PATHOLOGIES OF THE SOCIAL BRAIN

Allan Young
McGill University (Montreal)

This paper describes efforts to investigate the evolutionary origins of the human mind and brain. There is a vast literature on this subject and my account is necessarily schematic. I focus on efforts that approach mind, brain, and evolution through the portal of psychopathology and neuropathy. My history is explicitly tendentious: it follows a single pathway, from the birth of the *Jacksonian brain* to the discovery of the *social brain*.

The social brain is a consequence of two developments. The first development is the availability of functional neuroimaging technologies to investigators who were, up to this point, without access to the brain. With this equipment, it has become possible, for the first time, to see the mind working inside the brain at the researcher's command. The second development concerns psychiatry. The history of the mind and brain is simultaneously a history of psychopathology. This is true in the obvious sense that changing ideas about the mind and brain influence diagnostic and therapeutic practices. Bit it is also true in a less obvious sense. The history of the mind and brain leads through a series of mental disorders. Each disorder served, in its turn, as a window on the human condition and offered researchers a distinctive combination of features (cognitive, behavioural, epidemiological, etc.) and discursive opportunities. Today autism is a target for research, along with schizophrenia and sociopathy. Autism did not exist as a diagnostic classification until the mid-twentieth century. Once discovered, it metamorphosed through a series of clinical representations. It continues to evolve today, partly in response to experimental work on the social brain and partly in response to changes in psychiatry and the wider society. Once might say that the discovery of the social brain had to wait until autism was fully prepared.

In *The Descent of Man* (1871), Charles Darwin writes that the evolution of humankind is an unoriginal idea that is now shared by "eminent naturalists and philosophers… and especially by [Ernst] Haeckel …". Haeckel's investigations of comparative morphology persuaded him that the embryonic body compasses an evolutionary archive. A species' embryonic development

(ontogeny) recapitulates its evolutionary history (phylogeny), in that embryonic stages of higher species resemble the adult forms of lower species.[1] He organized his archive as a table of images (drawings): each vertical column represented a species (e.g., fish, salamander, pig, human) and each horizontal row bracketed a stage of embryonic development. The strongest resemblances are found in the earliest stage (topmost row); specialized features emerge in descending columns. In the era of molecular genetics, it is the top row – the conservation of features across species – that captures one's attention. The molecular biologist sees these similarities as evidence of shared modular body plans, Hox genes, and transcriptional regulators that have remained remarkably constant from among metazoan animals.[2] And each column reveals a similar evolutionary process, regulated by two factors: "the constraints that at every level control the systems involved, and the historical circumstances that control the actual interactions between the systems". Simpler organisms "are more dependent on constraints than on history. As complexity increases, history plays a great part".[3] François Jacob summarized the process this way:

> Natural selection …works like a tinkerer … who does not know exactly what he is going to produce, but uses … everything at his disposal to produce some kind of workable object. … [L]ike the rest of our body, our brain is a product of natural selection. … This evolutionary procedure – the formation of a dominating neocortex coupled with the persistence of a nervous and hormonal system partially, but not totally under the rule of the neocortex – strongly resembles the tinkerer's procedure. It is somewhat like adding a jet engine to an old horse cart. It is not surprising … that accidents, difficulties, and conflicts can occur. … [E]volution is far from perfection.[4]

Jacob vision is borrowed Levi-Strauss's description of *bricolage*:

> The *bricoleur* interrogates the heterogeneous objects that comprise his treasure, an inventory of tools and materials whose precise use and relations were determined in the past. He questions them in order to understand what each element can possibly 'signify' and thereby contribute to defining the assemblage [his *bricolage*] that might now come into existence and would differ from the source only in the internal arrangement of its

> parts. The poetry of the *bricolage* cannot be reduced to the mechanical execution of this operation, for the *bricoleur* speaks not only *with* these found objects, but also through them. [In this way,] he injects into this project something of his own character and life.[5 α]

Compare the texts: there is an obvious difference. Jacob's *bricolage* lacks 'poetry'. There can be no suggestion of 'the character and life of its author' (natural selection), unless we accept that Jacob is *also* the author. In which case, both versions are correct. The difference is inevitable and we must live with it, even in science. The following pages offer a history of the *bricolage* mentioned in Jacob's text, "our brain as a product of natural selection". To be more precise, I will trace a history of the assemblage of ideas, experimental practices, and new and old objects commonly called the 'social brain'.

PART ONE – JACKSONIAN BRAINS AND MODULAR MINDS

The Jacksonian Brain

My history of the social brain begins in the nineteenth century. In a series of publications culminating with the second edition of *Principles of Psychology* (1870), Herbert Spencer constructed an evolutionary theory of the nervous system and 'the form of life which we call Mind'. According to Spencer, the nervous system comprises a hierarchy of 'centres' (localized structures), each acquired as an adaptation in the course of evolution. The most archaic centres are automatic, highly organized, impossible to modify, and dedicated to rapid sensory-motor processing. More recently acquired centres are voluntary, flexible, slow and deliberative. The higher centres of the lower vertebrates are homologues of the lower centres of higher vertebrates; in the course of evolution, the 'seats of consciousness' in lower forms are transformed into reflex

---

α « Tous ces objets hétéroclites qui constituent son trésor, il les interroge pour comprendre ce que chacun d'eux pourrait « signifier », contribuant ainsi à réaliser, mais qui ne différera finalement de l'ensemble instrumental que par la disposition interne des parties ». « Mais il y a plus : la poésie du bricolage lui vient aussi, et surtout, de ce qu'il ne se borne pas à accomplir ou exécuter ; il « parle », non seulement avec les choses, … mais aussi au moyen des choses…. Sans jamais remplir son projet, le bricoleur y met toujours quelque chose de soi ».

centres in higher forms.[6] In short, the human brain encloses an evolutionary archive analogous to Haeckel's embryo.

The most celebrated British neurologist of his time, John Hughlings Jackson was a close reader of Spencer, and believed that his ideas might provide a key to the puzzling symptoms of neurological disease. In the *Croonian Lectures on Evolution and Dissolution of the Nervous System* (completed in 1887), Jackson described the neural hierarchy as a system of command and control, in which the higher (cortical) centres inhibit the spontaneous operations of lower centres. When higher centres are disabled by damage, fatigue, or toxins, the previously inhibited centres freely perform their evolved functions. In this way, neurological syndromes can be opportunities for reading the species' evolutionary archive.

Jackson adopted a 'concomitance' theory of mind and brain. The most recent evolved centres are the organic 'substrate' of human consciousness, but the mind is irreducible to the brain. He proposed two kinds of consciousness. *Subjective consciousness* is the unmediated experience of the world, constituted in bodily sensation (e.g. the regularity of heartbeat).[7] *Objective consciousness* is constituted in awareness of representations of experiences. Just as there is a neural hierarchy, there is a hierarchy of states of consciousness. The most evolved centres collect and re-represent the dispersed representations formed in lower centres. In neurological syndromes, the less evolved states of consciousness are exposed. Jackson coined the term 'dissolution' to label this process as it simultaneously affects brain and mind.[8]

During episodes of dissolution, observers can see reverse evolution in action. Jackson described how the process occurs during mild epileptic seizures. Normal consciousness regresses to a 'dreamy state' while the patient unconsciously enacts complex behaviour ('automatism'). This is followed by a state of 'crude sensations' characterized by nausea and fearfulness. Dissolution proceeds to a more thoroughly disinhibited state characterized by complete mental confusion ('mania').[9]

Darwinian psychiatry

Jackson died in 1911. In the years following his death, he was remembered mainly for his investigations of epilepsy and aphasia. There were a few exceptions. W.H.R. Rivers, who met Jackson while working as house physician at the National Hospital for the Paralyzed and Epileptic, at Queens Square, London. An Army doctor during World War I, Rivers published a

Jacksonian account of shellshock that was quickly forgotten.[10] The French psychiatrist Henri Ey adopted Jacksonian ideas about consciousness in his organo-dynamic theory of mental illness. While Ey's theory influenced the development of French psychosomatic medicine, its Jacksonian origins are generally ignored or unknown. Around the same time, the 1950s, Paul Maclean, at the NIMH, published his first account of the *triune brain*. This is a Jacksonian conception composed of three strata ('brains'): reptilian, paleomammalian, and neomammalian. Only MacLean's account of the limbic system (situated between reptilian and paleomammalian brains) penetrated the psychiatric mainstream. In brief, interest in evolutionary theories had evaporated by mid-century. This attitude extended beyond the Jacksonian brain. Even the evolutionary narratives of Freud, detailed in *Totem and Taboo* (1913) and *Moses and Monotheism* (1938) and basic to Freud's understanding of the human mind, were, by now, regarded as an embarrassment in psychoanalytic circles. In psychology, where decades of behaviouralism had transformed the human mind into a black box, the situation was essentially the same.

Interest in evolution renewed in the 1960s, stimulated by ethological research on primate social behaviour and the rise of sociobiology, defined broadly as the application of biology-based evolutionary theory to explain the psychological, social, and cultural dimensions of human behaviour. In 1967, John Price, a British psychiatrist, published a brief account in *The Lancet* tracing the origins of clinical depression to the social life of our Pleistocene ancestors. Recent research on baboons in East Africa provided a proxy (homology) for investigating early humans. Baboon society is hierarchical: the animals are innately aggressive and are ceaselessly competing for dominance. Price described them as being elated when going up the hierarchy and visibly depressed when going down. Their depression is an evolutionary adaptation (the 'yielding response'). It signals to the victor that the contest is over and the loser is ready to withdraw before suffering further injury; and it is evidence that physiological mobilization fuelling the behaviour has been switched off. The homologous symptom is the low energy associated with clinical depression. Baboons compete mainly for food and sex; the comparable symptoms are loss of appetite and libido and feelings of worthlessness.

Depression in humans and the yielding response in baboons are to be understood in social terms, as a ritual that "performs the function which death performs in unritualised fighting, and which the referee performs in culturally ritualised competition". This explains why depression is common and severe in old people. They have fallen to the bottom of the social hierarchy; they

are socially invisible; their depression retains its ritual form but has lost its ritual efficacy (an audience).[11]

Similar accounts have been published in the following decades. Each account begins with a psychiatric disorder – e.g. agoraphobia, postpartum depression, antisocial personality disorder. The disorder's origins are traced (by reverse engineering) to a challenge to group survival in Palaeolithic times, e.g. excessive mortality resulting from unrestricted in-group violence. The writer discovers a behavioural solution (adaptation) that matches the features of the target disorder. Each adaptation is presumed to be heritable (a behavioural phenotype) and therefore part of the human genome. Most of the identified syndromes are products of 'genome lag', i.e. our Pleistocene brains are maladapted to the demands of post-industrial society. Some of the evolved behaviour patterns generally work well but become dysfunctional and symptomatic in uncommon situations, e.g. when the survival response transmutes to posttraumatic stress disorder. Another class of syndromes remain adaptive (enhance fitness) but are incidentally a cause of pain and misery. For example, the low mood, low energy, and difficulty concentrating associated with depression are distressful but also inhibit self-defeating impulsive behaviour that might otherwise follow stressful life events (divorce, job loss, etc.).[12]

Modular minds

Pleistocene genotypes are connected to behavioural phenotypes (adaptations and syndromes) through the brain and the mind. When these writers occasionally mention neural mechanisms, it is only in passing. But they are divided regarding the mind. Darwinian psychiatrists, including John Price, Randolph Nesse and Isaac Marks[13], take mind for granted. Nothing is said about mental architecture, internal representations, etc. The evolutionary psychologists are different. Their object of inquiry is the mind, specifically a human mind divided into heritable components or 'modules'. Each mental module operates like a minicomputer dedicated to solving a single problem that would have affected reproductive success (fitness) in Pleistocene times. Each module generates an automatic behavioural response to a predetermined stimulus situation detected with the help of sensory 'filters' that are specific to this module. Each module is insulated from information and interference from other modules, but there is no consensus regarding the completeness of encapsulation. Some writers contend that limited intercommunication is possible and behavioural phenotypes may be products of multiple

modules operating in series. Certain psychopathologies, notably autism, are traced to a defective module that disrupts the flow of information within a series.[14]

The best known theory claims that the mind comprises a mosaic of modules: a 'confederation' of thousands of functionally dedicated mini-computers within which there is no place or need for a domain-free cognitive processor comparable to 'reasoning' in the generic sense. Each module represents an evolutionary adaptation and a corresponding narrative. There is no 'master narrative' of the evolution of the modular mind, but rather the aggregation of particular narratives, each keyed to a distinctive function and module.

Because modules are encapsulated (autonomous) and there is no supervening mental structure such as 'the ego', one might conclude that the modular mind lacks a master narrative. This is not entirely true however. There is a module, described as 'social intelligence' and 'theory of mind', that preserves certain properties of a general purpose processor. It is called 'domain-specific' because it is dedicated to transactions between people. But its putative specificity must be taken with a grain of salt, since the life world of humans, from the Pleistocene to the twenty-first century, is overwhelmingly a social world. To function properly (enhance fitness), a social intelligence module must be able to exploit information and coordinate responses generated by diverse domain-specific modules. In this context, 'coordination' means to select, sequence, and inhibit responses. As we shall see in a moment, these functions have a special potential for explaining the phylogeny and ontogeny of mental disorders, notably schizophrenia.

To understand the operations of the social intelligence module, one must first know its evolutionary narrative. I will postpone this narrative until Part Two, where I will describe how it provides a scaffold for constructing, via *bricolage*, a new brain-based evolutionary science of psychopathology. Before progressing to Part Two however, I want to consider developments that coincide with the emergence of the modular mind and return us to the subject of the Jacksonian brain.

The re-emergence of the Jacksonian Brain

In the 1980s, researchers identified a previously unreported class of neurons in the premotor cortex of rhesus monkeys. These 'mirror neurons' discharged when monkeys *performed* goal-directed actions and also when they passively *observed* other monkeys

performing these actions. Next, researchers identified visuomotor neurons that discharged when monkeys *listened* to sounds associated with these goal-directed actions. Neurons do not 'resonate' when an observed activity is not goal-related, e.g. merely lifting an arm. Later experiments showed that mirror neurons also discharge when monkeys observe facial emotions and sensations in other monkeys – the neuronal basis for 'emotional contagion'. [15]

The adaptive value of the arrangement is clear. Since the representations in the brain are multi-modal, non-human primates are able to infer the *goal* of an observed action even when visual information is incomplete. But apes and monkeys are unable to go one step further, to attribute *intentions* to the observed individuals. This capacity is limited to humans, as I shall explain in a moment.[16]

By the 1990s, researchers had demonstrated the existence of a human 'neural matching system' enabling people "to recognize, understand, and imitate the behaviour of others".[17] The system they describe is more extensive and complex than networks revealed in other primates. Further, findings connect mirror neurons to language and semantics. In work reported in the journal *Neuron*, investigators asked participants to passively read a list of action words that included 'lick', 'pick', and 'kick'. fMRI images showed a substantial match with the cortical regions that are activated when people move their tongues, fingers, and feet. According to the researchers, this is evidence that "words are processed by distributed neuronal assemblies with cortical topographies that reflect word semantics".[18]

Two decades of research lead to the conclusion that human activity is permeated by the operations of mirror neurons. The formation of the human mirror neuron system goes back 20 million years, to the emergence of the old world monkeys. Constituent elements (and functions) emerged at different points in time; the system's formation has been additive and the sequence is repeated in the cognitive development of children. Ontogeny recapitulates phylogeny:

1. The observer's mirror neurons resonate with an observed behaviour. The result is a transitory 'primary representation'. It is experienced by the observer as his own state, and is comparable to the 'subjective consciousness' described by John Hughlings Jackson. This ability is shared by monkeys, apes, and humans; the following capacities are limited to humans.

2. The action pattern is 'copied' by the observer's brain where it becomes a second-order representation and the basis for 'perspective-taking', seeing the target's view.

3. The copy is added to a repertoire of action patterns stored in the observer's brain. It is the basis of imitative behaviour, but can also be generalized to serve other situations (analogy, mimesis).

4. In order for the repertoire of copies to be adaptive and, in this way, incorporated into the human genome, the brain must possess additional capacities, namely the ability to filter competing sensory inputs; select and sequence action patterns; inhibit prepotent (automatic) responses; and monitor the over-all process.

5. Perspective-taking extends the boundaries of mirrored emotion states. Emotional contagion is the most primitive state from an evolutionary point of view and it occurs in both human and non-human primates, as a primary representation. *Empathy* is a second-order representation of emotion and therefore unique to humans. It is also neurologically complex, since it activates multiple networks of neurons, parts of the limbic brain, and episodic memory, and entails the on-going processing of autonomic sensation.[19]

6. The capacity to put copies off-line (intrinsic to perspective-taking) underpins developments in language – emergence of past and future tense, subjunctive and conditional tense – and powers of creative imagination.

Mirror neurons are necessary for operating this process but are insufficient for explaining it. An additional mechanism is required after Step 1, in order explain how the selection, sequencing, monitoring, and inhibition of action patterns occur; how copies are placed off-line; and how people are able to "mentally construct possible (e.g. planned) events in the future and reconstruct personal events from the past".[20] The required mechanism is called 'executive functions' (EF). Neuroimaging research locates EF in various parts of the neocortex, depending on the operation being investigated. Are the multiple locations connected and coordinated by a unifying circuitry? The executive function is identified by a set of operations that would seem to presume intentionality. Does a 'self 'or self-like element such as 'the will' emerge from these networks? I will return to this question in Part Two.

Jackson's understanding of the brain's 'higher centres' is quite similar. The centres are multiple and operate in the fashion of a 'committee' rather than a sovereign. There are additional similarities. The formation of the social brain (human mirror neuron system plus EF) dates back twenty million years, to the time when our monkey ancestors split from their antecedents. The brain's constituent elements emerged at different points during this long period: its formation

was additive rather than modular. The final system is hierarchical. Simple functions that emerged at an early stage (making copies for example) are controlled and inhibited by components (EF) that evolved later and are functionally complex and flexible.

The operations of the social brain are revealed in the unusual behaviour of autistic children. They exhibit deficits in capacities served by mirror neurons including empathetic engagement. They also exhibit EF deficits and are typically unable to create the off-line copies needed for constructing action plans based on imitation. The inability of EF to inhibit spontaneous motor neuron activation is reflected in *echolalia*, a condition in which autistic individuals automatically repeat what is said to them. And EF deficits explain an inability to reconstruct (narrate) the personal past, also reported to be common among autistic individuals.[21]

Comparable EF deficits explain symptoms associated with schizophrenia including delusions and 'disordered speech' (irrelevant chains of word-associations and digressive departures from the logical sequence of their utterances). In normal speech, the sequencing of words into sentences is made possible by the brain's ability (EF) to suppress tendencies to 'free association'.[β] Disordered speech in schizophrenia is evidence of divergence between the patient's initial non-verbal plan for his utterance and his actual utterance. The EF operation that is required to translate plan into an utterance has gone awry. The patient is bewildered by the discrepancy between intention (still in mind) and execution, and confronts an obvious question: "If I cannot control my speech and my voice is not saying what I want to say, then who or what is controlling it?" Schizophrenic delusions – evidence of thought insertion, thought broadcasting, thought blocking, etc. – are the patient's effort to understand his experience. "From the conclusion that some outside agent is controlling one's behaviour there follows inevitably the

---

[β] Comparable effects have been demonstrated experimentally. In one experiment, a group of people diagnosed with schizophrenia and a group of individuals without psychiatric diagnoses participate in Stroop-like tasks. Colour words first appear on a screen in concordant colours ('RED' is coloured red, 'BLUE' is coloured blue) and then in discordant colours. As each word appears, participants are asked to name the colour rather than the word. Individuals diagnosed with schizophrenia exhibit a longer response time, and the outcome is interpreted as evidence of a deficit in the in executive function (ability to inhibit competing sensory inputs). Functional neuroimaging during the tasks locates executive function in sectors of the pre-frontal cortex and may identify concomitant differences between schizophrenia and control groups.

question 'Why are they doing this to me and not to anybody else? What are their motives and why am I the target?'"[22]

PART TWO: THE EVOLUTIONARY NARRATIVE

The mind inside the social brain

The brain is different from the mind: it has always been visible. The ability see inside the brain, the functions that have been attributed to the brain and its parts: these things changed over time. But the brain's materiality was never in doubt. Until recently, the mind remained invisible, its ontological status debated. It could be observed only indirectly, through behaviour, self-reports, moments of introspection, etc. Functional neuroimaging changed this situation: it delivered an unprecedented object, the mind that can be seen inside the brain,. And unlike earlier kinds of functional neuroimaging such as PET (positron emission technology), fMRI is widely available, easily accessed by psychological researchers.

It is a remarkable cultural achievement and this is how it is produced in the laboratory. Participants are asked to complete a mental task: e.g. observe photos on a monitor (alternately neutral, grisly, erotic) or solve a semantic puzzle requiring analogical reasoning. A calibrated fMRI apparatus detects changes in metabolic activity in the participant's brain (indexed by alterations in blood flow). A computer program translates the data obtained into an image of localized structures, e.g. mirror neurons plus parts of the medial prefrontal cortex. The researcher interprets the images: he hypothesizes how the activated structures are functionally connected and then connects (explains) the identified function (mechanism) to the mental operation for which the experimental task was designed (e.g. analogical reasoning).

The procedure is used to compare people with and without psychiatric diagnoses. In which case, researchers will select experimental tasks that provide information concerning the deficits and abnormalities. Contrasting patterns of activation permit researchers to propose neural basis for the diagnostic features and, conversely, for normal capacities. The process, beginning with calibration, comprises a series of *interpretive operations*, each of which affects outcomes and interpretations downstream.

No single element in this process is more important than the choice and construction of the experimental task. And it is at this point the histories of the modular mind and the social brain now intersect, in a shared narrative describing the evolutionary origins of 'social intelligence'.

The evolutionary narrative

The narrative originates in an evolutionary puzzle and its solution. Why did hominid brains grow in size and power during the Pleistocene? Big brains are expensive: they consume large energy inputs and entail a prolonged period of infantile dependency. But big brains also have advantages: they are a medium for transmitting complex information (culture and technology) and creating a social environment within which individual learning could occur. In other words, they enhance group 'fitness', measured as reproductive success.

Group life in the Pleistocene was maintained through cooperation, altruistic behaviour (i.e. exchanges in which donors sacrifice overall fitness), and 'mind-reading'. Mind-reading is a cognitive capacity that permits me to observe and infer other people's intentions, predict their behaviour on these inferences, and detect efforts by 'cheaters' or 'defectors' (a term borrowed from game theory) to deceive me by reading my mind. To measure the mind-reading capacity in children and adults, clinicians and researchers employ 'false belief' tests.[χ23]

Cheating is profitable in the short run: defectors get more food, more copulation, better nourished children, and increased numbers of surviving offspring. Over time, defectors should predominate in society. If this were to happen, social behaviour would grow less predictable and group fitness would inevitably decline. Natural selection solved this problem by giving our ancestors a (modular) capacity to detect defectors and a corresponding urge to punish defectors even if this entails significant costs, such as reprisals, and no direct benefits to the individual.

[χ] Variations of the "Sally and Anne task" are widely used in research comparing normal and autistic children, and also normal and schizophrenic adults. Participants are asked to observe interaction between Sally and Anne in a series of panels. Sally places her ball in a basket. She leaves the scene. Anne transfers the ball to a box. Sally re-enters the scene. Participants are asked "Where will Sally look for her ball?" If the answer is 'the basket', then one concludes that the participant has read Sally's mind. If the answer is 'the box', there are two possibilities: the participant cannot read minds or she has a memory deficit. An additional question, "Where was the ball in the beginning?", is intended to determine whether memory capacities are intact.

The cheater detector is a temporary solution. Eventually defectors learn to use the detector to perfect more sophisticated deceptions. To maintain group fitness, natural selection provides an up-dated version of the cheater detector. A relentless 'cognitive arms race' is set in motion and it produces progressively more powerful versions of the social brain. And so, a competent lie is a significant achievement since it requires the ability to exploit the target's latest model cheater detector and the ability to suppress leaking information about one's intentions:

> [T]he liar is called upon to do at least two things simultaneously. He must construct a new item of information (the lie) while also withholding a factual item (the truth)….. [Since] the truthful response comprises a form of baseline, or pre-potent response, … responding with a lie demands some form of additional cognitive processing, that it will engage executive, prefrontal systems (more so than telling the truth) … [and] may be tested using functional neuroimaging. [24]

An unusual clinical condition called 'pseudopsychopathic personality syndrome' illustrates what happens when executive system connectivity is damaged (by orbitofrontal brain lesions) and unable to suppress the truth-telling response. These individuals have no problem imitating people or reading their minds, but they are incompetent social actors, routinely and indiscriminately tactless.

Finding the 'self' inside the social brain

Some sort of 'self 'is indispensable for the social brain is to make sense. A self is needed to take responsibility for the brain's executive functions. Who is coordinating and controlling the planning, strategizing, executing, and monitoring of actions, including speech? A self is needed to solve the 'frame problem'. The brain's data base is enormous and continually expands. The mind operates multiple kinds of inference (abduction, analogy, metoynym, synecdoche, etc.); this combination vastly complicates how a given situation (a battery of impressions entering consciousness via independent multimodal sensory inputs) might be interpreted or framed. Physical survival and the coordination of complex social relations will need an operator (the self, a central processor) to determine relevance and context. A self is likewise needed to solve for the puzzle of 'consciousness' (awareness, experience of subjectivity). When I say "I can shift my

attention at will", who actually does the shifting?"[25] And a self is needed because its pathologies are intrinsic to understanding disturbances of intentionality (when reasons, motives, and behaviour fail to mesh in, e.g. addictive behaviour, anorexia nervosa), depersonalization (when the individual is no longer the subject of his experiences), autobiographical memory (e.g. PTSD, confabulation), perception (delusions), and personality.

The self is indispensable, but is it also a thing-in-itself? Or is it merely a convenient fiction: a conceptual space housing empirical phenomena such as 'personal identity', 'self-awareness', and the 'social self'? And if the self is more than a fiction, how might it be imagined? John Searle writes that each of us is tempted to presume that there is a person-like entity inside his head, an 'inner-self' or a 'self-reflective ego' that does one's thinking, perceiving, remembering, and acting. This homunculus is a fallacy and leads to an infinite regress. If my subjective experience of my current situation is possible because an inner-self is observing the situation unfold before his eyes, then one has to suppose that inner-self has its own inner-self who performs the operations described in the preceding paragraph. And so on.[26]

How might mind and neo-Jacksonian brain be pulled together to form a satisfactory inner-self – a self that is not a homunculus and that is more substantial than the sum total of its executive functions? Todd Feinberg, a clinical neuroscientist, has recently outlined such a self. The inner-self is said to emerge from a "nested hierarchy of meaning created by the brain". Mental representations are continually created at different levels within the hierarchy. Representations produced at lower levels are jointly represented in awareness. In this fashion, successively higher levels of meaning are produced. Higher levels constrain (control) lower ones. The process is phylogenetic (consistent with the evolutionary ideas of Jackson and MacLean) and ontogenetic. The arrangement preserves the individual parts located at lower levels, but eliminates the independence of each part when it is operating within the framework of the nested hierarchy that constitutes the self. Thus there is no supreme commander directing the operation; control is generated from within the entire nested system. [27]

Feinberg creates *an inner-self for patients* (psychopathologies). With this concept, he can locate neural nodes at which the brain's nested hierarchy breaks down and corresponding functions fail to perform. Researchers have now located a great many self-referential 'regions of interest' in the brain. And even at this relatively early stage, the complexity of the 'system' is overwhelming and must be managed (reduced) by statistical operations such as cluster analysis.[28]

But does Feinberg's framework provide *an inner-self suitable for a normal brain*? His inner-self can explain *failures* of intentionality, but can it explain intentionality? Not yet. To do this, it would have to accommodate a relationship between the inner-self and autobiographical memory. The consequent demands placed on the system of nested hierarchies (the 'frame problem') would increase exponentially. How might a solution 'emerge' under these conditions? Would the system's neural complexity increase correspondingly? In which case, the term 'complexity' explains nothing. It simply confirms Jackson's original idea that brain and mind are related by concomitance rather than causality.

PART THREE – THE SPIRIT OF THE *BRICOLEUR*

> *The bricoleur speaks not only with these found objects, but also through them. He injects into his bricolage something of his own character and life.*

A mind for our times

The mind and brain just described are profoundly social in origins and capacities. The social brain evolved during the Pleistocene. According to the narrative, the physical environment and ancestral social life remained stable and homogeneous. Our ancestors lived in small local groups glued together by empathy, an ability to read minds, an irresistible urge to exchange valuables (food, labour, women), and a primitive sense of justice based on strong reciprocity (the need to repay in kind) and altruistic punishment.

The social brain has no interiority. Its 'inner-self' manages cognitive functions and behaviour, but cannot explain intentionality. The social brain accommodates 'facial emotions' (anger, fear, disgust, joy, sadness, shame) but ignores the complex, private emotions (guilt, remorse, etc.). It accommodates autobiographical memory, but mainly as a mechanism serving economic exchange. (Memory is the basis for assessing 'reputation', i.e. an individual's record of strong reciprocity and altruistic punishment.)

In the Late Pleistocene, the ancestral populations dispersed beyond the zone of origin in Africa. They adapted to heterogeneous physical environments and morphological and physiological differences emerged, but their social environment remained more or less unchanged. At this point, the social brain was complete, it evolved no further. Circumstances

altered during the Neolithic Period (the food production revolution) and, later, with the emergence of cities and polities based on a complex division of labour. Since that process began just 10,000 years ago, it is presumed that the intervening period was too brief to affect the social brain.

If true, these circumstances would eliminate the possibility of significant *heritable* racial differences affecting brain and mind, and we would attribute mental differences among populations to culture rather than biology. This assumption, that cultural differences explain mental differences, ignores a further possibility, intrinsic to human niche construction. Kim Sterelny has described how human groups construct the ecological niches that they occupy. Construction is the joint-product of biological inheritance (the social brain), cultural inheritance (technologies, systems of meaning, etc.), and physical environment. Adaptation to life within constructed niches can generate adaptive genetic changes some of which ultimately affect brain and mind, e.g. through the modification of metabolic pathways that impact on the development of the maturing nervous system. Increasing cultural differentiation generated by cumulative niche construction would make boundaries between groups less permeable. By reducing in-migration, circumstances would "accelerate the pace of evolutionary change, not just by accelerating environmental change, but also by making adaptive responses to environmental change less likely to be lost".[29]

Suggestive evidence that this effect may have occurred comes from two sources: significant differences observed in responses to psychoactive drugs among Europeans, Africans, and East Asians, and significant differences in the epidemiology of selected psychiatric disorders, notably schizophrenia.

The idea that the evolutionary narrative of the social brain should be extended into recent times has been consistently downplayed. For example, read the DSM-IV pages (Appendix I) devoted to 'culture specific syndromes'.[30] A less controversial proposal is that men and women are separated by heritable differences in brains and minds. (The contrast affects populations rather than individuals; it is possible for men to possess female brains and *vice versa*.)

Feminizing empathy

Baron-Cohen, world famous for his research on autism, has developed a self-report scale called 'the empathy quotient' (EQ). He uses it for measuring differences between normal adults

and adults with 'high functioning autism' (normal to superior intelligence). These autistic adults are predominantly male (10:1). They score significantly lower than normal individuals on the EQ, but do very well on a scale designed to assess one's 'systematizing ability', the drive to analyse and build systems.

> Systems can be technical (e.g., the workings of a machine), natural (e.g., the process of coastal erosion), abstract (e.g., mathematics), motoric (e.g., a guitar playing technique), taxonomic (e.g., a criterion for ordering compact discs), or social (e.g., a taxation system). When confronted with systems such as these we don't analyze them [or make predictions] in terms of emotions and mental states. Rather we examine relationships between components and correlations between events which then allow us to understand any underlying rules that may be relevant. [31]

When normal females are compared with normal males, the females generally score higher on the empathizing scale and lower on the systematizing scale. According to Baron-Cohen, the findings vindicate Hans Asberger's insight, in 1944, when he proposed that: "The autistic personality is an extreme variant of male intelligence … the male pattern is exaggerated to the extreme".[32] According to Baron-Cohen, there are three kinds of brain: a systematizing brain ('the extreme male brain'); an empathetic brain (common among females); and a balanced brain (in which capacities for empathy and systematizing are equal but reduced, as a consequence of a developmental 'trade-off').

He traces the evolutionary origins of empathetic and systematizing brains to a Pleistocene social division of labour. Good systematizers are skilled at using and making tools, hunting and tracking (understanding and exploiting natural systems), making exchanges and calculating cost-benefits, competing for social dominance (a tendency to aggression is an advantage, empathy a disadvantage), exercising leadership, and tolerating solitude. The female brain, on the other hand, is adapted to mothering (it 'reads' the needs of preverbal infants); sensitive to her male partner's emotional state and able to anticipate his next move (thus providing protection from spousal aggression); and inclined to gossip (enlarging the span of potentially useful information to which she has access).

Are heritable differences between male and female brains mirrored in women's occupational choices today? In 2005, the president of Harvard University, Lawrence Summers, "sparked an uproar … when he said that innate differences between men and women might be

one reason fewer women succeed in science and math careers", and wondered aloud whether the dearth of female professors in science and engineering at elite universities might be best explained by women's gender-specific desires and capacities.[33] In contrast, Baron-Cohen's autism-based theory of the female brain seems to have attracted no critical comments.

Other possibilities

The narrative of the social brain is pleasing. It confirms that human nature is unitary, despite differences among the world's peoples. It relates that society emerged spontaneously from an ancestral capacity for empathy and an irresistible interest in exchange. And it says that human evolution has provided us with moral foundations.

Empathy is said to be "the "glue" of the social world, drawing us to help others and stopping us from hurting others".[34] Mirror neuron research shows that empathy is aroused whenever we observe other people experiencing pain. Empathy is an essentially passive experience, but its emergence signalled the beginning of an evolutionary path leading upward to 'sympathy', the impulse "to take action to alleviate the other person's suffering".[35]

The story is a pleasing because it naturalizes certain cherished beliefs. Among my readers, there are people who say that the narrative is questionable for this same reason, that it tells us what we would like to believe. These people will imagine an alternative *bricolage*, infused with the spirit of, say, Nietzsche or Freud rather than Adam Smith. What might someone inspired by these men say about the evolutionary path from resonance to empathy to sympathy?

Do you remember the part of the narrative concerning defectors, detectors, and altruistic punishment? Good. You will also remember that, within the *bricolage*, this narrative does not stand alone. It supports practices and technologies (e.g. providing tasks for fMRI experiments) and is, in turn, brought to life by these same practices and technologies. Well then, the story of defectors, detectors, and altruistic punishment is supported by certain experiments, in which a participant is given a sum of money and told to divide it among other participants, according to any rule he chooses. The procedure is repeated with the same participants, but on each occasion a different person is asked to divide the sum. Sometimes the division is fair (consistent with social norms), sometimes it is selfish (a 'defector' is at work). The game sequence is structured so that it is possible to 'punish' defectors (e.g. by giving them no share), but this is expensive, since the enforcer must reduce the amount that he himself will now receive.

> Many people voluntarily incur costs to punish violations of social norms. Evolutionary models and empirical evidence indicate that such altruistic punishment has been a decisive force in the evolution of human cooperation. … Our [fMRI findings indicate] that people derive satisfaction from punishing norm violations and that the activation in the dorsal striatum reflects the anticipated satisfaction from punishing defectors.[36]

The observation that game players and ancestral humans "derive satisfaction from punishing norm violations" is open to multiple interpretations. In the researchers' account, the urge to punish is the end of an evolutionary road, since it explains how defectors would have been controlled. But one can, in a different spirit, imagine altruistic punishment as a new beginning. In this alternative evolutionary account, the pleasure ('satisfaction') that was originally derived from punishing violators is transmuted into a pleasure derived from punishing all sorts of people, and this pleasure combines with the urge to dominate or suppress potential rivals. And the acts of the leader resonate in the minds and brains of the followers – via the agency of mirror neurons, suggestion, emotional contagion etc. – so that he inspires both awe and dread in his followers. (Freud explored these possibilities in *Totem and Taboo*, *Group Psychology and the Analysis of the Ego*, *An Overview of the Transference Neurosis*, and *Moses and Monotheism*.)

Or take this example: mirror neuron research shows that watching the pain of others resonates in the brain of the observer. Baron-Cohen says that this experience may be the origin of sympathy. However the ability to experience someone else's suffering might equally explain the evolutionary origin of cruelty. There are ethological and ethnographic literatures that support this thesis. And there is no reason to suppose that heritable cruel urges would be maladaptive (reducing group fitness):

> Protean grouping patterns allow individuals to attack only when they have overwhelming power. Such tactical success allows them to kill safely and cheaply, and thereby win a likely increase in resources over the succeeding months or years. … [Thus] selection has favored a human tendency to identify enemies, draw moral divides, and exploit weaknesses pitilessly across boundaries.[37]

PART FOUR – THE *BRICOLEUR*'S TREASURE

The bricoleur interrogates the heterogeneous objects that comprise his treasure, an inventory of tools and materials whose precise use and relations were determined in the past. He questions them in order to understand what each element can possibly 'signify' and thereby contribute to defining the bricolage that might now come into existence

We can imagine an alternative narrative of origins, inspired by a different spirit. Must we choose between these narratives? Is it *either* Adam Smith *or* Sigmund Freud, or is it possible that the biological evolution of the human mind and social life proceeded along multiple pathways?

The 'either/or' solution assumes that human nature is pre-determined by the human genome, and that variations in mental capacities among populations (e.g. Congolese versus Japanese) are products of culture rather than biology. Molecular biology argues otherwise, namely that the pathway from genotype to phenotype is not predetermined. Multiple pathways evolved as adaptations to environmental variations in the past.[38] The pathway actually followed will depend on the organism's developmental (ontogenetic) history and its constructed niche. The same transcriptional sequence (DNA) can produce alternative cognitive or behavioural phenotypes. In other words, the minds and societies imagined by Adam Smith and Sigmund Freud may co-exist within the same human genome.

This was also François Jacob's point, quoted at the beginning of this paper, where he describes molecular genetics as *bricolage*. What about that other *bricolage*, the social mind, that is currently under construction by psychiatry, cognitive neuroscience, and 'neuroeconomics' (the theatre of altruistic punishments)? Look inside our *bricoleur*'s treasure of found objects. Can you find a disorder that might provide the *bricolage* with a second narrative or origins?

At first glance, depression seems to be the most likely candidate. Earlier, I told its narrative of origins: how ancestral humans organized themselves into hierarchies, how they competed for food and sex, how a 'yielding mechanism' evolved to preserve vulnerable individuals, and how this mechanism is now mirrored in the features of clinical depression.[39] In an addendum to this story, John Price made two more claims. When human groups are

commanded by unquestioned leaders and engaged in conflicts with external enemies, so that anger and aggression are directed outward, rates of depression decline. Second, the physician's social position is likewise high in the patient's hierarchy. And his "ability to comfort the disturbed patient lies in the fact that he is perceived as powerful, confident, assured, and yet strongly committed to the patient's interests".[40]

Good. We have found a narrative that is in the spirit of *Totem and Taboo* and radically different from the origin story based on schizophrenia and autism.[41] However depression is unavailable to our *bricoleur*, even though its narrative is supported by vast amounts of brain research conducted on humans, rodents, and monkeys over many years. The *bricoleur*'s problem is that this particular brain is significantly different from the social brain that is the core of his project. It is compatible with the social brain, but constructed along different lines, based on the patterning of neurotransmitters and neuroreceptors (serotonin, dopamine, cortisol, and ocytocin) rather than Jacksonian brain anatomy.[42]

The social brain is hierarchical and lacks interiority, while the depressed brain performs its tasks through the medium of subjective states. Low self-esteem and pessimism, rather than switches in the executive function, explain the yielding response (social and psychological withdrawl) in depression. One day, the social brain may become available to depression, but not yet.

CONCLUSION

The notion '*bricolage*', as used it in this paper, is vague and open-ended. This is its virtue. *Bricolage* is not a method of analysis, but a way to understand how objects, technologies, and representations have been brought together by manifold individuals (a collective *bricoleur*) to accomplish certain goals. I believe that it is an especially useful way to understand developments now taking place in psychiatry. Of course I may be mistaken about all of this: culture, psychiatry, etc. In which case there would be just a single *bricoleur*, me.

References to be provided

ENDNOTES

[1] Richardson and Keuck 2002; Gould 1977

[2] Kirschner and Gerhart 2005; Double and Wilkins 1998

[3] Jacob 1977: 1163.

[4] Jacob 1977: 1163-1164, 1166.

[5] Levi-Strauss 1962: 28, 32.

[6] Smith 1982a

[7] Antonio Damasio's 'somatic marker hypothesis' provides a modern analogue to Jackson's subjective consciousness.

[8] Smith 1982b

[9] Hogan and Kittiboriboon 2003; Meares 1999

[10] Interest in Rivers revived in the 1990s. For an account of his connections to Jackson, see Young 1999

[11] Price 1967; Price et al. 1994: 310, 312

[12] Nesse 2000.

[13] Marks and Nesse 1994; Nesse 2000.

[14] See Baron-Cohen 1985 on the modular origins of autism.

[15] Gallese et al. 1996; Gallese 2003b:174 ; Kohler et al. 2002; Decety 2002; Rizzolatti et al. 2002 ; Hirstein 2005:114:

> The [human] brain … has trouble keeping simulated emotions from affecting us, which is no doubt part of the reason we find it unpleasant to be with nervous or fearful people. [But] emotional infectiousness also works to synchronize our emotions with those with whom we interact, something that has social and interpersonal functions.

[16] Gallese 2003b:522-3; Filion et al. 1996; Provinelli and Preuss 1995

[17] Gallese 2002:36. The earliest research involving monkeys relied on transcranial magnetic stimulation (TMS). In the 1990s, TMS was displaced functional neuroimaging (fMRI). TMS measures neuronal activation (motor evoked potentials) in targeted body parts during experiments, and it produces graphic representations analogous to EEG results, rather than brain images characteristic of fMRI.

[18] Hauk et al. 2004:301, 305; also Pulvermüller et al. 2001, Pulvermüller 1999

[19] Gallese 2003b

[20] Williams et al. 2001:292; Ozonoff and Jensen 1999

[21] Williams et al. 2001:292; Ozonoff and Jensen 1999

[22] Maher 2003:18-19; also Gold and Hohwy 2000

[23] Baron-Cohen 1995

[24] Spence et al. 2004:1757

[25] Searle 2006:12

[26] Searle 2006:15

[27] Feinberg 2006:46

[28] See Northoff et al. 2006 for a meta-analysis of imaging studies of the self.

[29] Sterelny 2004: 239

[30] Cf. Lucien Lévy-Bruhl, *Les Fonctions metales* (1910) and *La Mentalité primitive* (1922)

[31] Baron-Cohen and Wheelwright 2004:302; the questions comprising the scales are provided in Baron-Cohen et al. 2003

[32] Cited in Baron-Cohen et al. 2003:363

[33] Bombardieri, Marcella. 2005. Summers' remarks on women draw fire. *Boston Globe*, 17 January.

[34] Baron-Cohen and Wheelwright 2004:163; Preston and de Waal 2002;

[35] Avenati et al. 2005; Baron-Cohen and Wheelwright op. cit.

[36] de Quervain et al. 2004:1254

[37] Wrangham 2004: 34-35.

[38] Sterelny; Jablonka

[39] Price et al. 1994: 310

[40] Price 1967: 244-245

[41] Randolph Nesse (2000) has proposed an additional evolutionary narrative for depression. Low mood characteristic of depression originates as a modular function that permitted humans to unconsciously disengage from situations that are futile, unprofitable, or dangerous. At such moments (e.g. divorce), one must avoid rash decisions. The start-up costs of new enterprises are high and people may ignore dangers or over-estimate the chances of success. “In this situation, pessimism, lack of energy, and fearfulness can prevent calamity even when they perpetuate misery” (p.17). Time slows down and one can consider the options. But the low mood system is prone to getting stuck in a positive feedback loop, resulting in severe depression that makes it harder to disengage from the precipitating condition.

[42] I am not suggesting that there has been no neuroimaging research relating to depression, but rather that the research is largely irrelevant to the narratives concerning the evolutionary origin of depression and the modern mind and brain.