Available online at www.sciencedirect.com
SCIENCE DIRECT®

Brain Research Reviews 43 (2003) 29–40

BRAIN RESEARCH REVIEWS
www.elsevier.com/locate/brainresrev

Review

A neurobiological mapping of theory of mind

Ahmad Abu-Akel*

Los Angeles, CA 90024, USA

Accepted 21 April 2003

Abstract

This paper attempts, based on a review of a wide range of clinical, biobehavioral and neuroanatomical studies, to account for the various theory of mind impairments observed in psychiatric and developmental disorders in a single neurobiological model. The proposed model is composed of a representational component subserved by posterior brain regions (temporal and parietal) and an application/execution component subserved by prefrontal regions. Information processed in posterior regions is relayed through a limbic–paralimbic system, which is essential for the implementation of theory of mind processes. In addition to its clinical implications, the proposed model accounts for (1) the ability to mentalize about both the self and others, (2) the nature of the anatomic connections of the various brain regions and their functional correlates, and (3) theories pertaining to the inferencing mechanisms used during mental representation/attribution.

Theme: Neural basis of behavior

Topic: Cognition

Keywords: Autism; Neurobiology; Schizophrenia; Theory of mind; Mental representation/attribution; Simulation/theory theory

Contents

1. Introduction 29
2. Theory of mind impairments 30
3. Neuroanatomy of theory of mind 31
 3.1. Representing mental states of self and other 31
  3.1.1. Brain regions specific for the representation of self-mental states 31
  3.1.2. Brain regions specific for the representation of others' mental states 32
  3.1.3. Brain regions common to the representation of self and others' mental states 33
 3.2. The integration: anatomical and functional connections 34
4. Theoretical and clinical implications 36
 4.1. Theoretical implications 36
 4.2. Clinical implications 36
5. Closing remarks 37
References 37

1. Introduction

Theory of mind (ToM), sometimes used interchangeably with other terms such as *mentalizing capacity*, is the ability to represent one's own or another's mental states such as intentions, beliefs, wants, desires, and knowledge. This ability is acquired by children around 4 years of age and continues to develop until around 11 years of age [7]. While for some thought processes mentalizing is used for introspection, it is mainly used to socialize with others [20,22]. This ability is believed to be an outgrowth of social intelligence. Social intelligences such as the ability to detect another agent's goal (such as in goal-directed

*Tel.: +1-310-824-7008; fax: +1-310-205-0310.
E-mail address: abuakel@hotmail.com (A. Abu-Akel).

0165-0173/03/$ – see front matter
doi:10.1016/S0165-0173(03)00190-5

actions), anticipate a course of action, distinguish between actions of the self and of others, learn through imitation, perceive emotions, share attention, and the ability for introspection and self-agency, are precursors or 'protoforms' of theory of mind knowledge [20,32,38].

Two accounts have been proposed to explain the inferencing mechanism underlying ToM: 'Theory Theory' (TT) and 'Simulation Theory' (ST). According to TT our knowledge of other minds is embodied in an explicit symbolic theory along the lines of folk psychology, with axioms and rules of inference, from which we may deduce what others know and want [47]. In contrast, proponents of ST contend that we mentally simulate others' thought processes and feelings, using our own mental state as a model of theirs [53]. To date, data accumulated to decide between these two competing theories are inconclusive and this issue remains an open empirical question.

Several neurobiological models have been proposed as bases for ToM [9,19,40]. These models and many others have invoked structures in posterior and anterior regions of the brain most consistently being the superior temporal sulcus and the medial prefrontal cortex, respectively. Limbic–paralimbic structures have also been suggested as part of the ToM circuitry most notably being the orbitofrontal cortex and the amygdala. For example, Brothers [19] provided a model for social cognition (an ability that encompasses ToM abilities) that consists of a three-node circuit connecting the orbitofrontal cortex, the superior temporal sulcus, and the amygdala. According to Brothers, if this circuit is broken at any point, autism can be produced, a disorder associated with significant deficits in theory of mind abilities. More recently, in their synthesis of the biological basis of social interaction, the Friths [40] suggested a theory of mind network that emphasizes the superior temporal sulcus, the medial prefrontal cortex (including the anterior cingulate cortex), and to some extent the amygdala.

In light of recent theoretical development and empirical findings, the present review is an attempt to provide a more elaborate model that considers (1) mental attribution to self versus mental attribution to others, (2) the pattern of connectivity between various brain regions that have been implicated in mentalizing abilities, (3) the inferencing mechanism used during mental states attribution (i.e. whether we infer mental states along the lines of ST or TT), and (4) the various ToM impairments observed in developmental and psychiatric disorders. Before embarking on this task a general description of the variety of impairments associated with this ability are summarized first.

2. Theory of mind impairments

Most studies evaluating ToM abilities in psychiatric and developmental disorders have adopted a dichotomous approach, wherein failure on ToM tasks indicates an absence of this ability.[1] That is, such research takes as its paradigm the presence or absence of conceptual abilities needed to mentalize about self and others, and the processes required for the successful application of these conceptual abilities. Within this paradigm two main lines of thinking have emerged. The first takes ToM impairment as a deficit in conceptualization and representational abilities. In this view, an individual lacking ToM is thought to be unable to represent mental states of others or of oneself [15,39,96]. The second argues that the deficit is one of application and performance. Here the individual demonstrates awareness of another's mental activity but

[1]The classic task used in testing for theory of mind is the 'unexpected transfer' task [103], which examines whether or not an individual can attribute a false belief to another. In this task the individual witnesses a scenario where a child hides an object in location A and leaves the room. In the absence of the child this object is displaced and hidden again in location B. The individual is then asked where the returning child will look for the object. If the individual answers that the child will look in location A, the individual is assumed to have an operative theory of mind. If the individual answers that the child will look in location B, it is assumed that the individual lacks an operative theory of mind.

[2]Similarities in ToM abilities between negative symptom schizophrenia and the autistic spectrum is plausible since there is evidence that similar to patients with autism, patients with negative symptom schizophrenia have abnormal social development in early childhood [34] during which period ToM abilities develop [67]. This is all the more plausible given that both populations exhibit similar communication problems as can be seen from the following speech samples where the second speech sample is similar to the first which U. Frith describes as a very limited conversation in its communicative potential and with no drive for continuity.

(1) A conversation with an able autistic patient (Ref. [39], pp. 118–119)

UF . . . you live in that lovely flat upstairs?
R Yes–suh [Ruth always emphasized the final consonant]
UF Is that really good?
R It is.
UF Do you do some cooking there?
R Yes. I do.
UF What kind of things do you cook?
R Anything.
UF Really. What is your favorite food?
R Fish fingers.

(2) A conversation with an individual with negative symptom schizophrenia with marked poverty of speech (Ref. [35], p. 44)

E How're you doing generally at the moment, Mr. D?
D All right.
E You're OK. How're . . . How've you been feeling in your spirits this past week?
D Not so bad.
E You're feeling all right. Do you have any spells of feeling sad or miserable?
D No.
E No? Nothing like that? That's good. Now tell me, Mr. D, do you have any special ideas about life in general?
D [Shakes head]
E No? Just ordinary ideas like the rest of us. No. Have you ever thought that you were a special person in any way?
D [Shakes head]
E Do you feel people stare at you and talk about you in some way?
D [Shakes head]

fails to apply this knowledge (attribute mental states) due to processing constraints [73,82,89]. The conceptual deficit view is generally ascribed to individuals with autism, while the application deficit view is generally ascribed to individuals with Aperger's syndrome and, according to some accounts, to negative symptom schizophrenia [17].[2]

Recent empirical findings and developments in the theoretical understanding of ToM suggest that there are additional varieties of ToM impairment that can be observed in developmental and psychiatric disorders. Of interest are two groups of patients with schizophrenia. The first group consists of patients with passivity phenomena who, despite their impaired ability to represent their own mental states such as their intention to act [28,70], seem to have no deficits in the ability to represent mental states in others [37,78].

The second group consists of patients with positive symptom schizophrenia, and in particular those with paranoid and delusional schizophrenia. It has been hypothesized that the ToM failure of these patients [26,37,91] is a consequence not so much from a *deficit* in the representation of mental states, as reported for example by Sarfati et al. [92], but rather from a *hypertrophy* of it, i.e. a hyper-theory-of-mind that often leads such patients to overattribute knowledge and other mental states to their interlocutors [1]. According to a different account, this abnormal mentalizing may be due to abnormal use of episodic memories used to help make predictions based on what we already know [26]. In either case, egocentrism or an autistic-like lack of consideration of the mental states of others clearly cannot account for the mentalizing behavior of this group of patients. It should be pointed out that while the ToM impairment in this group of patients is one of application, it is fundamentally different from that of patients with Asperger's syndrome and negative symptom schizophrenia—whereas patients with positive symptom schizophrenia attribute mental states abnormally, patients with Asperger's syndrome and negative symptom schizophrenia do not manifest their knowledge of mental states in real life situations [17,82].

By considering ToM behavior in various developmental and psychiatric populations, we are now able to reframe ToM impairments on a continuum where a specific impairment can be characterized as having: (1) no representational/conceptual understanding of mental states (e.g. autism), (2) representational understanding of mental states, but a deficit in the ability to apply/manifest this understanding (e.g. Asperger's syndrome and negative symptom schizophrenia), (3) representational understanding of mental states, but abnormal attribution/application of these mental states (e.g. delusional and paranoid schizophrenia), and (4) intact representational understanding of the mind of others, but impaired self (e.g. schizophrenic patients with passivity phenomena).

These clinical observations suggest that ToM impairments are either of competence or of performance. Such distinction can allow us to relate to the conceptual/application distinction across patient groups, as well as to those who apply them abnormally. Moreover, ToM impairments can be in either the ability to represent self and/or others' mental states. Such distinction can provide vital information for clarifying debates regarding which inferencing mechanism (ST or TT) is used during mind-reading. These performance–competence and self–other levels of distinctions should, therefore, be expressed in any neurobiological modeling of theory of mind. In other words, a given neurobiological model of ToM should express (shared or differential) networks subserving the ability to represent self and others' mental states, and similarly those pertaining to competence and performance.

The current paper is thus an attempt to provide a neuroanatomical circuit that accounts for different types of ToM impairments found in a variety of psychiatric disorders. While these disorders, specifically autism and schizophrenia, have different developmental histories they are similar, as this review will demonstrate, in that they impact the same mechanism that allows for the ability to mentalize [35].

3. Neuroanatomy of theory of mind

There is a plethora of studies that have investigated the involvement of brain areas during simple and complex forms of mentalization. However, many of these studies are localization studies and do not discuss the role of a particular pattern of connectivity between these areas. In proposing a functional neuroanatomy of ToM, the present paper, firstly, provides a review for the involvement of various brain regions in ToM processing. Secondly, it presents an integration of these seemingly independent areas into a functionally interconnected circuit that can explain the various manifestations of ToM impairments summarized above.

3.1. Representing mental states of self and other

The question of whether representing self or others' mental states employs the same or different brain regions has only been recently explored [102]. Studies concerned with the underlying neurophysiology of representing mental states provide data on the involvement of various brain regions that can be classified into three main groups: brain areas solely involved in the representation of self mental states, brain areas solely involved in the representation of mental states of others, and brain areas that are common to both.

3.1.1. Brain regions specific for the representation of self-mental states

Several studies suggest that areas in the right posterior parietal system, specifically the inferior parietal lobule (IPL), are responsible for representing one's own mental states. Evidence for this is provided in studies investigating the underlying neurophysiology during the attribution of mental states to self and the execution of self-generated acts. Vogeley et al. [102] provided the first direct evidence for the involvement of the posterior parietal region in the representation of self-mental states. In employing a two-way factorial design of stimulus material that involved ToM tasks of self and others, significant activation in the right temporoparietal regions (i.e. IPL/parietal operculum) was observed during tasks requiring the representation of one's own mental states only. (The authors in this study seem to use the term temporoparietal interchangeably with the IPL and the parietal operculum.) Significant activations were also observed in medial aspects of the superior parietal region. These regions were not active when subjects were required to infer the mental states of others. This suggests that representing self-mental states invokes a neural mechanism that is independent from that involved in representing others' mental states.

The specific involvement of the right posterior parietal region in representing self-mental states can also be inferred from studies investigating the underlying neurophysiology of self-generated acts. In a positron emission tomography (PET) study, Spence et al. [97] observed abnormal hyperactivation in the right IPL region during the execution of voluntary movements in schizophrenic patients experiencing passivity phenomena (i.e. experiencing beliefs that their thoughts and actions are under alien control). Furthermore, Iacaboni et al. [60] reported that the right parietal operculum was active during the execution of the act (i.e. the imitation of specific finger movements), but not when observing the same movement being executed by another individual. The activation of the right parietal operculum during execution only, the authors argue, suggests its involvement in the preservation of body identity during imitation, i.e. an awareness that 'it is my body that is moving' (p. 2528). The plausibility of the specialization of the right IPL in the representation of internally/self-generated acts finds support in studies reporting a specific involvement of this area in distinguishing self-produced actions from those generated by others [88], and studies showing that lesions to this area lead to the loss of corporeal awareness [14].

It should be noted that brain regions involved in representing self-agency or distinguishing self-produced actions from those generated by others can be considered part of the circuitry for representing self-mental states. These capacities are considered protoforms upon which ToM capacities may have evolved [16,38]. A corollary issue is that self-mental representation may be built on networks subserving body representation. Evidence illustrating the connection between the person's belief system and body representation is found in patients with neglected or disowned body parts who often use confabulatory explanations (reflecting a distorted belief system) to justify their existence [83,84].

Abnormalities in the parietal lobule have been reported for both patients with autism and schizophrenia. For example, in autism the parietal cortex is abnormally large compared to healthy controls [25], and has metabolic abnormalities [30]. Similarly, numerous structural MRI studies report abnormalities in the parietal cortex in schizophrenics, and particularly in the inferior parietal lobule (see Ref. [94] for review). These neuroanatomical abnormalities are congruent with the fact that these disorders are characterized by the inability to represent self-mental states.

3.1.2. Brain regions specific for the representation of others' mental states

Evidence from several studies in both primates and humans suggests that the superior temporal sulcus (STS) is specialized in the representation of the mental states of others. This assumption is motivated by evidence showing that cells in this region respond to what is generated by another and not by one's self. In primates, it has been shown that the STS responds selectively to sounds, hand and face movements generated by others, but not to similar movements of inanimate objects [72], nor to such sounds or movements generated by one's self [55]. More recently, Jellema et al. [63] provided evidence that two distinct cell populations in the STS (STSa, sometimes called STPa) in the macaque monkey have specific functions where one cell population is responsible for detection of attention of others and another is responsible for the detection of movements in others. The authors conclude that the combined effect of these functions allows for the detection of intentional action in others. Similar results have been obtained for humans [49,81]. For example, Puce et al. reported that the posterior part of STS is involved in the perception of eye gaze and mouth movement. This region, they noted, may be functionally related to adjacent superior temporal regions, which appear to be involved in the perception of others' hand and body movements. Moreover, the STS responds selectively to observation of goal-directed actions, such as movements that take objects as their intention (e.g. reaching, grasping, holding, tearing), but not to movements lacking such intentions [77].

Collectively, these abilities are especially important for inferring mental states in others and more generally for social interaction. For example, the representation of actions, their goals, as well as eye-gaze, which is used to control turn-taking in dialogues [46], and is involved in the deployment of mutual orientation and joint attention [9], are prerequisite for the development of mentalizing abilities both phylogenetically and ontogenetically. Moreover, recent neuroimaging data collected during the execution of ToM tasks by normals and patients of the autistic spectrum provide direct evidence for the role of the STS region in the representation of others' mental states [8,23,42]. Studies employing tasks requiring self-mental representation did not report activity in this region [51,68].

3.1.3. Brain regions common to the representation of self and others' mental states

Most researchers interested in the delineation of the neurobiology of ToM have designed studies that investigate brain activity during the attribution of mental states to others. Consequently, these studies have no say on whether involved brain regions are also involved during mental attribution to self. I will argue that regions which include the amygdala, the anterior cingulate gyrus (ACG), the orbitofrontal cortex (OFC), the ventral and dorsal medial prefrontal cortex (ventral/dorsal MPFC), and the inferolateral frontal cortex (ILFC), reported to be involved during the attribution of mental states to others, are also involved during the attribution of mental states to self. These regions can be classified into two main groups. The first group consists of structures in the limbic–paralimbic system and includes the amygdala, OFC, ventral MPFC, and ACG. The second group consists of structures in the prefrontal cortex (PFC) and includes the dorsal MPFC and ILFC.

Within the limbic–paralimbic system, amygdala damage (structural or functional) can impair ToM abilities, or abilities that are necessary for its integrity. For example, Kling and Brothers [64] showed that bilateral removal of the amygdala results in an extraordinary loss of social and affective behavior in monkeys. In humans, Allman and Brothers [4] showed that the amygdala is sensitive to direction of gaze and the expression of emotion in others' faces. Amygdala dysfunction in patients with autism [57] and schizophrenia [85,93] has been tied to their inability to represent their own emotion states or those of others [18]. Moreover, recent neuroimaging and lesion studies provide direct evidence for its involvement in the representation of mental states [8,31]. Baron-Cohen et al. [8] reported that compared to normals, patients with autism or Asperger's syndrome show no activation to the amygdala when making mentalistic inferences from the eyes, and Fine et al. [31] reported that an early left damage to the amygdala (particularly the basal nuclei) severely impairs the ability to represent mental states.

Next are the OFC and the ventral MPFC. These regions, though anatomically dissociable at some level, are presented here as a unitary complex since they sustain similar functions that pertain to the regulation and representation of socioemotional states [13,29]. The specific involvement of these two structures in mentalizing has been reported in several studies. In a single photon emission computerized tomography (SPECT) study, Baron-Cohen et al. [10] reported activation of the right OFC during recognition of mental state terms. In a later study, Stone et al. [98] reported that patients with a bilateral OFC lesion performed poorly on a story telling task requiring an understanding of both pragmatically appropriate social behavior (faux pas) and the effect of that behavior on the mental states of others. With respect to the ventral MPFC, Stuss et al. [100] investigated the role of the frontal lobes in theory of mind in 32 patients with limited focal frontal and nonfrontal lesions and reported that lesions to either the left or the right (though more to the right) ventral MPFC impair ToM abilities and particularly the ability to detect deception performed by others. The involvement of ventral MPFC in mentalizing is corroborated by Ohnishi et al. [71] who observed metabolic abnormalities in the ventral MPFC of children with autism. The role of these structures in the attribution of mental states to self is less clear. However, evidence for the involvement of the OFC and/or the ventral MPFC in processing one's self-mental states can be inferred from the socially inappropriate behaviors of patients with damage to these structures. These patients' deficient sense of self as can be inferred from the fact that they exhibit deficits in decision making including those with direct consequences to themselves, a lack of consideration for their personal habits, apathy towards criticism, and a lack of anxiety or concern for self [5,13,27]. Moreover, Gusnard et al. [51] showed that the ventral MPFC is active during the performance of a self-referential judgment task of affectively normed pictures. The involvement of these structures in emotional processing also suggests their involvement during the attribution of affective mental states to self.

Still within the limbic–paralimbic system, the anterior cingulate gyrus (ACG) plays a major role in attention and emotional processing (see Ref. [79] for review), and is involved during tasks requiring representation of mental states or related capacities [40,102]. Several neuroimaging studies have demonstrated activation of ACG during tasks requiring the attribution of mental states to others [33,42]. Commensurate with these studies is that patients with autism spectrum disorders have abnormal metabolic rates in the anterior cingulate gyrus [54,71]. Recently, however, Vogeley et al. [102] extended this finding and showed that the ACG is not only involved in the attribution of mental states to others but also to oneself. The involvement of ACG in the attribution of mental states to self can also be inferred from the findings of several other studies [51,95,97]. Shima et al. [95] reported that activity in the posterior cingulate gyrus (but anterior to the motor cingulate gyrus) was detected before the production of self-initiated movements. Moreover, based on PET evidence from schizophrenic patients with passivity phenomena (delusions of alien control) performing voluntary movements [97], and data from an fMRI study involving healthy participants performing self-referential mental activity tasks [50], the ACG appears to be involved in the representation of value attribution to internally/self-generated acts or thoughts. Taken together these studies strongly speak in favor of the dual involvement of the ACG in representing self and other mental states. Interestingly, a study of a single-unit response in neurosurgical patients revealed that neurons in ACG respond both when the subject is experiencing pain or simply observing another in pain [58].

Collectively, the above reviewed studies addressing these limbic and paralimbic regions, i.e. the amygdala, the ACG and the OFC and the ventral MPFC, strongly suggest that they have a dual role in regulating the social/emotional mental representations of both self and others. The role of this system in regulating social and emotional mental representations is most apparent in patients with schizophrenia and autism. Both autism [11,12] and schizophrenia [101] present severe abnormalities to the limbic system, which in turn significantly disrupt their ability to express and understand emotional and social cues adequately.

With respect to structures of the second group, several neuroimaging studies have consistently provided evidence for the involvement of the dorsal MPFC in mental attribution tasks [23,33,42,45,52,89]. Fletcher et al. [33] showed increased cortical activity predominantly in left dorsal MPFC (Brodmann's areas 8 and 9) of normal adults when presented with a narrative containing reference to mental states as compared to a narrative lacking such reference. Happé et al. [52], using Fletcher et al.'s [33] experimental design, found activation in the anterior cingulate and Brodmann's area 9 (but not 8) in both normals and patients with Asperger's syndrome, but these areas were significantly less active in the patients with Asperger's syndrome. In addition to the involvement of this structure in the attribution of mental states to others, it has been suggested that this region is also involved during the attribution of mental states to self [23,38]. This proposition finds supporting evidence in the study by Gusnard et al. [51] in which they reported that the presence of self-referential mental activity leads to increased activity in the dorsal MPFC from the baseline, and concluded that this region 'might be necessary for spontaneous as well as task-related self-referential or introspectively oriented mental activity' (p. 4263). Additional evidence is also found in studies requiring subjects to monitor their own mental states [65,68].

The ILFC has also been implicated in the neurobiology subserving ToM. In an fMRI study, Russell et al. [90] showed that during a mental attribution task, schizophrenic patients, compared to normal controls, have significantly less activation in the left inferior frontal cortex reaching into the insula (Brodmann's area 44/45/47). Moreover, in the ILFC there is a group of cells called mirror neurons (MNs) [21,60], which are thought to be involved in mental attribution processes [43,86].[3] Mirror neurons have unique properties in that they discharge congruent firing patterns both when the subject is executing specific goal-directed motor activities using facial, arm and hand movements, as well as when the subject passively witnesses those exact same goal-directed activities being performed by another. It is important to note that these neurons will not fire when observing an action in the absence of the acted-upon object. Accordingly, Rizzolatti and Arbib [86] proposed that this state-matching system, which is common to the observer and the observed, may play a role in the detection of mental activity (also see Ref. [59]). Since these neurons make no distinction with respect to agency (self or other) [62], it can be suggested that the ILFC plays a role in attributing mental states to both self and others.

3.2. The integration: anatomical and functional connections

To put it together, the neurobiological substrate subserving the ability to attribute mental states to oneself and to others is comprised of three main components: *posterior regions* which include the IPL and STS, *limbic–paralimbic regions* which include the amygdala, OFC, the ventral MPFC and ACG, and *prefrontal regions* which include the dorsal MPFC and ILFC. Within this network, post-Rolandic regions are reciprocally interconnected with pre-Rolandic regions [74,75]. Similar reciprocal connections exist between the limbic and paralimbic structures and the pre- and post-Rolandic regions implicated here [74,75]. For example, the cingulate gyrus has reciprocal connections with the prefrontal association cortex, the superior temporal sulcus, the posterior parietal cortex, and the amygdala [74,97]. The reciprocal connections among the regions implicated in the model suggest that the network for mental state attribution allows for feedback and recursion in processing. A schematic illustration of the neurobiological network subserving ToM is shown in Fig. 1.

Information in the brain is often fed forward from posterior to anterior regions through limbic and paralimbic structures [74]. For example, in tasks involving interpretation of actions, images are detected first in the STS region, relayed to paralimbic and limbic structures for emotional input, and then fed forward to frontal regions where mirror neuron circuits in the ILFC are activated to reproduce the action [43,63]. Similar to interpretation of actions, it is reasonable to conjecture that mental states of self or others are, respectively, first detected (represented) in the IPL and the STS regions, relayed to paralimbic and limbic structures for emotional input and integration of this emotional input with the incoming parasensory information. At this stage, the information is synthesized for evaluation of relevance or meaning for the individual and relayed to dorsal and lateral regions of the PFC for application and execution decisions.

The reciprocally interconnected limbic–paralimbic and neocortical systems [74] suggest that emotion and cognition can be interacting functions of the brain [65]. In this

[3] Mirror neurons were first observed in monkeys in the ventral premotor cortex (area F5), and there is good evidence for the existence of these neurons in Broca's areas or the frontal operculum in humans [21,60]. Broca's area (part of the ILFC) is the purported homologue of F5 in monkeys and is involved in the control of the oro-laryngeal, oro-facial and brachio-manual movements responsible for the physical production of speech (for discussion of the cytoarchitectonic and functional homologies between area F5 and Broca's area see Ref. [86]).

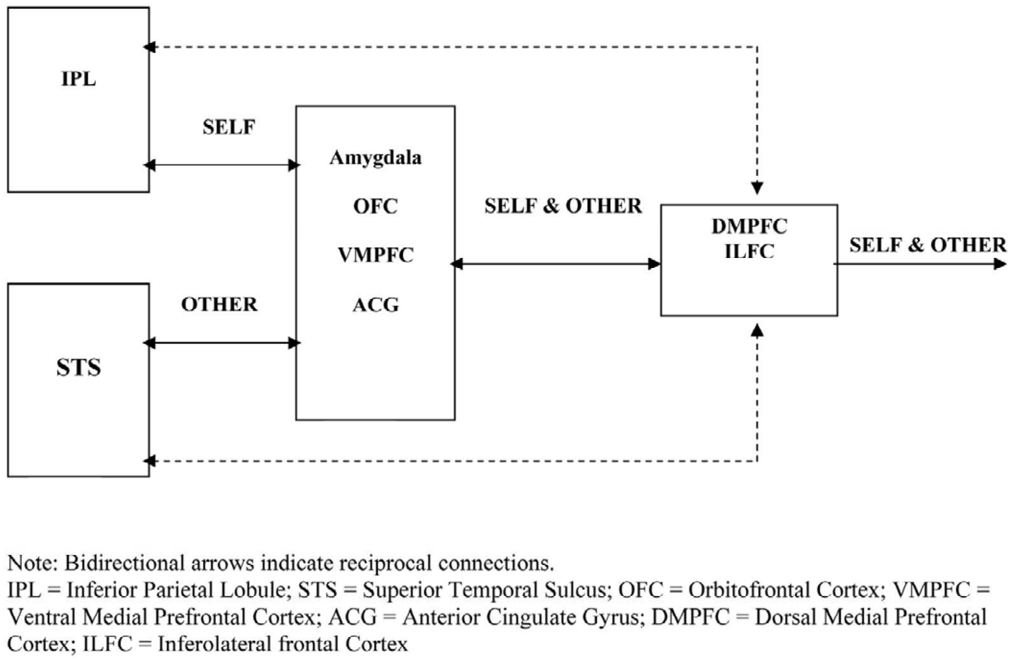

Fig. 1. Functional Neuroanatomy of representing and attributing self- and other mental states.

regard, Le Doux [66] proposes that by way of neural interactions among these various regions, affect and cognition can mutually affect each other. Indeed, incorporating emotional/motivational representations, or as referred to by Damasio 'somatic markers', influences cognitive processes in that they serve as biasing signals in decision-making processes [27], indicators of mental states such as intentions [24], and a support to conscious awareness [50]. Thus the experience of happiness (i.e. a self-mental state), for example, is created when representations of stimuli, self and affect are integrated. This integration is then used for the evaluation of behavioral consequences and the significance of the social situation. I argue that the involvement of the limbic–paralimbic network is not confined to representing affective mental states. There is strong evidence suggesting that damage to these regions also impairs general ToM abilities [8,31,98]. Concisely put, 'you cannot formulate and use adequate 'theories' of your mind and for the mind of others if something like the somatic marker fails you' (Ref. [27], p. 219).

It is suggested that processing of this interaction between cognitive and affective representation is most likely to take place in the OFC/ventral MPFC complex. The ventral MPFC region is viewed as a convergence zone that receives (primarily via the OFC) sensory data of all modalities (e.g. auditory, visual, somatosensory) as well as all kinds of perceptual syntheses of sensory input [28,80], which makes it an ideal infrastructure for synthesizing the diverse range of information needed for a complex mental activity such as mentalizing. At this stage, this highly synthesized information is forwarded via the ventral MPFC to various PFC regions [41,69], including the dorsal MPFC and the ILFC. These prefrontal regions, I argue, are primarily involved in the application process of mentalizing. To substantiate this claim, there is a need to show that the PFC sustains functions necessary for execution and application, or that impairment to the PFC would result in problems associated with application and/or execution of mental states.

Stuss and Benson's [99] model of brain functions suggest that the frontal lobe is involved in four basic functions. The first function is *sequencing* which is the ability to maintain meaningful information units. The second is *drive* that allows for the modulation and inhibition of cerebral activity. The third function is *control* (which includes planning, monitoring and anticipation) of cognitive and noncognitive activities. The fourth function is *self-analysis* that allows for evaluating the meaningfulness of the cognitive process to the individual. These functions of the frontal lobe strongly suggest that the PFC is involved in application and execution processes. Supporting evidence for this is provided by several studies on patients with damage to the PFC including the medial PFC. These studies report that while these patients can state socially appropriate thoughts and actions in laboratory settings (i.e. preserved representational abilities) they fail to apply this knowledge in their daily social interactions (i.e. abnormal application) [6,27]. This is strikingly similar to a study by Bowler [17] who shows that while patients with Asperger's syndrome as well as patients with negative symptom schizophrenia perform well on ToM tasks in a laboratory setting, they fail to demonstrate their knowledge of the other's mind in real life settings. Interestingly, abnormal activity has been observed in the dorsal MPFC (BA 8/9) of Asperger's patients while performing mentalistic tasks [52].

Additional support for the involvement of the PFC in application processes comes from a close examination of the ILFC. The ILFC and its mirror neuron circuit allows for mental state matching between self and other [43,63]. The ILFC also has motoric functions of speech, hand and face movements as well as functional relationships of output mechanism with the premotor cortex (Ref. [86], pp. 190–191). These functional properties of the ILFC suggest its involvement in the application of the attribution of mental states. To summarize, it is argued that (1) the functional properties of the ILFC, (2) the hindering effect of MPFC damage on the ability of patients with Asperger's syndrome and medial prefrontal damage to apply represented mental states and thoughts, and (3) the general prefrontal functions of sequencing, drive, control and self-analysis provide a plausible account for the involvement of the PFC in the application process of already represented mental states.

To conclude, the brain regions presented in this distributed network can be divided in terms of their involvement in the mentalizing process as follows: (1) representation of one's own mental states is subserved by the IPL region, (2) representation of others' mental states is subserved by the STS region, (3) mental states represented in these regions are fed forward to the limbic–paralimbic system for socioemotional regulation and interpretation, and (4) processed information is then projected to the MPFC and ILFC regions for application processing. This neurobiological model has important consequences for clinical and theoretical aspects of mentalizing which is the topic of the next section.

4. Theoretical and clinical implications

The potential success or soundness of a particular model depends, among other things, on its ability to relate to the theoretical/philosophical underpinnings of the phenomenon it observes, as well as its explanatory power of the various phenomenological manifestations of the phenomenon. As it pertains to the proposed model of the neuroanatomy of ToM, I will first address the inferencing mechanism that this model may support. In other words, how does this model account for Simulation Theory (ST) versus Theory Theory (TT)? Secondly, I will discuss the model's potential in predicting the various theory of mind impairments observed in psychiatric and developmental disorders outlined in Section 2.

4.1. Theoretical implications

If we indeed infer mental states via simulation processes along the lines of ST, the simulator needs first to introspect and directly be engaged in a pretense of not only being in the other's situation, but also being the other [48]. On the face of it, the mirror neuron system, described above and which appears to be involved in mind-reading, seems to fit the general framework of this theory [43]. However, there are two main problems with this proposal. Firstly, MNs do not make explicit whether a goal-directed act is self-generated or generated by another agent [62]. Secondly, but a corollary problem, is that Gallese and Goldman's mirror neuron/simulation account [43] would predict that there could not exist a situation wherein a person is capable of successfully attributing mental states to others but not to self since attributing to self, according to ST, is a prerequisite step to gain insight into others' mental states. Strikingly, however, schizophrenic patients with passivity features (e.g. delusions of control) were successful in inferring the intentions of others from indirect speech probes [26], and could correctly answer ToM questions addressing false beliefs in others [37,78]. This is problematic for ST given that these patients also have problems representing their own intentions to act [28,70], and are impaired in discriminating their own hand from an alien one [97], and misattribute internally generated acts to others [61].

From a neurobiological perspective it can be suggested that abnormalities in the IPL region, as found in these patients [97], hinders the ability to represent one's own mental states. This is analogous to cases of schizophrenia where overactivity in the parietal cortex, probably due to a loss of inhibitory control of the frontal cortex, leads to the inability to represent current and predicted states of one's own limbs [36]. The remaining intact regions, it seems, can still allow for the ability to attribute mental states to others through the STS and its connections with limbic, paralimbic and frontal regions. That individuals can attribute mental states to others while self is impaired implies these abilities are dissociable and that mental representation cannot be by way of simulation. It appears then, though by default, that TT is the likely inferencing mechanism by which we arrive at mental states.

4.2. Clinical implications

The clinical implications of the proposed model are quite diverse. In Section 2 I outlined, based on clinical data, the various ToM impairments often observed in psychiatric and developmental disorders. These include (1) an absence of the ability to represent mental states often ascribed to individuals with autism, (2) an absence of the ability to apply mental states often ascribed to individuals with Asperger's syndrome and negative symptom schizophrenia, and 3) an overattribution of mental states often ascribed to individuals with positive symptom schizophrenia of the delusional and paranoid types. The question that arises is which disruptions or lesions to the system produce which ToM impairments.

According to the knowledge we have, it would be impossible to determine the minimum damage needed to result in impairment to ToM abilities given that: (1) individuals with ToM impairments exhibit abnormalities in a number of brain regions, (2) these regions are anatomically interconnected and it is difficult to know the deficits that could result in other regions as a consequence of impairment to some particular region, and (3) there are almost no data on whether ToM processes are generated/modulated by specific cell populations within a given region.

With these limitations in mind, a number of predictions can, nonetheless, be made based on the proposed model. The model predicts that damage to the posterior regions which include the IPL and STS regions could lead to impairment in the ability to *represent* mental states of self and other, respectively. This suggestion is in accordance with recent findings suggesting the involvement of posterior regions in the representation of mental states, rather than frontal regions [6]. Specifically, damage to the IPL region may lead to the loss of the ability to represent mental states to self as is the case with schizophrenic individuals with passivity phenomena. Damage to both the IPL and STS regions may result in impairment to the ability to represent mental states about both self and others as in the extreme case of autism.

The model also predicts that damage to frontal regions is likely to impair *application or attribution* processes of mental states. Such damage can lead to the kind of application deficit found in patients with Asperger's syndrome and negative symptom schizophrenia, or that found in positive symptom schizophrenia. The difference between these types of application deficit is one of severity and is likely related to etiological differences in that positive symptom schizophrenia occurs much later in life[4]—earlier damage results in more severe application deficits, and later damage results in lesser application deficits.

One possible scenario is that application problems can result from dysfunctional inhibitory pathways between the PFC and limbic–paralimbic structures, particularly the amygdala [76,87]. This can result from either an over inhibition by the PFC of incoming information from limbic–paralimbic regions leading to the kind of application deficit in Asperger's syndrome and negative symptom schizophrenia, or loss of inhibition leading to hyper-theory of mind. Loss of inhibition can result from over excitation of the amygdala which could result from either a dysfunction in the inhibitory role of the amygdala in the spontaneous activity of neurons in regions of the PFC [76], or from a dysfunction in the ability of the PFC to recruit inhibitory interneurons of the amygdala (specifically of the basolateral nucleus of the amygdala) to attenuate sensory-driven amygdala-mediated affective responses [87]. The suggestion that damage to the amygdala could lead to some form of hypersociability [3] is in accord with the role of amygdala and its connection with PFC in controlling attribution processes of mental states.

Finally, as reviewed above, there is extensive evidence suggesting that damage to limbic–paralimbic regions severely impairs the ability to mentalize in both autism and schizophrenia. Specifically, dysfunction to this system can result in impairment in the conception of affective mental states. The involvement of the limbic–paralimbic circuit in the representation of affective mental states is in accordance with psychological models considering autism as a disorder of emotional deficit [56], and those considering it as a disorder of empathy [44]. It also supports recent empirical findings suggesting that patients with schizophrenia are also impaired in their ability to empathize [2]. Such deficits could be of application or of competence. An application deficit might result from a dysfunction in the inhibitory pathway of the PFC to the limbic system. On the other hand, a competence deficit might result from severe damage to the limbic system or its connections with posterior association cortices, namely the parietal lobule and the superior temporal sulcus.

[4]This is consistent with the view that, unlike patients with autism and negative symptom schizophrenia (see footnote 2 above), there is no evidence of poor childhood social functioning in patients with paranoid and related positive features. This suggests that ToM impairment in these patients is unlikely to be attributable to early childhood cognitive and/or social development (see Ref. [35]).

5. Closing remarks

In this paper, I have attempted based on a review of a wide range of clinical, biobehavioral, neuroanatomical and neuroimaging studies, to account for the various ToM impairments observed in psychiatric and developmental disorders in a single neurobiological model. The model provides some suggestions and predictions as to the neuroanatomical deficit that might be associated with various ToM impairments. Research under the guidelines of the proposed model can help achieve fuller understanding of pathologies that disrupt the ability to represent and attribute mental states and their neurobiological correlates.

References

[1] A. Abu-Akel, A.L. Bailey, The possibility of different forms of theory of mind impairment in psychiatric and developmental disorders, Psychol. Med. 30 (2000) 735–738, Letter.

[2] A. Abu-Akel, K. Abu-Shua'leh, 'Theory of mind' in violent and nonviolent patients with paranoid schizophrenia, Schizophr. Res., in press.

[3] R. Adolphs, D. Tranel, A.R. Damasio, The human amygdala in social judgment, Nature 393 (1998) 470–474.

[4] J. Allman, L. Brothers, Faces, fear and the amygdala, Nature 372 (1994) 613–614.

[5] S.W. Anderson, A. Bechara, H. Damasio, D. Tranel, A.R. Damasio, Impairment of social and moral behavior related to early damage in human prefrontal cortex, Nat. Neurosci. 2 (1999) 1032–1037.
[6] L. Bach, F. Happe, S. Fleminger, J. Powell, Theory of mind: independence of executive function and the role of frontal cortex in acquired brain injury, Cogn. Neuropsychiatr. 5 (2000) 175–192.
[7] S. Baron-Cohen, M. O'Riordan, V. Stone, R. Jones, K. Plaisted, Recognition of faux pas by normally developing children with asperger syndrome or high-functioning autism, J. Autism Dev. Disord. 29 (1999) 407–418.
[8] S. Baron-Cohen, H.A. Ring, S. Wheelwright, E.T. Bullmore, M.J. Brammer, A. Simmons, S.C. Williams, Social intelligence in the normal and the autistic brain: an fMRI study, Eur. J. Neurosci. 11 (1999) 1891–1898.
[9] S. Baron-Cohen, H.A. Ring, A model of the mindreading: neuropsychological and neurobiological perspectives, in: P. Mitchell, C. Lewis (Eds.), Origins of Understanding of Mind, Erlbaum, New Jersey, 1994, pp. 183–207.
[10] S. Baron-Cohen, H.A. Ring, J. Moriarty, B. Schmitz, D. Costa, P. Ell, Recognition of mental state terms. Clinical findings in children with autism and a functional neuroimaging study of normal adults, Br. J. Psychiatry 165 (1994) 640–649.
[11] M.L. Bauman, Autism: clinical features and neurobiological observations, in: H. Tager-Flusberg (Ed.), Neurodevelopmental Disorders, MIT Press, Cambridge, 1999, pp. 383–399.
[12] M.L. Bauman, T.L. Kemper, Neuroanatomic observations of the brain in autism, in: M.L. Bauman, T.L. Kemper (Eds.), The Neurobiology of Autism, Johns Hopkins Press, Baltimore, 1994, pp. 119–145.
[13] A. Bechara, A.R. Damasio, H. Damasio, S. Anderson, Insensitivity to future consequences following damage to human prefrontal cortex, Cognition 50 (1994) 7–12.
[14] G. Berlucchi, S. Aglioti, The body in the brain: neural bases of corporeal awareness, Trends Neurosci. 20 (1997) 560–564.
[15] D.M.V. Bishop, Annotation: autism, executive functions and theory of mind: a neuropsychological perspective, J. Child Psychol. Psychiatry 34 (1993) 279–293.
[16] S.-J. Blakemore, J. Decety, From the perception of action to the understanding of intention, Nat. Rev. Neurosci. 2 (2001) 561–567.
[17] D.M. Bowler, 'Theory of mind' in Asperger's syndrome, J. Child Psychol. Psychiatry 33 (1992) 877–893.
[18] B. Broks, Brain, self, and others: the neuropsychology of social cognition, in: G. Claridge (Ed.), Schizotypy: Implications for Illness and Health, Oxford University Press, Oxford, 1997, pp. 99–123.
[19] L. Brothers, The social brain: a project for integrating primate behavior and neurophysiology in a new domain, Concepts Neurosci. 1 (1990) 27–51.
[20] L. Brothers, Friday's Footprint: How Society Shapes the Human Mind, Oxford University Press, Oxford, 1997.
[21] G. Buccino, F. Binkofski, G.R. Fink, L. Fadiga, L. Fogassi, V. Gallese, R.J. Seitz, K. Zilles, G. Rizzolatti, H.J. Freund, Action observation activates premotor and parietal areas in somatotopic manner: an fMRI study, Eur. J. Neurosci. 13 (2001) 400–405.
[22] R.W. Byrne, A. Whiten, Cognitive evolution in primates: evidence from tactical deception, Man 27 (1992) 609–627.
[23] F. Castelli, F. Happe, U. Frith, C. Frith, Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns, Neuroimage 12 (2000) 314–325.
[24] B.G. Charlton, H.A. McClelland, Theory of mind and the delusional disorders, J. Nerv. Ment. Dis. 187 (1999) 380–383.
[25] G.A. Press Courchesne, R. Yeung-Courchesne, Parietal lobe abnormalities detected with MR in patients with infantile autism, Am. J. Radiol. 160 (1993) 387–393.
[26] R. Corcoran, G. Mercer, C. Frith, Schizophrenia, symptomatology and social inference: investigating the theory of mind in people with schizophrenia, Schizophren. Res. 17 (1995) 5–13.
[27] A.R. Damasio, Descartes' Error: Emotion, Reason, and the Human Brain, Putnma, New York, 1994.
[28] E. Daprati, N. Franck, N. Georgieff, J. Proust, E. Pacherie, J. Dalery, M. Jeanerod, Looking for the agent: an investigation into consciousness of action and self-consciousness in schizophrenic patients, Cognition 65 (1997) 71–86.
[29] R.J. Davidson, D.C. Jackson, N.H. Kalin, Emotion, plasticity, context, and regulation: perspective from affective neuroscience, Psychol. Bull. 126 (2000) 890–909.
[30] S.H. Fatemi, A.R. Halt, Altered levels of Bcl2 and p53 proteins in parietal cortex reflect deranged apoptotic regulation in autism, Synapse 42 (2001) 281–284.
[31] C. Fine, J. Lumsden, R.J.R. Blair, Dissociation between 'theory of mind' and executive functions in a patient with early left amygdala damage, Brain 124 (2001) 287–298.
[32] J. Flavell, Cognitive development: children's knowledge about the mind, Annu. Rev. Psychol. 50 (1999) 21–45.
[33] P.C. Fletcher, F. Happe, U. Frith, S.C. Baker, R.J. Dolan, R.S.J. Frackowiak, C. Frith, Other minds in the brain: a functional imaging study of 'theory of mind' in story comprehension, Cognition 57 (1995) 109–128.
[34] A. Forster, S. Lewis, M. Owen, R. Murray, Premorbid adjustment and personality psychosis, Br. J. Psychiatry 158 (1991) 171–176.
[35] C. Frith, The Cognitive Neuropsychology of Schizophrenia, Erlbaum, Hove, 1992.
[36] C. Frith, S.-J. Blakemore, D.M. Wolpert, Explaining the symptoms of schizophrenia: abnormalities in the awareness of action, Brain Res. Rev. 31 (2000) 357–363.
[37] C. Frith, R. Corcoran, Exploring 'theory of mind' in people with schizophrenia, Psychol. Med. 26 (1996) 521–530.
[38] C. Frith, U. Frith, Interacting minds—a biological basis, Science 286 (1999) 1692–1695.
[39] U. Frith, Autism: Explaining the Enigma, Blackwell, Oxford, 1989.
[40] U. Frith, C. Frith, The biological basis of social interaction, Curr. Dir. Psychol. Sci. 10 (2001) 151–155.
[41] J.M. Fuster, The Prefrontal Cortex, Raven Press, New York, 1989.
[42] H.L. Gallagher, F. Happe, N. Brunswick, P.C. Fletcher, U. Frith, C. Frith, Reading the mind in cartoons and stories: an fMRI study of 'theory of mind' in verbal and nonverbal tasks, Neuropsychologia 38 (2000) 11–21.
[43] V. Gallese, A. Goldman, Mirror neurons and the simulation theory of mind-reading, Trends Cogn. Sci. 2 (1998) 493–501.
[44] C. Gillberg, Autism and autistic-like conditions: subclasses among disorders of empathy, J. Child Psychol. Psychiatry 33 (1992) 813–842.
[45] V. Goel, J. Grafman, N. Sadato, M. Hallett, Modeling other minds, Neuroreport 6 (1995) 1741–1746.
[46] C. Goodwin, The interactive construction of a sentence in natural conversation, in: G. Psathas (Ed.), Everyday Language: Studies in Ethnomethodology, Irvington, New York, 1979, pp. 97–121.
[47] A. Gopnik, H. Wellman, Why the child's theory of mind is a theory, Mind Lang. 7 (1992) 145–171.
[48] R.M. Gordon, Folk psychology as simulation, Mind Lang. 1 (1986) 158–171.
[49] J. Grèzes, N. Costes, J. Decety, The effects of learning and intention on the neural network involved in the perception of meaningless actions, Brain 122 (1999) 1875–1887.
[50] S. Grossberg, The imbalanced brain: from normal behavior to schizophrenia, Biol. Psychiatry 48 (2000) 81–98.
[51] D.A. Gusnard, E. Akbudak, G.L. Shulamn, M.E. Raichle, Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function, Proc. Natl. Acad. Sci. 98 (2001) 4259–4264.
[52] F. Happé, S. Ehlers, P.C. Fletcher, U. Frith, M. Johansson, C. Gillberg, R. Dolan, R. Frackwiak, C. Frith, 'Theory of mind' in the brain: evidence from a PET scan study of Asperger syndrome, Neuroreport 8 (1996) 197–201.

[53] P.L. Harris, From simulation to folk psychology: the case for development, Mind Lang. 7 (1992) 120–144.
[54] M.M. Haznedar, M.S. Buchsbaum, T.C. Wei, P.R. Hof, C. Cartwrights, C.A. Beinstock, E. Hollander, Limbic circuitry in patients with autism spectrum disorders studies with positron emission tomography and magnetic resonance imaging, Am. J. Psychiatry 157 (2000) 1994–2001.
[55] J.K. Hietanen, D.I. Perrett, A role of expectation in visual and tactile processing within the temporal cortex, in: T. Ono, L.R. Squire, M.E. Raichle, D.I. Perrett, M. Fukuda (Eds.), Brain Mechanisms of Perception and Memory: From Neuron to Behavior, Oxford University Press, New York, 1993, pp. 83–103.
[56] P.R. Hobson, Autism and the Development of Mind, Erlbaum, 1993.
[57] M.A. Howard, P.E. Cowell, J. Boucher, P. Broks, A. Mayes, A. Farrant, N. Roberts, Convergent neuroanatomical and behavioral evidence of an amygdala hypothesis of autism, Neuroreport 11 (2000) 2931–2935.
[58] W.D. Hutchison, K.D. Davis, A.M. Lozano, R.R. Tasker, J.O. Dostrovsky, Pain-related neurons in the human cingulate cortex, Nat. Neurosci. 2 (1999) 403–405.
[59] M. Iacaboni, Mapping human cognition: thinking, numerical abilities, theory of mind, consciousness, in: A.W. Toga, J.C. Mazziotta (Eds.), Brain Mapping: The Systems, Academic Press, San Diego, 2000, pp. 523–534.
[60] M. Iacaboni, R.P. Woods, M. Brass, H. Bekkering, J.C. Mazziotta, G. Rizzolatti, Cortical mechanisms of human imitation, Science 286 (1999) 2526–2528.
[61] M. Jeannerod, The representing brain: neural correlates of motor intention and imagery, Behav. Brain Sci. 17 (1994) 187–245.
[62] M. Jeannerod, The 25th Bartlett Lecture: to act or not to act: perspectives on the representation of actions, Q. J. Exp. Psychol. A 1 (1999) 1–29.
[63] T. Jellema, C.I. Baker, B. Wicker, D.I. Perret, Neural representation for the perception of the intentionality of action, Brain Cogn. 44 (2000) 280–302.
[64] A.S. Kling, L. Brothers, The amygdala and social behavior, in: J.P. Aggleton (Ed.), The Amygdala: Neurobiological Aspects of Emotions, Memory, and Mental Dysfunction, Wiley, New York, 1992, pp. 232–241.
[65] R.D. Lane, G.R. Fink, P.M.-L. Chau, R.J. Dolan, Neural activation during selective attention to subjective emotional responses, Neuroreport 8 (1997) 3969–3972.
[66] J.E. Le Doux, Emotion: clues from the brain, Annu. Rev. Psychol. 46 (1995) 209–235.
[67] A. Leslie, Pretense and representation: the origins of 'theory of mind', Psychol. Rev. 94 (1987) 412–426.
[68] P.K. McGuire, E. Paulesu, R.S.J. Frackowiack, C. Frith, Brain activity during stimulus independent thought, Neuroreport 7 (1996) 2095–2099.
[69] E.K. Miller, The prefrontal cortex and cognitive control, Nat. Rev. Neurosci. 1 (2000) 59–65.
[70] J. Mlakar, J. Jensterle, C. Frith, Central monitoring deficiency and schizophrenic symptoms, Psychol. Med. 24 (1994) 557–564.
[71] T. Ohnishi, H. Matsuda, T. Hashimoto, T. Kunihiro, M. Nishikawa, T. Uema, M. Sasaki, Abnormal regional cerebral blood flow in childhood autism, Brain 123 (2000) 1838–1844.
[72] M.W. Oram, D.I. Perrett, Responses of anterior superior temporal polysensory (STPa) neurons to 'biological motion' stimuli, J. Cogn. Neurosci. 6 (1994) 99–116.
[73] S. Ozonoff, S.J. Rogers, B.F. Pennington, Asperger's syndrome: evidence of an empirical distinction from high functioning autism, J. Child Psychol. Psychiatry 32 (1991) 1107–1122.
[74] D.K. Pandya, E.H. Yeterian, Architecture and connections of cortical association areas, in: A. Peters, E.G. Jones (Eds.), Cerebral Cortex, Vol. 4, Plenum, New York, 1985, pp. 3–61.
[75] D.K. Pandya, E.H. Yeterian, Architecture and connections of cerebral cortex: implications for brain evolution and function, in: A.B. Scheibel, A.F. Wechsler (Eds.), Neurobiology of Higher Cognitive Function, Guilford Press, New York, 1990, pp. 53–84.
[76] J.M. Perez-Jaranay, F. Vives, Electrophysiological study of the response of medial prefrontal cortex neurons to simulation of the basolateral nucleus of the amygdala in the rat, Brain Res. 564 (1991) 97–101.
[77] D.I. Perrett, M.H. Harries, R. Bevan, S. Thomas, P.J. Benson, A.J. Mistlin, A.J. Chitty, K.J. Hietanen, J.E. Ortega, Frameworks of analysis for the neural representation of animate objects and action, J. Exp. Biol. 146 (1989) 87–114.
[78] G. Pickup, C. Frith, Theory of mind impairments in schizophrenia: symptomatology, severity and specificity, Psychol. Med. 31 (2001) 207–220.
[79] M.I. Posner, M.E. Raichle, Images of Mind, Scientific American Library, New York, 1994.
[80] J.L. Price, S.T. Carmichael, W.C. Drevets, Networks related to the orbital and medial prefrontal cortex; a substrate for emotional behavior?, Prog. Brain Res. 107 (1996) 523–536.
[81] A. Puce, T. Allison, S. Bentin, J.C. Gore, G. McCarthy, Temporal cortex activation in humans viewing eye and mouth movement, J. Neurosci. 18 (1998) 188–199.
[82] D. Raffman, What autism may tell us about self-awareness: a commentary on Frith and Happe, Mind Lang. 14 (1999) 23–31.
[83] V.S. Ramachandran, Anosognosia in parietal lobe syndrome, Conscious Cogn. 4 (1995) 22–51.
[84] V.S. Ramachandran, S. Blakeslee, Phantoms in the Brain, William Morrow, New York, 1998.
[85] G.P. Reynolds, The amygdala and the neurochemistry of schizophrenia, in: J.P. Aggleton (Ed.), The Amygdala: Neurobiological Aspects of Emotion, Memory, and Mental Dysfunction, Wiley-Liss, New York, 1992, pp. 561–574.
[86] G. Rizzolatti, M.A. Arbib, Language with our grasp, Trends Neurosci. 15 (1998) 188–194.
[87] J.A. Rosenkranz, A.A. Grace, Dopamine attenuates prefrontal cortical suppressions of sensory inputs to the basolateral amygdala of rats, J. Neurosci. 21 (2001) 4090–4103.
[88] P. Ruby, J. Decety, Effect of subjective perspective taking during simulation of action: a PET investigation of agency, Nat. Neurosci. 4 (2001) 546–550.
[89] J. Russell, N. Mauthner, S. Sharpe, T. Tidswell, The 'windows task' as a measure of strategic deception in preschoolers and autistic subjects, Br. J. Dev. Psychol. 9 (1991) 331–349.
[90] T.A. Russell, K. Rubia, E.T. Bullmore, W. Soni, J. Suckling, M.J. Brammer, A. Simmons, S.C.R. Williams, T. Sharma, Exploring the social brain in schizophrenia: left prefrontal underactivation during mental state attribution, Am J. Psychiatry 157 (2000) 2040–2042.
[91] Y. Sarfati, M.-C. Hardy-Baylé, C. Besche, D. Widlöcher, Attribution of mental states to others in people with schizophrenia: a non-verbal exploration with comic-strip, Schizophr. Res. 25 (1997) 199–209.
[92] Y. Sarfati, M.-C. Hardy-Baylé, E. Brunet, D. Widlöcher, Investigation of theory of mind in schizophrenia: influence of verbalization in disorganized and non-disorganized patients, Schizophr. Res. 37 (1999) 183–190.
[93] F. Schneider, U. Weiss, C. Kessler, J.B. Salloum, S. Posse, W. Grodd, H.W. Mueller-Gaertner, Differential amygdala activation in schizophrenia during sadness, Schizophr. Res. 34 (1998) 133–142.
[94] M.E. Shenton, C.D. Chandlee, M. Frumin, R.W. McCarley, A review of MRI findings in schizophrenia, Schizophr. Res. 49 (2001) 1–229.
[95] K. Shima, K. Aya, H. Mushiake, M. Inase et al., Two movement-related foci in the primate cingulate cortex observed in signal-triggered and self-paced forelimb movements, J. Neurophysiol. 65 (1991) 188–202.
[96] B. Sodian, U. Frith, Deception and sabotage in autistic, retarded and normal children, J. Child Psychol. Psychiatry 33 (1992) 591–605.
[97] S.A. Spence, D.J. Brooks, S.R. Hirsch, P.F. Liddle, J. Meehan, P.M. Grasby, A PET study of voluntary movement in schizophrenic patients experiencing passivity phenomena (delusions of alien control), Brain 120 (1997) 1997–2011.

[98] V.E. Stone, S. Baron-Cohen, R.E. Knight, Frontal lobe contributions to theory of mind, J. Cogn. Neurosci. 10 (1998) 640–656.

[99] D.T. Stuss, F.D. Benson, The Frontal Lobes, Raven Press, New York, 1986.

[100] D.T. Stuss, G.G. Gallup, M.P. Alexander, The frontal lobes are necessary for 'theory of mind', Brain 124 (2001) 279–286.

[101] S.F. Taylor, I. Liberzon, L.R. Decker, R.A. Koeppe, A functional anatomic study of emotion in schizophrenia, Schizophr. Res. 58 (2002) 159–172.

[102] K. Vogeley, P. Bussfeld, A. Newen, S. Herrmann, F. Happé, P. Falkai, W. Maier, N.J. Shah, G.R. Fink, K. Zilles, Mind reading: neural mechanism of theory of mind and self-perspective, Neuroimage 14 (2001) 170–181.

[103] H. Wimmer, J. Perner, Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception, Cognition 13 (1983) 103–128.