The increasing visibility and sense of intellectual opportunity associated with neuroscience in recent years have in turn stimulated a growing interest in its past. For the first time, a general reference book on the history of science has seen fit to include a review of the history of the brain and behavioral sciences as a thread to be reckoned with within the broader narrative tapestry. On the one hand, this looks like a welcome sign that a new historical subfield has “come of age.” On the other hand, when one settles down to the task of composing a “state of the art” narrative, one realizes just how much these are still early days. The bulk of available secondary literature still swims in a space between nostalgic narratives of great men and moments, big “march of ideas” overviews, and an unsystematic patchwork of more theorized forays by professional historians into specific themes (e.g., phrenology, brain localization, reflex theory).¹

The challenge of imagining a comprehensive narrative is made all the more formidable by the fact that we are dealing here with a history that resists any easy or clean containment within disciplinary confines. The paper trail of ideas, experiments, clinical innovations, institutional networks, and high-stakes social debates not only moves across obvious sites of activity such as neurology, neurosurgery, and neurophysiology but also traverses fields as


This chapter is also indebted to the contributions of Hannah Landecker of the Science, Technology, and Society Program at MIT, who worked with me assiduously through the conceptualization and partial drafting of earlier versions. Because she was less involved in the last stages, including the final writing, of this chapter, she, with characteristic humility and integrity, asked me not to list her as coauthor.
The Brain and the Behavioral Sciences

(only apparently) distinct as medicine, evolution, social theory, psychology, asylum management, genetics, philosophy, linguistics, anthropology, computer science, and theology.

We are also dealing with a history that challenges us to engage one of the largest questions that may be asked by historians: What has been the outcome of the effort by human beings over the past two centuries to apply the categories of scientific understanding to themselves — beings caught between a universe of social and moral realities and a universe that seems to stand outside of such realities and that they have learned to call “natural”? On all sorts of levels, our fractured understandings of ourselves meet and jostle together uneasily in this history, and any approach that fails to recognize this will in some fundamental sense miss the point.

GHOSTS AND MACHINES: DESCARTES, KANT, AND BEYOND

Questions of where to begin a story are always contested, and we have chosen to discover a “beginning” to the history of the brain and behavioral sciences in the seventeenth century, the time when the new natural philosophers of Europe had begun to converge on a model of a universe in which everything appeared capable of being accounted for in terms of matter and motion and described using the language of mathematical geometry; everything, that is, except perhaps those same philosophers themselves — those little spots of consciousness that peered through telescopes, scribbled calculations, pondered infinity, and longed for immortality, all while living inside a body that decayed, grew sick, and could be rendered dead without a moment’s warning.2

How was the scientist to understand the place of his own conscious mind in a world of matter and motion? Did his soul alone transcend the physical laws of the universe, interacting with the body (perhaps via a specific location, or special “seat”), while itself remaining untouched by the ravages of mortality and the prison cell of mechanical determinism? The notorious mind–body dualism of René Descartes — about which more ink has been spilled than can begin to be reviewed here — appeared to offer this promise.3 Descartes’ “ghost in the machine” (as Gilbert Ryle would much later famously mock it) began with a reflex model of physiology to account for most intelligent functions in humans and all intelligent functions in animals4 but then posited

4 Certainly reflex theory is invariably conventionally depicted as having its “origins” with Descartes. Georges Canguilhem has argued, however, that this is a retroactive construction of origins, which began after the establishment of mechanist theory around 1850. He credits the notion instead to
the existence in human beings alone of “something else” – a kind of pure thinking substance or rational soul that was able to move the body directly, at will, via the so-called animal spirits. The machine-body interacted with this soul, but the soul was the final authority in all volitional psychological events. “The will is so free in its nature that it can never be constrained,” Descartes asserted.

But how creditable was this idea? In the eighteenth century, the French philosopher Voltaire would ask sardonically how it was that the great Newtonian heavens conform without exception to the commands of physical law but there remains in the universe “a little creature five feet tall, acting just as he pleases, solely according to his own caprice.”

Indeed, in the early twentieth century, the philosopher Alfred North Whitehead would reflect on the incoherences inherent within that original Cartesian vision of a dualistic universe and the enduring problems they had made for future efforts to think well about human minds and the living world.

During the seventeenth century there evolved the scheme of scientific ideas which has dominated thought ever since. It involves a fundamental duality, with material on the one hand, and on the other hand mind. In between there lie the concepts of life, organism, function, instantaneous reality, interaction, order of nature, which collectively form the Achilles heel of the whole system.  

Some people went on the offensive against Descartes early on. In the mid-eighteenth century, the French physician and philosopher Julien Offray de la Mettrie (1709–1751) took up arms against Descartes’ dualistic metaphysics and proposed to simply eliminate one half of the binary opposition, resolving both mind and matter into a single materialistic formula. In his soon to be notorious (Man the Machine), written in 1748, book L’Homme Machine he pushed the point: “Since all the faculties of the soul depend to such a degree on the proper organization of the brain and of the whole body that apparently they are this organization itself, the soul is clearly an enlightened machine.”

In this era, to call the soul a machine – enlightened or otherwise – meant that you believed all of its thoughts and behaviors were products of the same impersonal laws of matter and motion that had been shown by the great Isaac Newton to govern the stars and planets. For every Alexander Pope who celebrated the clarifying intellectual power of Newton’s accomplishments (“Nature and Nature’s laws lay hid in night; God said, ‘let Newton be,"

1 This quotation and the quotation from Descartes in the previous paragraph were cited in Daniel Robinson, The Enlightened Machine (New York: Columbia University Press, 1980), p. 12.
3 Julien Offrey de la Mettrie, Histoire naturelle de l’ame (A La Haye: Chez Jean Neaulme, 1745).
and all was Light"), there was a Friedrich Schiller who shuddered at the deterministic prison it appeared to make of the universe ("Like the dead stroke of the pendulum, Nature – bereft of gods – slavishly serves the law of gravity"). Good or bad, in the late eighteenth century, Immanuel Kant came forward to insist that actually, in the case of living creatures – including especially human beings – Newtonian categories of mechanistic causality fell short. To make sense of the presenting realities of life and mind, Kant said, human judgment was forced to postulate another principle of causality, which he called “natural purpose” ("Naturzwecke"). This was a form of explanation in which the working parts of an organism were to be understood in terms of the teleology or purposive functioning of the organism as a whole.8

For a time, this piece of the Kantian legacy would offer a touchstone to a new generation of researchers who aimed to find a “third way” between Cartesian theistic dualism on the one side and crude materialistic reductionism on the other. Figures such as Karl Ernst von Baer and Johannes Müller in Germany and Thomas Laycock in England worked within a naturalistic framework that historian Timothy Lenoir has characterized as “teleological mechanism” – a framework that had room for at least some of those unstable conceptual categories identified by Whitehead that since the seventeenth century had haunted the fault line between those two monoliths of our metaphysics, “mind” and “matter.”9

THE PIANO THAT PLAYS ITSELF: FROM GALL TO HELMHOLTZ

By the early nineteenth century, however, this first antireductionist science of mind, life, brain, and body would come under increasingly successful attack by a new generation of workers. The story here is complex, internally contentious, and not seamless. One strand begins at the start of the nineteenth century with the work of Franz Joseph Gall, who would become renowned (and also derided) for his system of “organology” or phrenology. This system was rooted in three fundamental principles: The brain is the organ of the mind (not an obvious proposition at the time); the brain is a composite of parts, each of which serves a distinct mental “faculty”; and the sizes of the different parts of the brain, as assessed chiefly by examining the bumps on the skull, correspond to the relative strengths of the different faculties served.10

10 Franz Josef Gall, On the Functions of the Brain and Each of Its Parts, trans. W. Lewis, 6 vols. (Boston: Marsh, Capen and Lyon, 1835). To track the further development of phrenological thinking, see J. G. Spurzheim, Phrenology or the Doctrine of the Mental Phenomena, 2nd American (Philadelphia:
Anne Harrington

Gall was certainly not the first to interest himself in the relationship between organic structure and different aspects of psychic activity in the brain. Before Gall, the philosopher-naturalist Charles Bonnet had gone so far as to declare that anyone who thoroughly understood the structure of the brain would be able to read all the thoughts passing through it “as in a book.” Bonnet, though, working in a Cartesian mode, had imagined the brain’s presumed different organs as vehicles that the immaterial soul manipulated at will, like a pianist at the keyboard. Where Gall most clearly broke from his predecessors was in his decision to eliminate this pianist, this overruling soul, and posit instead a brain composed of some thirty self-animated organs that together generated the totality of the human mind and personality. Within Gall’s system, the piano was to play itself.

Originally ridiculed in the historical literature as a pseudoscience of “bump-reading,” the past thirty years have witnessed a partial rehabilitation of phrenology, both as an approach to brain–behavior relations that primed the pump for enduring work to come and as an anticlerical and politically potent force that expressed itself in institutional sites ranging from the asylum to the popular lecture hall.¹¹

For the purposes of this chapter, however, it will suffice to emphasize a different kind of point: that Gall’s work contributed to and, even more, exemplified a spreading approach to mind–brain relations characterized by two interconnected strategic principles: (1) to break mind down to its functional building blocks is to know it, and (2) if you can ground a piece of mind in its presumed corresponding piece of brain, then you can claim it for science. This way to truth was not a necessary one (a different approach, for example, would be chosen by evolutionary biology), but it did help launch empirical programs both in the laboratory and the clinic that would prove highly productive.¹² Indeed, with the advent of new “imaging” technologies that allow one to “see” different parts of the living brain “light up” in response to tasks and stimuli, the approach is more alive than ever.

Whatever challenge Gall and his ilk offered to Christian dualistic theologies, one thing this first generation of workers rarely, if ever, seriously questioned was the Kantian insistence that living organisms need to be understood teleologically: that the characteristics of mind are not just products of cause

Lippincott, 1908). For a sense of what, at the time, represented the most trenchant critique of this approach to the brain, see J. P. M. Frouen, Phrenology Examined, trans. C. L. Meigs (Philadelphia: Hogan and Thompson, 1846).


¹² This is a point developed by Susan Leigh Star in her book Regions of the Mind: Brain Research and the Quest for Scientific Certainty (Stanford, Calif.: Stanford University Press, 1989).
The Brain and the Behavioral Sciences

but are what they are for reasons. A broader shift away from this kind of approach to mind and brain began to gather force in the late 1840s (ironically, during the same time that new evolutionary ideas were reinstating concerns with functional utility elsewhere in the life sciences). We can track the shift by following the rise and growing influence of a closely knit group of “organic physicists” working in Germany – Hermann von Helmholtz, Emil du Bois-Reymond, Ernst Brücke, and Karl Ludwig. These were men who had come of age under the influence of the “teleological mechanists” and had also together resolved to rebel against their teachers – seeking instead to build a science in which all explanations of living processes would ultimately find translation into the new causal-material understandings of the physical sciences. As these men famously put it in 1847:

[N]o other forces than the common physical-chemical ones are active within the organism. In those cases which cannot be explained by these forces, one has either to find the specific way or form of their action by means of the physical mathematical method or to assume new forces equal in dignity to the chemical-physical forces inherent in matter, reducible to the forces of attraction and repulsion.13

From the synthesis of organic substances such as urea in the laboratory, to the establishment of cell theory, to new mechanistic understandings of embryological development, a series of milestone events in the mid-nineteenth century acted together to invest the biophysicists’ cause with considerable momentum. Of all of the apparent success stories, however, none was more historically salient for the vision than the establishment in the late 1840s of the law of conservation of energy, or the first law of thermodynamics, associated especially with physiologist-turned-physicist Helmholtz.14 “The law in question,” explained Helmholtz in an 1862 popular lecture on the topic, “asserts that the quantity of force which can be brought into action in the whole of Nature is unchangeable, and can neither be increased nor diminished.” In other words, all forms of energy (mechanical, kinetic, thermal) were equivalent and could be transformed into one another. There was nothing special, nothing “extra” that was needed to understand life, including the lives and minds of human beings. As the medical physiological Rudolf Virchow put matters in 1858: “[T]he same kind of electrical process takes place in the nerve as in the telegraph line; the living body generates its warmth through combustion just as warmth is generated in the oven; starch is transformed into sugar in the plant and animal just as it is in a factory.”15

14 Hermann von Helmholtz, Über die Erhaltung der Kraft: Eine Physikalische Abhandlung (Berlin: George Reimar, 1847).
Anne Harrington

IMAGINING BUILDING BLOCKS: FROM LANGUAGE TO REFLEX

Even as the biophysicists were gaining ground from their base within the German-speaking countries, a revised vision of the brain as a collection of modular mental functions would begin to find a new life in Descartes’ birthplace, France. In the 1860s, the French neuroanatomist and anthropologist Paul Broca used certain clinico-anatomic evidence to persuade his colleagues, and much of the international scientific community, that at least one of the phrenological mental faculties – the “faculty of articulate language” – in fact had a discrete “seat” in the brain, and that this seat lay in the third frontal convolution of the (as became more clear a few years later, exclusively left) frontal lobe of the human cortex.16

There is a lot that is not obvious about Broca’s ability to turn the tide of international opinion in favor of a localizationist approach to brain function when opinion had been so solidly opposed to it for almost two generations preceding. On the face of things, the elements with which he had to work do not appear particularly auspicious: a small handful of patient cases, mostly of older people whose multiple ailments clouded the clinical presentation of speech loss, murky autopsy data that required considerable equivocation to make the evidence “come out right,” and critics standing ready with apparently more plentiful and less ambiguous counterevidence. To bring this success story into focus, a rich “contextual” reading therefore appears necessary. The language localization efforts, for example, were undertaken during a time in France when republicanism was on the rise and the monarchy and Catholic Church were on the defensive. Thus, the French neurologist Pierre Marie, at the turn of the new century, recalled how medical students in France had quickly seized on the new doctrine of localizationism because, by its materialistic radicalness and distastefulness to the older generation, it seemed to represent scientific progress, free thought, and liberal politics. In Marie’s words: “For a while, among the students, faith in localization was made part of the Republican credo.”17


The language localization success story also comes into clearer focus when one locates it inside a larger effort within French racial anthropology of the time to determine the biological bases of the “known” mental differences existing among the different races. One widely accepted assertion from this work held that members of the allegedly evolutionarily superior white European races also possessed a considerably more developed frontal area than the “primitive” nonwhite human races, who were supposed to have larger posterior brain regions. Broca’s close colleague, the French neuroanatomist Pierre Gratiolet, had gone so far as to classify the Caucasian, Mongoloid, and Negroid races in terms of their allegedly dominant brain regions: as “frontal race,” “parietal race,” and “occipital race,” respectively.

Given this, we can begin to see why Broca might have been so motivated to seek the seat of a faculty such as language (used to such stunning effect by fellow Europeans from Shakespeare to Voltaire to Goethe) in the frontal lobes. And we can also begin to appreciate the logic whereby the ultimate localization of articulate language in the frontal region of the left hemisphere alone (and the corresponding brain-based link made between language and right-handedness) would contribute to a broader discourse in which the brain’s right hemisphere became the “savage,” the “female,” the “mad,” and the “animal” side of the brain. We are here concerned with a “brain” that is functioning in part as a flexible symbolic resource, a concrete metaphor for the carrying out of a society’s moral and political work. Parenthetically, this would be no less true in the 1970s, when the “split-brain” operations, associated with the work of people such as Roger Sperry, Joseph Bogen, and Michael Gazzaniga, reopened questions about our brain’s two hemispheres and their possible different “cognitive styles” (see also the section on “Technological Imperatives”).

Now the plot thickens further. Back in the 1820s, anatomists Charles Bell and François Magendie together had demonstrated that the spinal cord was functionally dual, with the posterior nerves acting as a channel for (incoming) sensory information and the anterior nerves acting as a channel for (outgoing) motor responses. In this way, these men helped establish an apparent material basis in the nervous system for “reflex” action. An important project of the 1830s and 1840s focused on systematically extending the new sensory-motor reflex model of nervous functioning to ever higher levels of the nervous system. The cerebral cortex itself, however, had been exempted from this creeping colonization, honored as a more or less mysterious physiological

---

terrain that serviced “mental” functions. In localizing language, Broca accepted this view as much as anyone else.19

Then, in the 1870s, two German researchers, Gustav Fritsch and Eduard Hitzig, demonstrated that the cerebral cortex also plays a role in sensory-motor activity. Applying electrical currents to the brains of dogs, the two Germans were able to produce crude movements of the body, and found moreover that specific brain regions seemed responsible for specific movements.20 Now, if the cortex possessed “motor centers,” as Fritsch and Hitzig’s work suggested, then it was logical to suppose, by analogy with the workings of spinal and subcortical structures, that it possessed sensory centers as well. And indeed the effort to identify these cortical motor and sensory centers in laboratory animals dominated experimental physiology in the last three decades of the nineteenth century. Parts of this work would ultimately not only advance laboratory research agendas, but help lay the foundations for the rise of neurosurgery at the end of the century.21

But what did this kind of localization work imply for the effort to correlate mind with matter? Could it be true, as the English neurologist David Ferrier said in 1874, that “mental operations in the last analysis must be merely the subjective side of sensory-motor substrata?”22 In the 1870s, a young German psychiatrist named Carl Wernicke attempted an answer to this question in a way that also explicitly gestured back to the biophysicists’ dream of creating an explanatory language for mind and brain that looked ultimately to the explanatory languages of the physical sciences for its orientation.

This is how it all worked. Using the anatomy of sensory-motor “projections” established in the anatomy lab by his teacher Theodor Meynert, Wernicke envisioned a cortex in which the back was specialized for processing and storing sensory data, and the front consisted of motor projections and centers. Within this schema, the form of language loss associated with “Broca’s area” was reconceptualized as a “motor” deficit, while Wernicke posited a more fundamental, sensory basis for language comprehension and generation in the (posterior) temporal region of the brain. Language and rational thought were generated within this brain through hypothesized physicalist processes, whose varied forms of breakdown could be charted using paper and pencil. Sensory-motor centers were supposed to communicate

---

19 For a good introduction to the intellectual issues at stake in the reflex story, see Clarke and Jacyna, Nineteenth-Century Origins of Neuroscientific Concepts. For an analysis of the place of reflex theory within a larger set of culturally resonant debates about control, inhibition, and regulation, see Roger Smith, Inhibition: History and Meaning in the Sciences of Mind and Brain (Berkeley: University of California Press, 1992).


with one another along “association fibers,” exchanging “impressions” like so many electrical pulses along a telegraph line and combining in accordance with the established psychological “laws of association.” For a new generation, this way of thinking about the brain – parsimonious, monistic, and predictive – would feel like a coming of age. It would lay the foundations for asylum-based research into a whole slew of newly conceived discrete brain disorders (the aphasias, the agnosias, the apraxias), an effort that old-timers would later nostalgically remember as a “golden era” in the history of clinical exploration of higher brain function.

ELECTRICITY, ENERGY, AND THE NERVOUS SYSTEM
FROM GALVANI TO SHERRINGTON

Nevertheless, at the beginning of the twentieth century, the Spanish neuroanatomist Santiago Ramon y Cajal recognized a fundamental shortcoming in the localization theories of his time: “However excellent, every physiological doctrine of the brain based on localizations leaves us absolutely in the dark over the detailed mechanisms of the psychological acts.” Cajal’s histological work identifying different types of nerve cells and the geography of their connections would, on the one hand, take the localizationist project of “mapping” the nervous system to a new level. However, he recognized that knowledge of the intricate anatomy he was untangling needed to be accompanied by an understanding of the “nature of the nervous wave, the energy transformations which it brings about or suffers at the moment when it is borne.”

As early as the mid-eighteenth century, confidence had been growing that the nervous force would turn out to be electrical in nature. The larger story to be told here does more than take us into the early history of what would become electrophysiology. It also opens doors for us into a series of tangled Enlightenment and Romantic era debates about the relationship between the organic and the inorganic, man and the cosmos, and brings esoteric science and popular culture into a common conversation over the efficacy and meaning of new therapeutic practices that began to circulate under the name of “animal magnetism” or mesmerism.

For our purposes, we must leave that story aside and identify a more conventional reference point in the historical record: 1791, the year that Luigi Galvani in Italy came out in print with experiments that he believed had demonstrated that the nerves contained intrinsic electricity. In this classic work, a frog’s leg was pierced and held by a brass hook through the thigh. When at rest, the foot would drop to make contact with a silver strip. On contact, a current was created, causing the leg muscles to contract and the foot to lift. This broke the current, causing the leg to drop again to the silver strip.

Galvani’s interpretation of the meaning of this experiment was challenged by his Italian colleague Alessandro Volta. Volta felt that Galvani had not demonstrated the existence of an inherent animal electricity but merely revealed the possibility of creating an electric current between dissimilar metals (the brass hook and silver strip) separated by a moist medium (the frog’s flesh). He could produce the same kind of phenomenon, he showed, using what he called an “artificial electric organ” – disks of different metals separated by pasteboard sheets soaked in brine, or the first wet-cell battery.  

Galvani’s work may not have been definitive, but others – again, with the Italians taking an early lead – would make the case more definitively. Then, in the 1840s, du Bois-Reymond clinched the case with his work illustrating “negative variation” in the nerve: action potential that generated a constant current following nerve stimulation. Du Bois-Reymond’s contemporary von Helmholtz then went on to measure the speed of neural electrical conduction and found it surprisingly slow – a mere eighty-five miles per hour. Not only had the nervous energy been domesticated inside the conceptual categories and experimental apparatus of nineteenth-century physics; it was looking positively tame.

Meanwhile, conceptualization of the matter of the cellular architecture of the nervous system was growing through the assiduous work of histologists. Camillo Golgi in the 1870s used silver staining to visualize nerve cells at newly high levels of definition, and he felt the evidence argued for a nervous system that functioned as a continuous network (the “reticular theory”). But Cajal, working at around the same time, disagreed. He thought the microscopic evidence showed that nerve cells were not linked but rather were discrete entities that communicated with one another by some yet to be determined process (the “neuronal theory”). Cajal’s view would win the day, and it would provide a foundation for relating the anatomy and physiology of the nervous system.

---

Anne Harrington

system in new, more integrated ways. Suddenly, one could begin to see how electrical messages passing through the physical architecture of the nervous system might be purposefully directed, diverted, inhibited, and augmented at different neuronal junctions, like a train having its course set and reset at various railroad switch points.27

The potential of neuronal theory began to be realized early in the twentieth century with the work of physiologist Charles Sherrington. Working with dogs, Sherrington aimed to map the complex pathway taken by an electrical nerve impulse as it moved from a sensory receptor on the periphery (in this case, a tactile receptor on the skin) into the spinal cord and brain, and back out over a motor pathway to produce a response (scratching). These studies led him to a way of thinking that emphasized how reflex action at one level of the nervous system could modify (stimulate or inhibit) reflex action at another level.28 These processes were understood to result from interactions between electrical impulses and modulatory chemical signals emitted at individual nerve junctions (that Sherrington named “synapses”).

During these same years, in Russia, physiologist Ivan Petrovich Pavlov (1849–1936) built on these new physiologically grounded understandings of reflex in another way, highlighting a crucial distinction between what came to be called conditioned reflex actions and unconditioned reflex actions (dogs salivating in the presence of meat powder versus dogs salivating when they hear a bell that had previously been merely paired with meat powder). This work helped set the stage for the emergence of behaviorist approaches to Anglo-American and Russian psychology during the early years of the twentieth century – approaches that, ironically enough, would largely eliminate considerations of brain and biology from the experimental picture in order to focus on clarifying strategies of prediction and control of behavior.29

Yet back in England, surveying the results of a lifetime of physiological work, Sherrington had concluded that none of these new understandings of low-level nervous functioning – to which he had so fundamentally contributed – had anything to say about high-level processes such as mind and consciousness. These, he insisted – to the dismay of at least some of his colleagues – had a soul-like reality that transcended the physical. It was evident that, even in the twentieth century, data from the clinic and laboratory alone


28 Roger Smith has impressively explored the broader cultural and semantic field within which concepts of inhibition were developed and played out in physiology, psychiatry, and elsewhere and discusses Sherrington’s work in that context. See Roger Smith, _Inhibition: History and Meaning in the Sciences of Mind and Brain_ (Berkeley: University of California Press, 1992).

Anne Harrington

were not going to be sufficient to resolve ongoing debates about the final nature of our humanness.30

HAUNTED BY OUR PAST: THE BRAIN IN EVOLUTIONARY TIME

If studies of the brain in the late nineteenth century served as one important lightning rod for debates about our nature and fate as human beings, their importance certainly would be matched, if not bested, by the new evolutionary ideas associated with Charles Darwin. But how did the two traditions interact? Alfred Russel Wallace – cofounder with Darwin of the theory of evolution by natural selection – introduced a note of tension into the relationship early on by suggesting that, in fact, the human brain represented a dilemma for the new evolutionary theory because it was capable – even in “savages” – of far greater feats of intellectual prowess and acts of ethical refinement than would have been required for mere survival. It was therefore difficult to see how it could be a product of mere natural selection. It was as if, instead, the brain had been “prepared” in advance (perhaps by an “Overruling Intelligence”) in such a way as to enable the subsequent flowering of human civilization. Charles Darwin’s own comments on Wallace’s actions here deserve to be recalled I hope, he told his friend, that you have not “murdered yours and my child” too completely.31

More consistent with the secular, anticlerical temper of the day was the virtuoso 1874 lecture by Thomas Henry Huxley “On the Hypothesis that Animals are Automata and Its History,” which brought together reflex theory and evolutionary theory to argue for a shockingly modern metaphysics of mind–body relations. This “conscious automata” theory denied any efficacious place for consciousness or “free will” in human life. The view here was that consciousness simply accompanies us in our lives like “the steam-whistle which accompanies the work of a locomotive engine.”32


Meanwhile, in some quarters, the following question began to be asked: How could one begin to orient the empirical projects of the brain sciences to do better justice to the fact that the brain, too, is not just an object in space but an evolved process in time—a four-dimensional entity? A way of thinking about this problem would ultimately be found in an image of hierarchy. The British neurologist John Hughlings Jackson in the 1870s had been among the first to articulate clearly the idea that different levels of the brain might serve as a kind of archaeological record of the biological history of a species, with lower and higher levels corresponding to earlier and later phases of evolutionary development.33

But that was not all. Jackson’s temporal view of brain functioning was also predicated on the assumption that more recently evolved layers of function—in humans, associated with rational thought and moral control—were the most vulnerable ones. This meant that, in cases of shock or damage, the more refined layers broke down first, and one was then witness to a welling up of the suddenly unmasked primitive levels of brain functioning. “Dissolution” was Hughlings Jackson’s term for this cascading down the nervous system to more primitive automatic and emotional states of functioning. In an era of growing social unrest, this was a model of brain functioning destined to embed itself in larger political concerns of the day. When Jackson’s colleague Henry Maudsley imagined the unregulated “lower centres” of the brain to be “like the turbulent, aimless action of a democracy without a head,” he was only one of many to worry that outbursts of animalistic physiology might account for everything from street riots to crimes of passion.

In psychiatric asylums, ideas like these came to serve as important resources for a renewed effort to see madness as a medical disorder with a biological underpinning, and thereby to reassert the status of asylum psychiatry as a medical science, when it had been increasingly denigrated since the late nineteenth century as a mere custodial profession. What Shorter calls the era of the “first biological psychiatry”34 had a strong hereditarian orientation that came in a distinctively fatalistic flavor—biology was destiny, as the materialists had long insisted, but sick biology, “degenerate” biology, was perhaps especially so. It was not really until the 1940s that biological psychiatry would start to be identified with a slew of biological interventions, from shock treatment to surgery,35 and not until the 1960s that the current identification of biological psychiatry with pharmaceutical interventions would begin to take hold.

It is true that by the second decade of the twentieth century, especially in the United States, optimistic social engineering programs would join forces

---

both with behaviorist thinking in psychology and with an Americanized interpretation of psychoanalysis to make a strong counterargument for the capacity of proper socialization and education to ameliorate human vulnerabilities (the “mental hygiene” movement). Nevertheless, even within this new cultural setting, the older Darwinian-inspired image of mind as an unstable struggle between “higher” and “lower” levels would persist in covert ways. It would be incorporated, for example, into the psychoanalytic concept of “regression” and serve as the rationale for the psychoanalytic distinction between primary and secondary mental processes, expressed by Freud himself in the vivid image of the conscious, rational ego struggling to maintain some sort of check over the unconscious, passion-driven “id.”

Back in the more esoteric world of university laboratory research, the basic vision of a “higher” mind functioning as an inhibitory force over the “animal” below would continue to leave its imprint on emerging mid-century understandings of the brain. A high-profile laboratory program headed by John Fulton at Yale University studied hierarchical processes of inhibition and disinhibition in the brain, all conceptualized within an evolutionary framework. Building on the work of anatomist James Papez, one of the members of this laboratory team at Yale, physician and physiologist Paul MacLean, conceptualized a system of integrated subcortical brain structures that he felt acted as the “emotional” center of the brain – mediating survival-enhancing behavior, including drives to mate and care for one’s young, and acting in other respects very much like a Freudian instinct-driven unconscious. MacLean ultimately called this system the “limbic system.”

Evolutionary thinking shaped brain science thinking in a somewhat different way with the work of Harvard psychophysiologist Walter Bradford Cannon on the role of the sympathetic-adrenal system in the arousal processes associated with the “fight or flight” emotions (especially rage and fear). Cannon saw this part of the nervous system as one half of a regulatory system (the other half would be called the “parasympathetic system”) involved in maintaining a state of responsive balance or “homeostasis” in the organism as a whole. Beginning in the late 1950s, the Cannon “fight or flight” model would be pressed into service as an organizing framework for a remarkably complex tangle of science, clinical practice, and cultural moralizing about a new psychophysiological experience called “stress” – now discovered in everything.

The Brain and the Behavioral Sciences

from monkeys who developed ulcers in laboratory settings, to “Type A” executives in corporate boardrooms ripe for heart attacks.\(^5\)

THE SUBJECT STRIKES BACK: HYSTERIA AND HOLISM

Even as everything seemed to be going well for the expansionist ambitions of the brain sciences, there were also some growing cracks in the larger citadel. The “subject,” who was to be domesticated within the current conceptual categories of brain anatomy and physiology, was in a range of ways refusing to lie down and behave the way she was supposed to.

Space permits us to do no more than gesture here in a couple of relevant directions. The first of these takes us back to the last decades of the nineteenth century, to a time when Europe’s leading neurologist, Jean-Martin Charcot, had resolved to bring the conceptual categories and clinical methodologies of neurology to elucidate the physiological logic of one of his era’s most baffling disorders: hysteria. At first, everything seemed to go well – even brilliantly. Order began to emerge out of chaos. Symptoms were cataloged, and physiological “laws” were described. Photographs were made of patients to provide the evidence Charcot needed to prove – as he put it – that the laws of hysteria that he had discovered were “valid for all countries, all times, all races” and “consequently universal.”

But, as things unfolded, it turned out that this was a physiology whose laws, far from being “universal,” were in the end so local that they basically only unfolded inside the walls of Charcot’s asylum, the Salpêtrière. Using hypnosis (which Charcot had also helped rehabilitate), rivals of Charcot showed that one could reproduce all the symptoms of hysteria, and one could also change them or make them disappear. As this came out little by little, Charcot became a target of ridicule, and his disciples scattered. The entire neurological edifice of hysteria, rooted in the visible, the objective, the universal, slowly crumbled – all of its contours now chalked up to some invisible and obscure psychological process that people were beginning to call “suggestion.”

In the space of confusion and humiliation that opened up here, people such as Freud came in and reinterpreted hysteria not as a disease of the “brain” but as a disease of the “mind.” And out of this moment of choice one sees the rise of a new kind of Cartesian logic that would get variously institutionalized and elaborated through such twentieth-century distinctions as “neuroses” versus “psychoses,” “psychiatry” versus “neurology,” “talking therapies” versus “drugs,” and somatic disorders that are “all in your head” versus

somatic disorders that are “real.” We are still living today with the legacy of those institutionalized metaphysical sortings. Nowhere is this more clearly seen than in our current approaches to managing what is called “the placebo effect.” We are so convinced of the power and ubiquity of this phenomenon that we require all new drugs to be tested against dummy versions of themselves; at the same time, we are committed to seeing all placebo effects as “imaginary” or “unreal.”

At about this same time, other kinds of discontents were afoot in the neurology clinic. Particularly in the German-speaking countries, evidence was being mobilized against the diagnostically useful model of mind and brain functioning laid down by Wernicke and his generation. Much of the energy fueling the opposition drew on the anomalies and challenges raised for Wernicke’s model by the problem of “recovery” — the evidence for the brain’s capacity to heal itself. Increasingly, it would be said that the simple fact that brain-damaged people could get better over time, could regain lost speech and movement, was simply incompatible with the nineteenth-century “machine” model of the nervous system as a purely mechanical apparatus operating according to fixed laws of reflex and association. The fighting words were spoken: Machines did not repair themselves after suffering damage, and functions that “resided” in certain fixed regions of the brain could not reappear if those brain regions had been permanently destroyed. For this reason, and others, it had become clear that human beings were actually “more than machines” — enlightened or otherwise (pace la Mettrie) — and the brain and behavioral sciences of the future (these rebellious voices from the clinic declared) were going to have to take into account all the ways in which this was so.

The 1920s began also to see laboratory-based challenges to the prevailing view of the cortex as a hard-wired structure in which highly determined nerve connections and brain areas served specific functions. The failure of the American psychophysicologist Karl Lashley to find any specific site in the rat cortex where the memory (“engram”) of a learned behavior could be localized helped usher in a “new view” of the cortex dominated by principles of functional “equipotentiality” and “mass action.” In the 1930s, work on amphibians by Paul Weiss further suggested that when nerve centers to limbs were cut and rearranged, orderly coordination could nevertheless be reestablished. The brain in these years (in part also for reasons that have to do


The Brain and the Behavioral Sciences

with political and cultural congeniality) appeared to be a marvelously plastic structure. Not biology but the environment – from family life to laboratory conditioning – appeared to “call the shots” in the life of a mind and brain.41

TECHNOLOGICAL IMPERATIVES AND THE MAKING OF “NEUROSCIENCE”

That environmentalist perspective would only begin to change in the late 1950s when new projects in both the laboratory and the clinic began to argue for the relative incapacity of the brain to rewire itself after damage, and the extent to which specific functions did have a hard-wired “place” in the cortex. New technologies, new experimental paradigms, and renewed cultural openness to interpreting ambiguous data all probably contributed to this swing back toward a kind of biological determinism. One complex expression in the 1970s was the explosion of interest in so-called split-brain research and lateralized hemisphere functioning. California psychologist Roger Sperry and his colleagues had first studied epileptic patients in whom connections between the cerebral hemispheres had been severed for therapeutic reasons. It appeared that each severed hemisphere possessed a more or less independent sphere of consciousness – often the left brain literally did not know what the right was doing. Moreover, the two hemispheres responded to the environment and computed information differently: The left hemisphere was specialized for language and (some began to argue) for analytic, piecemeal thinking in general; the right hemisphere was specialized for visual-spatial information processing and (it was argued) “holistic” (creative, artistic) thinking in general. These studies not only stimulated new kinds of research into higher brain function; they also produced a (perhaps peculiarly American) cultural dialogue on the relative virtues of what was called “left brain” versus “right brain” thinking.42

Otherwise in the postwar era, technological innovation would soon drive research at least as much as theoretical preoccupation. For example, with the development of the microelectrode in the 1940s, much basic neurobiological research went to the cellular level. In the 1960s, Harvard researchers David Hubel and Torsten Wiesel used microelectrodes to record activity in single nerve cells across the cellular columns of the primary visual area of the cortex (the anatomy of which had been worked out by Johns Hopkins neuroanatomist Vernon Mountcastle). They stunned the research community


42 For a useful overview of this literature and these events, see Sally Springer and Georg Deutsch, Left Brain, Right Brain (San Francisco: W. H. Freeman, 1993); Anne Harrington and G. Oepen, “‘Whole brain’ Politics and Brain Laterality Research,” Archives of European Neurology, 259 (1989), 141–3.
with their conclusions that different individual cells “saw” differently or, more precisely, had different built-in capacities to respond to visual stimuli – what they called “pattern specificity.” In other words, it seemed that the specific instructions by which the brain came to know the world were written as far down as the individual cell level.43

Beginning in the late 1980s, the dominant molecular focus in basic neurobiological research would begin to be partly overshadowed by excitement over new neuroimaging technologies that promised insights into the contributions made by specific neural structures to more global brain functioning. In the 1940s, Seymour Kety had used nitrous oxide to track changes in cerebral blood flow, suggesting that there might be ways to watch the “living brain” in action. This work was one step in a chain of technological developments that ultimately led to the anatomical views created by computer tomography (CT) and the dramatic colored brain pictures produced by positron emission technology (PET), and more recently by functional magnetic resonance imaging (fMRI). Slowly, a new sort of celebratory rhetoric, peppered with “final frontier” imagery, spread across the disciplinary culture of brain science. In the end, the secrets of mind and brain would be resolved, not through philosophical subtleties, but through new technological devices that would allow us to go and see where no man (or woman) had gone and seen before.44

Today, most brain and behavior science research is still sustained by a commitment to playing for technological high stakes and a pride in its own forward-looking identity. Brain science has “the future in its bones” (to recall the famous line of C. P. Snow),45 and it knows it. Nevertheless – more than it often likes to admit – the living flesh and blood of its practices and thinking remain fed by its discipline-divided and ethically contentious past. Despite the high hopes of multidisciplinary integration envisioned in the 1960s by Francis Schmitt’s Neurosciences Research Project (NRP) – which led to the coinage of this new word “neuroscience” – all the new projects and understandings do not map seamlessly onto one another. For example, updated notions of hard-wired localization coexist with models of the nervous system as a self-updating dynamic system of “neural nets” (work associated with such names as Gerald Edelman).46 Models of mind developed within the sanitized walls of computer science (so-called artificial intelligence) juggle uneasily against models of mind thrashed out in the less regulated worlds of primatology research and biological anthropology. Studies of the neurochemistry of the nervous system – including the discovery in the 1970s of the

endorphins, the brain’s “natural opiates” (by Solomon Snyder and Candace Pert)\(^{47}\) – have led some to question the extent to which the nervous system can even be properly said to exist as an independent entity. Perhaps instead it will need to be reconceived as part of a more complex system of inter-connected biochemical processes, including those that regulate endocrine and immune functions. In this last vision, the “mind” emerges, not just as a product of the brain, but in some sense of the entire human organism.

At the same time, ongoing political debates over possible brain-based determinants of sexual orientation, violence, intelligence, and supposed mental disorders (from depression to attention-deficit disorder) suggest that the moving horizon of brain and mind research will continue to be drawn into our society’s changing political and cultural imperatives and preoccupations. Today, as in the past, our questions about what we think it means to be human, and all the ways we think science can help us answer that question, simply feel too urgent for us to keep them separate from – even if we want to or think we should – our human lives, part of some imagined domesticated world of disinterested inquiry alone.