limited evidence for complex social inference skills such as theory of mind. Furthermore, risk factors and the specific neural underpinnings of these deficits in this population are unclear. We investigated these issues using a comprehensive range of social inference tasks (emotion identification and comprehension of sincere, deceitful and sarcastic social exchanges) in individuals with temporal lobe epilepsy or lobectomy (n = 87). We observed deficits across patient groups which were partly related to the presence of mesial temporal lobe sclerosis, early age of seizure onset and left lobectomy. A voxel-based morphometry analysis conducted in the pre-operative group confirmed the importance of the temporal lobe by showing a relationship between left hippocampal atrophy and overall social inference abilities, and between left anterior neocortex atrophy and sarcasm comprehension. These findings are in keeping with theoretical proposals that the hippocampus is critical for binding diverse elements in cognitive domains beyond canonical episodic memory operations, and that the anterior temporal cortex is a convergence zone of higher-order perceptual and emotional processes, and of stored representations. As impairments were frequent, we require further investigation of this behavioural domain and its impact on the lives of people with epilepsy.

Keywords: theory of mind; mentalizing; emotion; hippocampus; temporal pole

INTRODUCTION

Temporal lobe epilepsy (TLE) is characterized by anterior temporal epileptiform discharges and, frequently, by structural abnormalities such as mesial temporal lobe sclerosis (MTS) found primarily in the hippocampus. Individuals with TLE are commonly referred for a unilateral anterior temporal lobectomy (ATL), which encompasses the hippocampus, amygdala and anterior temporal neocortex, to treat their epilepsy when refractory to medication. From a cognitive standpoint, episodic memory deficits are common in this population, particularly when MTS is present and/or following surgery (Chelune, 1995; Jones-Gotman et al., 2010; Sherman et al., 2011). Although considerable research efforts have been devoted to language and memory functioning in TLE and ATL, less attention has been paid to other cognitive skills governed by anterior and mesial temporal structures. Such a contending cognitive domain is social inference, which we conceptualize as including a range of abilities from emotion identification to theory of mind skills. Paradigms used to assess these skills include simple tasks that focus on a single type of inference and/or cue (e.g. facial emotion identification from photographs) to more complex tasks involving multiple inferences (emotion, thoughts and intention) based on the interpretation and integration of various cues (discourse, intonation, facial expression, gestures, context and social norms). Examples of such tasks include social faux-pas detection, false-belief reasoning and sarcasm comprehension.

Converging evidence from neuroimaging and neurological populations supports the involvement of temporal lobe regions in social inference. Specifically, fMRI studies in healthy participants demonstrate engagement of the amygdala, hippocampus and anterior temporal neocortex in addition to the engagement of the superior temporal sulcus and of prefrontal regions (Frith and Frith, 2006; Carrington and Bailey, 2009; Spreng et al., 2009). The relative involvement of these temporal regions is a function of task characteristics. For instance, amygdala activation is particularly salient when aversive emotional material is used (e.g. fearful faces; Zald, 2003), and anterior temporal neocortex engagement is observed most consistently with contextually rich narratives (e.g. video clips, stories, cartoons; Carrington and Bailey, 2009). A few studies also report hippocampal activation during real-time social interactions (Schilbach et al., 2006; Polosan et al., 2011), event construction (Mitchell et al., 2005; Rabin et al., 2010) and disambiguation of social cues (i.e. facial expressions) in working memory (Ross et al., 2013), as well as a correlate of perceived irony (Akimoto et al., 2014). In neurological populations, decline in social and emotional functioning is a hallmark of fronto-temporal dementia (FTD), and recent brain volumetric studies have linked atrophy in the temporal poles, parahippocampal gyrus and amygdala, in addition to orbitofrontal regions, to impairments in emotion recognition and sarcasm comprehension (Kipps et al., 2009; Rankin et al., 2009). A similar relationship between temporal pole volume and performance on a social faux-pas task was found in schizophrenia (Herold et al., 2009). Atrophy in the hippocampal formation has also been related to theory of mind in children with traumatic brain injury (Dennis et al., 2013), and decreased hippocampal engagement was observed during a social inference task in individuals with schizophrenia (Andreasen et al., 2008).

In TLE and ATL, reduced facial emotion recognition has been reported frequently (Meletti et al., 2003, 2009; Hlobil et al., 2008; Broicher et al., 2012), but only five studies have been published to date on complex social inference abilities. Four of these studies demonstrated poorer performance relative to controls or other non-temporal epilepsy groups on social faux-pas detection (Schacher et al., 2006; Giovagnoli et al., 2011; Broicher et al., 2012; Li et al., 2013), the attribution of social behaviours to moving triangles (Broicher et al., 2012) as well as false belief reasoning, processing of implied meanings and cartoon theory of mind tasks (Li et al., 2013). The fifth study (Shaw et al., 2007) reported intact social faux-pas detection in a small patient group (n=19) both prior to and after surgery; however, performance was at ceiling.
These findings demonstrate that TLE and ATL groups have reduced social inference abilities and given that they perform more poorly than other epilepsy groups, this deficit is thought to have a neural underpinning rather than a psychosocial cause (e.g. stigma, social isolation). However, the specific neural substrates remain unclear. Individuals with MTS have impoverished social inference skills relative to individuals with other temporal lesions or with non-lesional epilepsy (Meletti et al., 2003; Giovagnoli et al., 2011), and the few studies that compared ATL with TLE patients (most of whom had MTS) failed to detect significant differences (Schacher et al., 2006; Shaw et al., 2007; Hlobil et al., 2008; Tanaka et al., 2013). These suggest that medial temporal lobe (MTL) damage is the critical common factor. While the amygdala is an obvious candidate region given its well-established role in emotional processing, evidence of impaired emotion recognition in the absence of amygdala abnormalities (Adolphs et al., 2001) and preserved emotion recognition despite amygdala damage (Siebert et al., 2003; Fowket et al., 2006) challenge the claim that such damage is necessary or sufficient. Based on neuroimaging findings in healthy controls (Carrington and Bailey, 2009; Spreng et al., 2009), it is plausible that both hippocampal and anterior neocortical disruptions contribute to social inference deficits in TLE and ATL. Such deficits may also be related to extra-temporal damage, as there is some overlap between regions showing structural changes in TLE (Keller and Roberts, 2008; Li et al., 2012) and regions involved in social inference skills in healthy individuals (Frith and Frith, 2006; Carrington and Bailey, 2009; Spreng et al., 2009) and in individuals with brain dysfunction (Stone et al., 1998; Stuss et al., 2001; Kipps et al., 2009; Rankin et al., 2009; Dennis et al., 2013). As for the effect of laterality of epilepsy or MTL damage, data to date are highly inconsistent (right worse than left (e.g. Meletti et al., 2009; Li et al., 2013), left worse than right (e.g. Carvajal et al., 2009; Giovagnoli et al., 2011) and no laterality effect (e.g. Bonora et al., 2011; Tanaka et al., 2013).

In addition to lesion type and location, timing of damage may represent a key risk factor for social inference deficits. Developmental research in healthy cohorts indicates an age range in which different abilities are acquired, from early childhood for emotion recognition to teenage years for advanced skills such as sarcasm comprehension (Kolb et al., 1992; Herba and Phillips, 2004; Glenwright and Fexman, 2010). Damage sustained during these critical periods may result in social inference deficits that persist in adulthood, but damage sustained after these periods may have less or no effect. Indeed, younger age of seizure onset and longer duration of epilepsy are associated with emotion recognition (Hlobil et al., 2008; Meletti et al., 2009) and with complex social inference deficits (Schacher et al., 2006; Giovagnoli et al., 2011) (but see Broicher et al., 2012; Li et al., 2013 for alternative findings). Results relating MTS and social inference deficits are in keeping with this given that these structural changes often result from a febrile illness in early childhood (Cendes, 2004) and that the effect of MTS on emotion recognition was no longer significant once age of onset and duration were controlled for (Meletti et al., 2009).

Nonetheless, evidence from temporal lobe stroke and FTD shows that adult-onset temporal lobe damage can negatively impact social inference skills (Kipps et al., 2009; Rankin et al., 2009; Xi et al., 2013), which raise the possibility that ATL surgery can exacerbate social inference deficits.

In the present study, our goal was to address the gaps and inconsistencies in this literature by investigating a comprehensive range of social inference tasks in individuals with TLE and ATL. We used The Awareness of Social Inference Test (TASIT; McDonald et al., 2003), a standardized test that includes a measure of emotion identification and complex social inference tasks requiring the attribution of emotions, intentions, actions and thoughts to characters during sincere, deceitful and sarcastic social exchanges. We believe this task to be particularly sensitive to temporal lobe damage given the emotional content and narrative base of the video vignettes as well as findings relating poor performance to amygdala and temporal pole atrophy in FTD (Kipps et al., 2009; Rankin et al., 2009). The TASIT has also been used in a number of other clinical populations (e.g. TBI, autism, schizophrenia; McDonald et al., 2003; Sparks et al., 2010; Mathersul et al., 2013), but not in TLE and ATL.

To investigate the neural substrates of social inference in TLE and ATL, we compared groups based on surgical status and we correlated behaviour to brain volumetric measures in TLE. We also investigated the influence of other variables such as the presence of MTS or other lesions, age of seizure onset, duration of epilepsy, intellectual abilities, sex and age. We hypothesized that social inference will be impaired in TLE and ATL, and that this deficit will be associated with an early age of onset, MTS and surgery. In TLE, we also predicted that reduced medial (hippocampus and amygdala) and anterior temporal volumes will be related to poor social inference skills.

Materials and Methods
Participants
Fifty individuals with intractable unilateral TLE identified as surgical candidates based on their multidisciplinary evaluation (video-EEG, 3TMRI and neuropsychological assessment) and 37 individuals who had undergone a unilateral anterior temporal lobectomy were recruited from the Epilepsy Surgery Program at Toronto Western Hospital to participate in this study. Demographic and clinical data are presented in Table 1. From the pre-operative group, 27 had right-sided seizure onset (RTLE) and 24 had left-sided seizure onset (LTLE), and from the post-operative group, 19 had a right surgery (RATL) and 18 had a left surgery (LATL). All ATL surgeries included resection of the hippocampus, amygdala and various degrees of the temporal neocortex (i.e. 30 patients had standard en-bloc excisions removing ~4.5 cm of the anterior inferior and middle temporal gyri, and four LATL and three RATL had a more restricted neocortical resection). All patients had left-language lateralization based on functional MRI. Participants from the ATL group were tested at least 6 months before surgery.

Table 1: Demographic and disease characteristics

<table>
<thead>
<tr>
<th>Groups</th>
<th>N</th>
<th>Male: Female</th>
<th>Age Mean (SD)</th>
<th>Education</th>
<th>FSIQ Mean (SD)</th>
<th>Age of onset Mean (SD)</th>
<th>Duration of illness Mean (SD)</th>
<th>MRI findings (MTS:non-MTS lesion: normal)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LTLE</td>
<td>24</td>
<td>13:11</td>
<td>38.9 (11.9)</td>
<td>15.6 (2.7)</td>
<td>107.7 (11.6)</td>
<td>22.4 (13.2)</td>
<td>16.5 (14.1)</td>
<td>7:6:1</td>
</tr>
<tr>
<td>RTLE</td>
<td>26</td>
<td>14:12</td>
<td>38.0 (12.7)</td>
<td>13.5 (1.2)</td>
<td>102.8 (12.4)</td>
<td>18.0 (13.2)</td>
<td>19.6 (15.3)</td>
<td>14:8:1</td>
</tr>
<tr>
<td>LATL</td>
<td>18</td>
<td>11:7</td>
<td>42.5 (12.9)</td>
<td>14.6 (2.7)</td>
<td>102.8 (12.0)</td>
<td>16.7 (16.2)</td>
<td>24.0 (18.0)</td>
<td>15:2:1</td>
</tr>
<tr>
<td>RATL</td>
<td>19</td>
<td>7:12</td>
<td>38.9 (9.6)</td>
<td>14.4 (2.2)</td>
<td>105.4 (8.0)</td>
<td>16.6 (13.0)</td>
<td>20.0 (12.0)</td>
<td>12:3:4</td>
</tr>
</tbody>
</table>

Note: Mean and s.d.; MTS: MTLE; N: sample size; FSIQ: full-scale intellectual quotient; LTLE, left temporal epilepsy; RTLE, right temporal epilepsy; LATL, left anterior lobectomy; RATL, right anterior lobectomy.
post-operatively, were seizure-free and remained on anti-epileptic medications. None of the participants had major psychiatric disorders that could interfere with their participation or had a full-scale intellectual quotient (FSIQ) below 78 (<1.5 SD below normative mean) on the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999). FSIQ scores were also used in our behavioural analyses.

In addition, 15 age-matched control participants with no history of neurological or psychiatric condition were recruited from the community. We also used the TASIT normative data based on much larger samples (n = 88 Part 1, n = 98 Part 2 and n = 123 Part 3) to assess frequency of impairment in patients given that such analyses are problematic when few or no participants have impairments. All TLE/ATL participants (n = 87) and healthy controls were included in the behavioural study, and only TLE participants with MTS or a normal MRI were included in the brain volumetric analyses (RTLE: n = 22, LTLE: n = 18).

**Social inference task**
Participants completed version A of the TASIT (McDonald et al., 2003). Part 1, the Emotion Evaluation Test, includes 28 short video clips in which an actor portrays one of seven emotional states (happy, surprised, sad, angry, anxious, disgusted and neutral) and participants are required to identify the corresponding emotion using a seven-alternative forced-choice. Parts 2 and 3 include video-vignettes depicting social interactions between characters, and participants are required to answer yes–no questions pertaining to the character’s emotional state, belief and intention (what the character wants the partner to think or feel), and meaning of the exchange (whether a literal or non-literal interpretation is correct). Part 2 includes five sincere exchanges in which correct interpretations can be based on the literal meaning of the dialogue, as well as 10 sarcastic exchanges, in which the content of the dialogue is incongruent with other paralinguistic cues and is intended to be interpreted as opposite to the literal meaning. Part 3 includes eight sarcastic and eight deceitful interactions. Part 3 vignettes are ‘enriched’ relative to Part 2 in that additional evidence is provided to support the correct meaning of the interaction (visual cues or prologue in which the truth is revealed). This enrichment is critical for deceitful exchanges given that the interaction would otherwise be interpreted as sincere.

**Behavioural analyses**
Analyses were done on the total number of correct answers on the emotion evaluation (Part 1), sincere (Part 2), deceitful (Part 3) and sarcastic items (combining Parts 2 and 3). We used non-parametric Kruskal–Wallis analyses to compare scores from the five groups (controls, RTLE, LTLE, RATL and LATL), and, if significant, we conducted post-hoc pairwise comparisons using Conover–Inman, a Fisher’s least significant difference method performed on ranks. Individual scores on the emotion evaluation, sincere, deceitful, sarcasm-part 2 and sarcasm-part 3 items were also compared with the published normative data and were transformed into binary scores representing the presence or absence of a clinically relevant impairment (i.e. ≤5% percentile of normative sample). The frequency of impairments was compared between the combined clinical group (TLE + ATL) and the published normative sample using the Yates-corrected chi-square as an omnibus test, which is a conservative approach used when frequencies are small, and odds ratio are provided as indicators of effect size. When this comparison yielded statistically significant results, it was followed by post-hoc pairwise analyses between the five groups (normative group and four patient groups) using the same statistical test.

To investigate the relationship between social inference and other variables, we used Spearman correlations for continuous variables (FSIQ, age, age of onset and duration of epilepsy) and Kruskal–Wallis for categorical variables (sex and lesion types). We also used forward stepwise multiple linear regressions to develop models of the best predictors of social inference abilities in TLE and ATL. For these analyses, we set tolerance level at 0.01 and required a probability P < 0.15 for a variable to be entered in the model. Variables with the highest correlations with the dependent variables were entered sequentially in the model until further additions failed to improve the model. All statistical analyses were conducted in SYSTAT (Systat Software Inc.)

**Neuroimaging**
Scan acquisition (TLE groups). Anatomical scans were acquired on a 3-T Sigma MR System (GE Medical Systems) as part of the standard pre-operative clinical workup. These consisted of 3D spoiled gradient-recalled T1-weighted sequence with axial acquisition, 146 slices, 220 mm FOV and 256 x 256 matrix, which resulted in a voxel size of 0.78125 x 0.78125 x 1.0.

Voxel-based morphometry analyses. We used VBM8 (http://dbm.neuro.uni-jena.de/vbm/) to preprocess MRI scans, which involved segmentation, DARTEL normalization, the creation of modulated images to provide an index of grey matter (GM) volume at each voxel and spatial smoothing using a 8 mm FWHM isotropic Gaussian kernel. Statistical analyses of the smoothed modulated images were done in SPM8 (http://www.fil.ion.ucl.ac.uk/spm/). For whole-brain analyses, we used a mixed general linear model (GLM) based on our behavioural analyses. We included three factors representing social inference deficits (emotion, deceit and sarcasm scores), four covariates including laterality of the epilepsy focus and variables showing a unique contribution to behaviour in our regression analyses (age, FSIQ and age of onset). Threshold masking was set to 0.1, and extent threshold was set to P < 0.001 based on random field theory, which corresponded to clusters ≥53 voxels for the conjunction (global null) analysis combining all tasks of social inference, and to clusters ≥77 voxels for single contrasts for individual tasks. The global null conjunction analysis does not require that each contrast is significant individually, but rather that contrasts’ effects are in a consistent direction and jointly significant (Friston et al., 2005).

**RESULTS**

**Demographic data**
As shown in Table 1, there were no significant differences between groups (LTLE, RTLE, LATL, RATL and controls) in age [KW(3) = 1.92, P = 0.76] and education [KW(3) = 7.66, P = 0.11], and no differences between patient groups in terms of FSIQ [KW(3) = 2.14, P = 0.54], epilepsy duration [KW(3) = 2.57, P = 0.46] and age of seizure onset [KW(3) = 2.99, P = 0.39]. However, the proportion of individuals with lesions on their pre-operative MRI differed between groups [Pearson χ²(3, N = 87) = 8.91, P = 0.03] in that lesions were more frequently observed in the LATL group than in the LTLE group [Yates-corrected χ²(1, N = 43) = 6.32, P = 0.01] and specifically, MTS (83% vs 29%).

**Emotion evaluation**
As shown in Figure 1, emotion recognition scores differed across the five groups [KW(4) = 29.22, P < 0.001]. All patient groups performed more poorly than controls [LATL: Conover–Inman = 5.98, P = 0.001; RATL: Conover–Inman = 2.24, P = 0.03; LTLE: Conover–Inman = 3.14, P = 0.002; RTLE: Conover–Inman = 4.13, P < 0.001], and LATL

---

1 Sarcasm scores for Part 2 and 3 were compared to normative data separately because normative data for a combined score are not published. These scores were combined for the rank analyses for brevity and because this measure has increased range and is normally distributed.
group performed below other patient groups [RATL: Conover–Inman = 4.01, \( P < 0.001 \); LTLE: Conover–Inman = 3.40, \( P = 0.001 \); RTLE: Conover–Inman = 2.46, \( P = 0.02 \)]. There was also a trend for poorer emotion recognition in RTLE than RATL [Conover–Inman = 1.88, \( P = 0.06 \)]. Not only did TLE and ATL groups obtain weaker scores relative to controls, but as shown in Table 2, 25% of them obtained frankly impaired scores vs 5% of the normative group. Specifically, impairments were more frequent in RTLE and LATL relative to healthy individuals, but not in LTLE or RATL. Deficits were also more frequent in LATL than RATL. Analyses of specific emotions are reported in supplementary materials.

**Comprehension of sincere exchanges**

There were no differences in rank scores between groups [KW(4) = 1.86; \( P = 0.76 \); see Figure 1], nor were there differences in the frequency of impairments between patients and healthy individuals (Table 2). This finding suggests that patients are able to make social inferences accurately when multiple sources of information (e.g. discourse content, paralinguistic and contextual cues) are congruous and when the literal meaning is true.

**Comprehension of deceitful exchanges**

In deceitful exchanges, cues are consistent with one another with the exception of one fact which allows the interpretation of the social exchange as deceitful rather than as sincere. As shown in Figure 1, performance of every patient group was weaker than that of controls (main effect: KW(4) = 28.53, \( P < 0.001 \); pairwise comparisons: LATL: Conover–Inman = 5.10, \( P < 0.001 \); RATL: Conover–Inman = 3.83, \( P < 0.001 \); LTLE: Conover–Inman = 3.01, \( P = 0.003 \); RTLE: Conover–Inman = 5.56, \( P < 0.001 \)], and both LATL and RTLE performed more poorly than LTLE [LATL: Conover–Inman = 2.55, \( P = 0.01 \); RTLE: Conover–Inman = 2.87, \( P = 0.005 \)]. However, as shown in Table 2, the frequency of deficits was not significantly different between patients and the normative group. The discrepancy between analyses of ranks and of frequency of impairment suggests that patients have only mild difficulties with this type of social exchange.

**Comprehension of sarcastic exchanges**

In sarcastic exchanges, the discourse content is incongruent with paralinguistic cues and social context, and this ambiguity must be resolved by integrating all cues to derive the accurate meaning of the
Table 2 Frequency of social inference impairments

<table>
<thead>
<tr>
<th>Social inference tasks</th>
<th>Normative group (NG)</th>
<th>Overall TLE/ATL group (n = 87)</th>
<th>Patient subgroups</th>
<th>Main effect and pairwise subgroup comparisons statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>LTLE (n = 24)</td>
<td>RTLE (n = 26)</td>
</tr>
</tbody>
</table>
| Emotion                | 5%                   | 25.3%*                        | 16.7%             | 30.8%*       | 44.4%*       | 10.5%         | Main effect (NG vs overall TLE/ATL): Odds ratio = 5.62
|                        |                      |                               |                   |              |              |              | Yates-corrected $\chi^2(1, N = 175) = 11.43, P = 0.001$
|                        |                      |                               |                   |              |              |              | Pairwise subgroup comparisons:
|                        |                      |                               |                   |              |              |              | LTLE vs NG: Yates-corrected $\chi^2(1, N = 112) = 1.77, P = 0.18$
|                        |                      |                               |                   |              |              |              | RTLE vs NG: Yates-corrected $\chi^2(1, N = 114) = 10.14, P = 0.001$
|                        |                      |                               |                   |              |              |              | LATL vs NG: Yates-corrected $\chi^2(1, N = 107) = 17.42, P < 0.001$
|                        |                      |                               |                   |              |              |              | RATL vs NG: Yates-corrected $\chi^2(1, N = 106) = 0.07, P = 0.79$
|                        |                      |                               |                   |              |              |              | LTLE vs LATL: Yates-corrected $\chi^2(1, N = 37) = 3.81, P = 0.05$
|                        |                      |                               |                   |              |              |              | Main effect (NG vs overall TLE/ATL): Odds ratio = 0.44
|                        |                      |                               |                   |              |              |              | Yates-corrected $\chi^2(1, N = 185) = 0.37, P = 0.54$
|                        |                      |                               |                   |              |              |              | Pairwise subgroup comparisons:
|                        |                      |                               |                   |              |              |              | NA
|                        |                      |                               |                   |              |              |              | Main effect (NG vs overall TLE/ATL): Odds ratio = 1.68
|                        |                      |                               |                   |              |              |              | Yates-corrected $\chi^2(1, N = 210) = 0.49, P = 0.48$
|                        |                      |                               |                   |              |              |              | Pairwise subgroup comparisons:
|                        |                      |                               |                   |              |              |              | NA
|                        |                      |                               |                   |              |              |              | Main effect (NG vs overall TLE/ATL): Odds ratio = 6.68
|                        |                      |                               |                   |              |              |              | Yates-corrected $\chi^2(1, N = 185) = 14.71, P < 0.001$
|                        |                      |                               |                   |              |              |              | Pairwise subgroup comparisons:
|                        |                      |                               |                   |              |              |              | LTLE vs NG: Yates-corrected $\chi^2(1, N = 122) = 2.27, P = 0.13$
|                        |                      |                               |                   |              |              |              | RTLE vs NG: Yates-corrected $\chi^2(1, N = 124) = 6.14, P = 0.01$
|                        |                      |                               |                   |              |              |              | LATL vs NG: Yates-corrected $\chi^2(1, N = 117) = 24.81, P < 0.001$
|                        |                      |                               |                   |              |              |              | RATL vs NG: Yates-corrected $\chi^2(1, N = 116) = 3.68, P < 0.001$
|                        |                      |                               |                   |              |              |              | LTLE vs LATL: Yates-corrected $\chi^2(1, N = 42) = 3.90, P = 0.05$
|                        |                      |                               |                   |              |              |              | Main effect (NG vs overall TLE/ATL): Odds ratio = 3.73
|                        |                      |                               |                   |              |              |              | Yates-corrected $\chi^2(1, N = 210) = 7.18, P = 0.007$
|                        |                      |                               |                   |              |              |              | Pairwise subgroup comparisons:
|                        |                      |                               |                   |              |              |              | LTLE vs NG: Yates-corrected $\chi^2(1, N = 147) = 2.09, P = 0.15$
|                        |                      |                               |                   |              |              |              | RTLE vs NG: Yates-corrected $\chi^2(1, N = 149) = 1.64, P = 0.05$
|                        |                      |                               |                   |              |              |              | LATL vs NG: Yates-corrected $\chi^2(1, N = 142) = 3.89, P = 0.05$
|                        |                      |                               |                   |              |              |              | RATL vs NG: Yates-corrected $\chi^2(1, N = 141) = 1.25, P = 0.26$

Note: LTLE, left temporal epilepsy; RTLE, right temporal epilepsy; LATL, left anterior lobectomy; RATL, right anterior lobectomy; NG, normative group; *main effect = TLE/ATL vs NG, P ≤ 0.05. *Difference between two clinical subgroups, P ≤ 0.05; post-hoc pairwise comparisons were carried out only when the main effect was significant.

Interactions. Sarcasm scores (which are displayed in Figure 1) differed significantly between groups [KW(4) = 17.98, P = 0.001]. All patient groups performed more poorly than controls (LATL: Conover–Inman = 4.26, P < 0.001; RATL: Conover–Inman = 3.06, P = 0.003; LTLE: Conover–Inman = 2.68, P = 0.009; RTLE: Conover–Inman = 3.78, P < 0.001), and the LATL group performed more poorly than the LTLE group [Conover–Inman = 1.95, P < 0.05]. In both Sarcasm-Part 2 and Part 3 (see Table 2), patients showed more frequent deficits than the normative group. Specifically, relative to controls, deficits were more frequent in RTLE and LATL. RATL also had more frequent deficits in sarcasm-Part 2, and there was no difference between the LTLE and normative groups. Strikingly, 50% of the LATL participants were impaired on sarcasm-Part 2, which was significantly greater than the frequency of deficits noted in the LTLE group, suggesting that left surgery may be associated with worsening of social inference deficits.

Relationship with individual characteristics

FSIQ correlated positively with emotion recognition (r = 0.34, P = 0.001) and with comprehension of deceitful (r = 0.39, P < 0.001) and sarcastic exchanges (r = 0.55, P < 0.001). Age showed a modest negative correlation with sarcasm comprehension (r = -0.22, P = 0.04) and emotion identification (r = -0.20, P = 0.06), but not with deceitful items (r = -0.07, P = 0.55). There were no differences between male and female patients on any measures [emotion: KW(1) = 3.08, P = 0.08; deceit: KW(1) = 2.46, P = 0.12; sarcasm: KW(1) = 1.31, P = 0.25].

Relationship with disease-related characteristics

To investigate the association between pre-operative lesions and social inference deficits, patients (TLE + ATL) were divided into three groups based on pre-operative MRI findings: normal, MTS and non-MTS lesions. MRI findings were not related to emotion identification [KW(2) = 3.36, P = 0.19], but were related to performance on deceitful [KW(2) = 17.80, P < 0.001] and sarcastic items [KW(2) = 13.16, P = 0.001]. Performance of individuals with a normal MRI and those with non-MTS lesions did not differ [deceit: Conover–Inman = 0.13, P = 0.89; sarcasm: Conover–Inman = 1.65, P = 0.10], but individuals with MTS performed more poorly than the other two groups (normal MRI: deceitful: Conover–Inman = 3.97, P < 0.001; sarcasm: Conover–Inman = 2.21, P = 0.03; non-MTS lesions: deceitful: Conover–Inman = 3.50, P = 0.001; sarcasm: Conover–Inman = 3.71, P < 0.001).

These results suggest that MTS is a significant risk factor for developing social inference difficulties. Given that the majority of individuals in our LATL group had MTS pre-operatively, it is unclear whether their widespread social inference difficulties were attributed to left MTS only and whether these difficulties were exacerbated by surgery. To address this issue, we compared individuals with LTLE and MTS (n = 7) with individuals with LATL who had pre-surgical MTS (n = 15), and observed poorer performance in the post-surgical
reverse relationships were true for epilepsy duration (emotion: group for emotion identification \[\text{KW}(1) = 3.17, \text{one-tailed } P = 0.04\] and sarcasm \[\text{KW}(1) = 2.76, \text{one-tailed } P = 0.04\], but not for deceit \[\text{KW}(1) = 1.30, \text{one-tailed } P = 0.13\].

As for other disease-related variables, age of onset correlated with emotion identification (\(r_s = -0.04, P = 0.72\)), and deceit (\(r_s = 0.38, P < 0.001\)) and sarcasm comprehension (\(r_s = 0.21, P = 0.05\), and the reverse relationships were true for epilepsy duration (emotion: \(r_s = -0.11, P = 0.33\); deceitful: \(r_s = -0.38, P < 0.001\); sarcasm: \(r_s = -0.38, P < 0.001\)). These variables are not independent; as expected, duration and age of onset are correlated (\(r_s = -0.66, P < 0.001\)), and patients with MTS had longer epilepsy duration \([\text{KW}(1) = 17.03, P < 0.001]\) and earlier epilepsy onset \([\text{KW}(1) = 9.60, P = 0.002]\) than other patients. To address the relative unique contribution of the various variables to social inference abilities, we ran separate stepwise multiple regression models for each social inference measure reduced in TLE and ATL (emotion, deceit and sarcasm) including variables that showed a relationship with at least one of these measures (age, FSIQ, age of onset, duration and MTS). Of these variables, age (\(\beta = -0.24, P = 0.01\)) and FSIQ (\(\beta = 0.38, P < 0.001\)) explained 21.9\% \((P < 0.001)\) of the variance of the emotion recognition measure; FSIQ (\(\beta = 0.29, P = 0.004\)), age of onset (\(\beta = 0.23, P = 0.02\)) and the presence of MTS (\(\beta = -0.23, P = 0.02\)) accounted for 29.3\% \((P < 0.001)\) of the variance for deceit comprehension, and FSIQ (\(\beta = 0.43, P < 0.001\)), age (\(\beta = -0.22, P = 0.02\)) and the presence of MTS (\(\beta = -0.18, P = 0.06\)) explained 34.0\% \((P < 0.001)\) of the variance for sarcasm comprehension. Importantly, duration had no significant unique contribution to any social inference measure once MTS and age of onset were modelled.

Summary of behavioural findings

Our data suggest that early MTL damage may be especially detrimental to the acquisition of social inference abilities, and that more extensive anterior temporal damage in adulthood (left surgery), may exacerbate social inference deficits. However, based on recent findings showing additional structural changes outside the temporal lobes in TLE (Keller and Roberts, 2008; Li et al., 2012), it is unclear whether social inference deficits in TLE result from disruption in the peak location of MTS (hippocampus) or in other brain regions. Analyses of the relationship between grey matter integrity and social inference in TLE can help address this question.

Grey matter volume and social inference in TLE

Social inference regions. We used a global-null conjunction analysis of GM volume in a GLM model in which we included three factors for social inference abilities (emotion, sarcasm and deceit scores), four covariates including laterality of the epilepsy focus and variables showing a unique contribution to behaviour in the previous regression analyses (age, FSIQ and age of onset). This analysis revealed a jointly consistent positive relationship between social inference and GM volume in the left hippocampus, which is consistent with the effects of MTS in our previous behavioural analyses. As shown in Figure 2, the relationship between the left hippocampus and social inference was driven primarily by the LTLE group, as there was a significant correlation between social inference performance and signal change in the LTLE, but not in the RTLE group. Social inference was also associated with volumes in extra-temporal regions including the left thalamus, left superior frontal, left occipital and right precentral/postcentral regions (Table 3).

Atrophy and specific social inference task

Using the same GLM model, we investigated main effects by examining regions uniquely correlated with different types of social exchanges. Emotion recognition correlated with volumes in right parietal and left inferior frontal regions (Table 3). However, no relationship specific to emotion emerged in temporal lobe regions, including the amygdala. No regions were uniquely related to comprehension of deceitful exchange. Volumes in the right posterior temporal lobe and the left anterior temporal neocortex (Figure 2 and Table 3) correlated positively with performance on sarcasm comprehension. The relationship between sarcasm comprehension and left anterior neocortical volume...
in TLE provides converging evidence of this structure’s importance in combination with our previous observation of poor social inference skills in individuals with left ATL.

**DISCUSSION**

In keeping with previous findings (Meletti et al., 2003, 2009; Schacher et al., 2006; Hibbel et al., 2008; Giovagnoli et al., 2011; Broicher et al., 2012; Li et al., 2013), we observed social inference deficits in individuals with TLE and ATL. The severity of these impairments is partly related to the presence of MTS and to early age of seizure onset as well as to age and intellectual abilities. Deficits were more widespread and severe in the LATL group relative to other TLE groups, even LTLE patients with MTS, indicating that the structural integrity of the left anterior temporal cortex is key to social inference abilities. This conclusion is supported by voxel-based morphometry (VBM) analyses in pre-operative groups, which demonstrated that grey matter volume in two regions excised in LATL correlated with social inference skills: (i) left hippocampal volume correlated with performance on all social inference tasks, and (ii) left anterior neocortical volume correlated with sarcasm comprehension.

Core roles of these regions in supporting social inference have been proposed. With respect to the hippocampus, specific processes in the service of memory operations that have been attributed to this region can be readily extended to the domain of social inference. In particular, several theories emphasize its role in relational binding or in integrating disparate elements of experience in the formation and retrieval of episodic memories (Eichenbaum, 2006; Mayes et al., 2007). These binding processes have been ascribed to hippocampal engagement in such disparate domains as retrieval of autobiographical memories (Moscovitch et al., 2005; Cabera and St Jacques, 2007), future simulation (Eichenbaum and Fortin, 2009; Addis and Schacter, 2011), scene construction (Hassabis and Maguire, 2007), complex perception and working memory (Yonelinas, 2013), open-ended problem-solving (Sheldon et al., 2011) and language processing (Duff and Brown-Schmidt, 2012). Of note, individuals with unilateral TLE have been shown in our lab and others to have deficits in several of these domains. In the current study, hippocampal involvement was observed across all social inference tasks, each of which requires the integration of multiple elements (dialogue, facial expressions, voice prosody and contextual information) to perform accurately.

The anterior temporal neocortex has been hypothesized to store and process social information, in addition to general semantic knowledge. The major theories regarding the role of this region in social inference posit that it is a ‘convergence zone’ in which current complex perceptual inputs and emotional responses are combined with prior social knowledge (e.g. traits, scripts, emotional tone) to enable accurate interpretation of life situations (Frith and Frith, 2006; Olson et al., 2013). These views account for the consistent engagement of this region in complex social inference tasks that use socially relevant, dynamic and contextually rich narratives such as video clips, stories and animated cartoons (Carrington and Bailey, 2009). Our behavioural and VBM results show that the anterior temporal neocortex is particularly important in sarcasm comprehension, which is also consistent with findings in FTD and other neurological populations (Kipps et al., 2009; Rankin et al., 2009). The major difference between sarcasm and other social inference conditions is the requirement that knowledge about paralinguistic cues (e.g. voice tone, facial expression) and of contextually appropriate behaviour be accessed and applied to interpret the situation accurately, which may place higher demands on convergence of current and stored information.

Although our results highlight the importance of the medial and anterior temporal cortices to social inference, our VBM analysis identified additional brain regions of importance, including posterior temporal, frontal, parietal, occipital and subcortical regions. The integrity of these structures has been shown to be compromised in TLE with MTS. A recent meta-analysis demonstrated that up to 40% of existing VBM studies of TLE showed decreased volume in the thalamus, parietal and superior/dorsal frontal regions and 25% of studies reported volume reductions in the inferior temporal lobe (Keller and Roberts, 2008). Furthermore, activation in many of these regions, including frontal (BA6, 10 and 45), inferior temporal (BA 37) and occipital (BA19) cortex has been reported in social inference fMRI studies in healthy individuals (Carrington and Bailey, 2009; Spreng et al., 2009).

However, the ventromedial prefrontal cortex (VMPFC), which has been shown to be necessary in supporting social inference in focal lesion studies (Stone et al., 1998; Stuss et al., 2001; Stuss and Anderson, 2004; Shamay-Tsoory et al., 2005, 2006; Lee et al., 2010), and the superior temporal sulcus, which is reliably reported in imaging studies (Frith and Frith, 2006; Carrington and Bailey, 2009; Spreng et al., 2009), did not emerge from our analyses. This negative finding may be due to the nature or severity of the damage specific to temporal lobe epilepsy rather than to the specific tasks used. Although studies in FTD and other neurodegenerative diseases (Kipps et al., 2009; Rankin et al., 2009) also failed to identify a relationship between grey matter volume in these regions and sarcasm comprehension, there is considerable evidence for VMPFC involvement in deceit and sarcasm comprehension tasks. Specifically, patients with large VMPFC lesions clearly demonstrate impairments on such tasks (Shamay-Tsoory et al., 2005, 2006), and VMPFC activation is reported in fMRI studies using sarcasm tasks in healthy individuals (Shibata 2010; Spotorno, 2012) and is reduced in other clinical populations with impaired social inference abilities such as autism (Wang et al., 2006) and schizophrenia (Rapp et al., 2013). Functional alterations resulting from structural damage elsewhere may nonetheless be present in these regions in TLE, and could potentially be observed using functional neuroimaging techniques in future studies. This account is particularly likely for the medial prefrontal cortex which receives input from the anterior temporal neocortex via the uncinate fasciculus, a white matter tract that has been shown to have reduced structural integrity in MTS (Scanlon et al., 2013) and is severed in temporal lobectomy.

As our data show that social inference impairments are common in every TLE/ATL patient group, it is important to achieve a better understanding of these difficulties and their possible exacerbation with surgical treatment in order to guide potential rehabilitation efforts. Although our findings suggesting poorer social inference following left temporal surgery must be interpreted with caution given
Social inference in epilepsy

the small sample that contributed to the current analysis, our results nonetheless warrant a deeper investigation of the impact of surgery laterality in a longitudinal design. Future studies should also investigate the key components of social inference deficits in TLE and ATL, for instance, to confirm whether the same relational binding processes ascribed to the hippocampus do indeed underlie the various difficulties observed in social inference, autobiographical memory, prospection, working memory, etc., and thus, provide a unified account of cognitive deficits in TLE. The relationship between social cognitive skills involved in interactive social behaviour and psychosocial issues should also be explored. As demonstrated in a recent study, deficits in social inference have a significant impact on overall quality of life and coping ability, and may contribute to psychosocial problems in TLE (Giovagnoli et al., 2013). As such, further systematic investigation of these impairments is crucial given the profound impact they can have on the life of individuals with epilepsy.

REFERENCES


