In brains we trust: How neuroeconomists stylize trust, the brain, and the social world
Klaassen, P.

Citation for published version (APA):
Klaassen, P. (2014). In brains we trust: How neuroeconomists stylize trust, the brain, and the social world

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Genesis and development of localizationism

I use history as a brain scientist uses a rat, cutting through it in order to follow the mechanisms that may allow me to understand at once the content of a science and its context.

Latour (1988, p.12)

The value of such a pre-idea resides neither in its inner logic nor in its ‘objective’ content as such, but solely in the heuristic significance which it has in the natural tendency of development. And there is no doubt that a fact develops step by step from this hazy proto-idea, which is neither right nor wrong.

Fleck (1979, p.23)

3.1 Introduction

In chapter 1 we saw that trust has been taken into the neuroscience laboratory in order to study it using, amongst other things, fMRI (functional magnetic resonance imaging). This invites us to ask how a social phenomenon such as trust has ended up in the brain. In different ways chapters 4 and 6 are concerned with this, but so is the present chapter. That is to say, both the type of evidence fMRI constitutes and, specifically, the role images play in this will be more thoroughly discussed in chapter 6. Before turning to this epistemologically key issue of analyzing fMRI in its capacity as an active linkage, however, in this chapter I indulge in a brief exercise of Fleckian “genealogy.” I examine the history
and pre-history of fMRI neuroscience in order to better understand how the conditions of possibility for the immense popularity of fMRI localization studies have developed from a proto-idea that long preceded any of the requirements for being realized scientifically.¹

Thus, here I address one prominent aspect of the answer to the question of how trust ended up in the neuroscience laboratory is dealt with—namely, that part of the answer that has to do with the history and philosophy of the idea that mental or psychological functions can be located in the brain. As such, this chapter serves two purposes. Firstly, it adds to the understanding of localizationist fMRI research, which is an essential ingredient of the neuroeconomics of trust. And secondly it serves to illustrate further what it means to engage in historical epistemology, Fleck-style.

With this focus on localizationism and its contemporary major technological implementation, fMRI, this chapter is not exclusively about neuroeconomics but speaks to a much broader field of research, encompassing much of cognitive neuroscience as well other parts of social neuroscience. Indeed, as far as the involvement of fMRI in neuroeconomics is concerned, neuroeconomics constitutes but one case of what is in fact a much larger phenomenon: The rise of neuroscience that was demonstrated in chapter 1 and to which the project at hand owes its significance. To a large extent this rise goes hand in hand with that of fMRI and likely has to be understood partly as a function thereof.

Moreover, the upsurge of fMRI as research technology has, with good reason, been interpreted as being tantamount to an epidemic. As an illustration, in 1990 the technology had not been invented yet; however, by October 2010 a total number of 53,662 peer-reviewed articles had been published with the keywords “fMRI”, “functional MRI” or “functional magnetic resonance imaging” (Rose & Abi-Rached 2013, p.74). In another characterization by Rose and Abi-Rached, this trend is taken to constitute a genuine “industry of visualization,” given that vital to fMRI’s success are the brain images it creates (see figure 5.3 on page 111 for an example).²

It would require much more than one single chapter to thoroughly deal with this topic, and here I will only scratch the surface of the history of ideas, technologies, and brains, as well as the scientists and philosophers, that would have to figure much more prominently if this exercise were to get the space it is entitled to. The limited space the topic is granted here can be justified insofar as much of the preliminary work has already been completed by a large number of historians, philosophers, sociologists and ethnographers. Their work allows me to achieve my own rather restricted goals to clarify the idea that psychological capacities, qualities and functions are instantiated or performed—in or by—specific, localizable parts of the brain, and to understand the implementation of this idea in fMRI research technology.³

The present chapter is structured as follows. In section 3.2 some background to the idea...
of localizationism is outlined; among other things I point out its continuous struggle with non-localizationism. This section functions as a sort of preliminary introduction to the exploration, in section 3.3, of the proto-idea of localizationism. This means that via nineteenth-century ideas and practices of phrenology, we are led first toward one major inspirational source of phrenology—Johann Gottfried Herder (1744–1803)—and then to even older ideas on the significance of the brain to human nature, as exemplified by Locke’s ideas on the brain’s role in consciousness. After focusing mainly on the history of ideas in these two sections, I will discuss a variety of technologies that have been essential in the success of localizationism (section 3.4). Then, in section 3.5 I move in earnest to the twentieth-century development of fMRI and the victory—temporary or definitive—of localizationism that it signifies. This section is dedicated to the explication of some of the basics of fMRI. (In chapter 6 I will return to fMRI to investigate this technology in its capacity as an active linkage in neuroeconomics.) Finally, in section 3.6 I will conclude with a brief summary of this history and how it relates to the Fleckian framework outlined in chapter 2.

3.2 Localization

Neuroimaging experiments such as those discussed in chapter 1 are premised first and foremost on the idea that mental or psychological functions, capacities and powers can be localized in specific parts of the brain. Intriguingly, this localizationist idea is much older than the technological means that seem to support it, such as EEG (electroencephalography), PET (positron emission tomography) and fMRI. Thus, the idea of localizationism has preceded the technologies through which the living, thinking and “acting” brain can be investigated, and which can allegedly show that localizationism is true. Moreover, although the conviction that psychological functions can be localized may be pervasive today, it does not enjoy unanimous consent within today’s neuroscience community.

Looking at the history of theories concerning the brain and the localization of functions, in 1959 it was observed that

> a swing of the pendulum tends to occur between [those theories] which hold that specific functions are controlled by specific parts of the brain, and [...] theories [...] which hold that the brain acts as a single functional unit. (Tizard 1959, p. 132)

In our age of fMRI localization, it is clear as to what side the pendulum has swung, but even today there is an alternative to the localization strategy, which is called dynamical systems theory. Neuroscientists favoring the dynamical systems approach abandon serious
hope in functional localizationism, opting instead opting for a wholesale characterization of the entire neural system (cf. Bunzl et al. 2010, p.51). However, without obvious (technological) handles on and understandings of connectivity, and with an estimate of 100 billion neurons populating any human brain, each neuron in turn having an average of 7000 synaptic connections to other neurons, the computational complexity for this approach to brain dynamics is perplexing. It is anything but self-evident that dynamical systems theory will become a vocal competitor of the “visual industry” of localizationist fMRI research anytime in the near future. Be this as it may, the challenges for today’s version of the “whole brain theory” of the structure and function of the brain do not by themselves comprehensively explain what our present concern with localization is due to. I will turn to that now.

Phrenology

As has been pointed out time after time, historical sources propagating localizationism can be traced back at least to late eighteenth-century Vienna. Since early modern times it has been widely agreed that the brain is the “seat of the soul”—not so much the soul’s materialization, as the place where soul and body come into contact with each other. But in the last decade of the eighteenth century the Viennese physician Franz Joseph Gall (1758–1828) developed the idea that the brain, and especially the cerebral cortex, should be seen not as a unitary seat of the soul but instead as a composite structure made up of a large variety of different “organs,” each having their own specific functions. According to Gall, the size of those organs corresponds to their power. But that is not all, and it is certainly not what Gall became most famous for. Instead, his fame is due to the last of the seven premises upon which he claimed his system was based; that is, the idea that

[from the genesis of the bones of the skull from infancy to the greatest age, the shape of the exterior surface of the skull is determined by the shape of the brain; therefore so far as the outer surface of the skull and the inner coincide, and no exception is made for the usual contours, particular aptitudes and tendencies can be concluded. (Gall (1798), pp.322-323, cited in Van Wyhe 2002, p.23)]

It is to this premise that we owe the well-known “phrenological maps,” one of which is shown in figure 3.1. Such maps allowed Gall to make inferences about people’s mental or psychological propensities on the basis of examinations of their skull, which could then reveal characteristics of their underlying brains. Because Gall first developed his so-called organology on the basis of his research in prisons and the asylum, among the first of his generalizations was the idea of a faculty for murder (“Würgsinn”) and a faculty for theft (“Diebsinn”) (Van Wyhe 2002, p.21). But it was not
long before Gall’s system had come to name 27 distinct psychological faculties, which purportedly were identifiable through the shape of the skull.

Figure 3.1: Frontispiece of Spurzheim’s *The Physiognomical System of Drs. Gall and Spurzheim* (1815).

After having lectured on his system for some six years in Vienna, Emperor Franz II banned Gall’s lectures because of the danger they posed. Amongst other things, it was thought that they “might lead to materialism and thereby go against the ‘first principles of morality and religion’” (Van Wyhe 2002, p.25). In the wake of the French revolution, a fear for radical politics, such as exemplified by Jacobinism, was undoubtedly part of the reason for this ban. As for Gall’s system, the ban on his lectures did it more good than bad. It gave him much publicity and motivated him to go on a lecture tour which took him to German, Danish and Dutch courts, major universities and scientific societies. Eventually, he settled
From Gall’s original formulation of localizationism, I now head in two opposite directions, one going back in time and toward the history of ideas, and one leaving Gall behind and moving toward the intertwined histories of ideas, scientific practices and technological developments.

### 3.3 Toward a proto-idea of localizationism

Gall’s novelty with regard to the localization of mental abilities and inclinations was more akin to a change of attention than to a radically new point of view. Gall located so-called inner qualities in specific parts of the cerebral cortex, but for this he built on work by the Swiss biologist Charles Bonnett (1720–1793), who in 1770 defined the brain as “an Organ of intricate composition, or rather really an Assemblage of different Organs” (Bonnet (1770), quoted in Clarke & Jacyna 1987, p.228).

Moreover, performing physiognomical readings of inner qualities from external appearances, as Gall divined them from the shape of the skull, was a common practice at the time and in fact, is something that had been coming and going ever since Greek antiquity. Lastly, Gall did not initially get his motivation from his empirical investigations—which surely he performed. Instead, as a medical student Gall had read Johann Gottfried von Herder, and it was Herder who had convinced Gall that “vital ‘powers,’ which coursed through all of Nature, needed to have material ‘organs’ through which they could interact with the material world” (Van Wyhe 2002, p.19). Herder even went so far as to state that he was certain that on the agreement of these parts will be created a valuable science to which a physiognomy based on conjecture would not easily attain. The foundations of the external form are inside; for everything has been fashioned by the organic powers operating from within outwards. (Herder (1965), p.205; cited in Clarke & Jacyna 1987, 230-231)

Here, then, we have stumbled upon something that is truly a proto-idea of what has become obvious beyond all doubt with the advent of fMRI and similar technologies: without having to open up skulls, mental properties, which arise from what allegedly is their underlying neurally specialized materialization can be investigated in living humans. Of course Herder’s formulation in itself hardly resembled anything we would now accept as scientific vindication, and at the time Herder wrote, one would have been hard pressed to develop the idea into a truly scientific, empirical hypothesis of recognizable form today. Herder’s idea, therefore, does not so much constitute a solid and inevitable foundation or source
of today’s localizationist neuroscience. But as Fleck remarks about the proto-idea behind Wassermann’s achievement of devising a blood test for syphilis, and as the epigraph to this chapter goes:

The value of such a pre-idea resides neither in its inner logic nor in its ‘objective’ content as such, but solely in the heuristic significance which it has [...]. And there is no doubt that a fact develops step by step from this hazy proto-idea, which is neither right nor wrong. (Fleck 1979, p.25)

The same holds in this case. Herder’s idea cannot be thought of as foundation or source, but it owes its status as proto-idea to the fact that Gall mobilized it as an indispensable resource in the development of his phrenological system. A lot of hard work followed, much of which was not positively inspired by Gall’s ideas. Indeed, from Gall onward, research has moved in many different directions. It has been governed by a variety of sometimes incompatible presuppositions and bound up with a host of different scientific, societal and technological developments. In the end, however, Herder’s view that “organic powers” within “work outwards” was solidified such that it is now an almost unavoidable neuroscientific fact. In other words, Herder’s idea should not be conceived of as a seed that could have developed into only one type of plant—quite the opposite. If it were not for all the work done in many, many different scientific and technological fields after him, Herder’s formulation would never have constituted anything akin to an event worthy of recognition in the history of science; it was its future that accomplished this. As argued for and exemplified by the historical epistemologies of Ludwik Fleck (1979), Michel Foucault (2002), Bruno Latour (1988), and Nikolas Rose (2010), history is hardly ever best understood on the model of chronological linearity. Thus, in chapter 4 I will illustrate how the meaning of historical events is as much in the hands of those who lived the events as it is in the hands of those who mobilized and in the process transformed them—or, for that matter, as it is in the hands of those who ignored them and allowed them to wither away. Rather than presenting a teleological or Whiggish narrative, then, I would here emphasize that our present is much more the outcome of a series of contingencies, accidents and chance encounters of a large collection of disciplines, disciples, technologies, ideas, (social) demands and intellectual challenges, than it is the inevitable outcome of a process which started with Herder’s idea. In this discontinuous genealogy, the name of Herder, as well as his above formulation, functions as a symbolic landmark, not as an ultimate foundation (cf. Abi-Rached & Rose 2010). Though contingencies will feature more extensively in the remainder of this chapter, let me point out that what matters, and why it is suitable to call Herder’s idea a proto-idea of today’s localizationist neuroscience, is that in all its otherness from today’s conception of localizationist neuroscience, it helps us understand the emergence of the historical conditions of possibility for thinking of the relationships between...
the parts of the brain and their psychological functions. Thus, the concept of proto-idea
does not suggest that what can be recognized as having developed from a proto-idea fol-
lowed from it by necessity and was inevitable. Rather, an idea such as Herder’s can be
conceived of as a proto-idea, only because of the largely contingent series of events that
happened to it.

“Brainhood” as a condition of possibility for localizationism

However, if we are to capture the development of not only the possibility of the rather spe-
cific idea and practice of localizationism, but also the more general contemporary concern
with the brain in relation to how we make sense of ourselves, we should continue a bit
further down the chronological road. Because of the latter issue’s relevance to the scientific
and social prominence which is granted to neuroscience today, it is well worth considering
this independently.

In analogy with the use of “personhood” as a term for conveying the idea that one is a
particular person, historian of science Fernando Vidal has coined the term “brainhood” to
refer to the practice of relating to ourselves and those around us in terms of our brains and
the conditions of our brains. The idea is that people have come to think that they do not
(merely) have brains, in the same sense in which we can say people have feet or have eyes,
but that they are their brains. Indeed, even if most people today have a rich repertoire
of ways of thinking about themselves, it is clear that the idea of brainhood is popular
and significant and can be found in many places (see e.g. Gazzaniga 2005, Swaab 2010);
according to Vidal the idea has grown more and more common in the West since the mid-
twentieth century.

It would seem obvious that this practice of relating to ourselves and our kin in terms of
(conditions of) our brains is a direct consequence of a string of events that took place in
the psychological and neurosciences, beginning in the 1960s, and of, especially, the brain
images (most prominently from fMRI research) that have been circulated widely through-
out both the scientific and the popular press ever since they first emerged. To name but
a small set of relevant episodes illustrating this, consider the commercial success of block-
buster psychopharmaca such as chlorpromazine for treating psychosis, selective serotonin
reuptake inhibitors such as fluoxetine (commonly known by its commercial name Prozac)
for treating depression and the psychostimulant methylphenidate (known for instance as
Ritalin) for treating ADHD (attention deficit hyperactivity disorder).

However, Vidal makes a strong case in arguing that the idea of brainhood is in fact not best
thought of as the result of all the investments in and returns from neuroscientific research
over the last half century. On the contrary, he argues that the affirmation of identity such
that we are our brains is not the result of empirical science but, instead, is a much older
(philosophical) assumption. Reversing cause and consequence, it is this old philosophical assumption that we are our brains that has provided a rationale for, and thereby constitutes part of, the conditions of possibility necessary for the sciences of the brain to flower, rather than the other way round (Vidal 2009).

In order to denote those who think of themselves along the lines of brainhood, Vidal has coined the term “cerebral subject.” And indeed Vidal shows that the cerebral subject is not at all a radically novel character who only entered the scene with the rise of the neurosciences in the 1960s. Quite the contrary: The cerebral subject is a “figure of modernity” owing to such (natural) philosophers as René Descartes (1596–1650), Thomas Willis (1622–1675) and, perhaps most prominently, John Locke (1632–1704). As Vidal writes:

In a radical philosophical innovation introduced in the second edition of his Essay Concerning Human Understanding (1694: book 2, ch. 27), Locke separated substance and personal identity. The identity of the man, he wrote, consists in ‘a participation of the same continued life, in succession vitally united to the same organized body’ (§ 6). The person, in contrast, is ‘a thinking being, that has reason and reflection, and can consider itself as itself, the same thinking thing, in different times and places’ (§ 9). (Vidal 2009, p.13)

What counts with regard to the identity of man is not so much his body, but his consciousness. Locke’s philosophical originality lies in his elaboration of the idea that bodies are just things we have rather than things we are. In order to make his case, Locke conducted many thought experiments. Among the more famous is his consideration of what it would mean to have one’s consciousness located in one’s little finger:

If my consciousness is located in my little finger, and this finger were cut off my hand, then, the philosopher claimed, ‘it is evident the little finger would be the person, the same person; and self then would have nothing to do with the rest of the body’ (§ 17). (Vidal 2009, p.13-14)

Importantly, however, Locke fully realized that consciousness is not localized in little fingers, and that it cannot exist all on its own. Thus,

[4]though Locke thought-experimented with a conscious little finger [...] he knew it was the nerves that conveyed sensory informations ‘to their Audience in the Brain, the mind’s Presence-room’ (Essay: 2.3.1). (Vidal 2009, p.14)

By first identifying personhood in terms of consciousness and subsequently identifying the brain as the mind’s presence-room, the cerebral subject is brought to life. This is what is required for the idea of brainhood, and it is this idea—that our brain is definitive of who and what we are—that has subsequently proven to be a noteworthy motivational force for
many philosophers and scientists to further our understanding of the brain. Furthermore, driven by the assumption of localizationism, discussed above, the brain sciences have come to provide all types of detail on the function and structure of the nervous system. This neuroscientific research obviously strengthened and fed into the idea of the cerebral subject but, as Vidal argues, this empirical work did not fundamentally alter the idea, let alone invent it.  

3.4 Phenomenotechnologies of localizationism—from artificial lesions to silver staining

Now let me take us in the opposite direction and move on to an era in which Gall is part of the past rather than of the future, and when technology and scientific practice are as vital as ideas are. Let me focus, in other words, on some relevant instances of what Gaston Bachelard has called “phenomenotechniques”; namely, techniques through which concepts become scientific and through which objects of research—or epistemic things—are realized, are made real (Bachelard 2002, p.70). Bachelard uses this notion to emphasize that objects of scientific research are not simply found, but rather co-constituted by technologies. As Hans-Jörg Rheinberger explains,

[The concept [of phenomenotechnique] aims at conceiving of technology not as an eventual byproduct of scientific activity, as a derivative product through which science manifests itself in society, but as constitutive of the contemporary scientific modus operandi itself. Insofar as the technological mode of action is engaged in the core of the scientific enterprise, the technological object itself acquires an epistemic function. (Rheinberger 2005, p.315)]

This notion provides a welcome supplement to Fleck’s view of style and active linkages, since Bachelard rightly emphasizes the importance of technology and does so from within a comparable philosophical scheme.  

With the brain and the self aligned and with the idea in place that the diverse psychological qualities of the self can be discerned in various parts of the cerebral cortex, the nineteenth-century science of the brain took off promisingly. Johann Spurzheim (1776–1832), who had long been Gall’s assistant, and George Combe (1788–1858), an Edinburgh lawyer who became Britain’s main spokesman of phrenology, are the most well-known of the many scholars who promoted the new science born of combining the idea that the brain consists of many functionally specialized organs with craniometry, or the measurement of skulls. Phrenology soon became institutionalized in all the obvious ways: through scientific societies, publications in scientific periodicals, books, lecture series and of course the
realization of its own journal (Van Wyhe 2004, p. 68). As such, together with mesmerism and energy science, phrenology became one of the nineteenth century’s more important cultural and ideological manifestations of naturalism, only to be displaced as its major ambassador by Charles Darwin’s evolutionary biology (cf. Van Wyhe 2004).

The true value of the event of phrenology cannot be assessed without taking into consideration the many ways in which its naturalism was implicated in for instance novel divisions of expertise, medical reform, secularization and racial science, as so many good histories have taught us. However, for the sake of brevity the various ways in which the esoteric and the exoteric were mangled in phrenological practice will not get center stage here.

What cannot be denied is that, in spite of its long-continuing popularity among the laity (especially in the United States), phrenology had already come into disrepute by the 1830s. This may have been partly the result of its association with showmanship—Spurzheim, Gall and other phrenologists did not refrain from spectacular performances—and fortune-telling, which however the more serious phrenologists did not partake in. More importantly, though, were probably Pierre Flourens’ “ablutions,” or artificial lesion studies, performed on animals, which purportedly falsified phrenology’s core conviction that the cerebral cortex contains many highly functionally specialized parts, each responsible for a certain type of behavior.

While Gall’s findings were based solely on correlations from real life or naturalistic observations, Flourens’ views were based on meticulous experimentations (Young 1970, pp. 58-63). Explicitly going against Gall’s localizationist thesis, Flourens asserted that “[u]nity is the great reigning principle” (Flourens (1824), quoted in Tizard 1959, p. 133). Thus he concisely captured the central idea of what have been called (aggregate) field or diffusionist theories of the brain: such theories are premised on the conviction that the brain does not function with a “pointillist division of labor” (Star 1989, p. 4), but should instead be seen as a unitary system, perhaps even in concert with the body it is part of and the (social) surroundings the person whose brain it is lives in (cf. Tizard 1959). Even though Flourens distinguished several different parts of the brain that could be related to somewhat confined functions (e.g., he found that movement is governed by the cerebellum), what is much more important in the present context is that Flourens’ field theory was compatible with early modern theories about the brain that we encountered before. Specifically, these were theories that had been in vogue since the sixteenth century, and according to which, the brain should first and foremost be thought of as the “seat of the soul” or as “sensorium” (i.e., as that part of the material body where it came together with the immaterial mind or soul). The unitary immaterial spirit postulated by such theories was thought to have a similarly unitary material physiology. In accordance with this assumption, Flourens dismissed Gall’s idea of very detailed localizations of mental capacities and inclinations.
Flourens experimented mainly on pigeons, and extrapolated his results to human brains. His belief was that the brain’s two hemispheres operated as indivisible wholes that showed few functionally specialized localizations. He was also of the opinion that the hemispheres were insensible to stimulation. No matter the combined success of these views in countering phrenology, the setback they meant for localizationism did not last long.

First, the Frenchman Paul Broca (1824–1880), who just like Gall was a physician, brain anatomist and craniometrist, gave another swing to the pendulum of localizationist versus diffusionist, or field theories, of the brain (cf. Star 1989, Tizard 1959). Based on his autopsies of patients who had suffered from impaired speech abilities, Broca in 1861 publicized his eponymous brain region, the third frontal convolution of the frontal lobe, which he claimed was implicated in speech. With that publication Broca revitalized localizationism—indeed, the idea of functional specialization of specific locales in the brain once again “became a mandate and program for subsequent researchers” (Star 1989, p.5).15

Today, any neuroscientist will still be able to point out “Broca’s area,” perhaps confirming the present reign of localization theory, or localizationism, over diffusionist theory. However, it is widely agreed upon that Broca’s research leading up to the establishment of the fact that Broca’s area was involved in speech was, at best, of dubitable quality (cf. Young 1970). Furthermore, the claim Broca became so famous for had been made before him in 1825 by Jean-Baptiste Bouillaud (1796–1881), a student of Gall’s—who, around 1800, had himself anticipated the claim. But apparently Broca’s localizationist claim came at the right time and, especially, from the right person. That is to say, it came at a time when phrenology suffered from a bad reputation, but from someone who was not closely associated with phrenology.16

In addition to Broca’s finding—or, more accurately, successful claim—a second major blow was delivered to the antilocalizationism Flourens had promoted. It was delivered by a German duo who are still very well-known in the psychological and neurosciences for the finding: physiologist and anatomist Gustav Fritsch (1838–1927) and neurologist and psychiatrist Eduard Hitzig (1839–1907). Together they demonstrated, in their experiments on dogs and monkeys, the electrical activity of the brain and, specifically, that stimulating various discrete areas of the cerebral cortex produces distinct motor reactions.

Without exaggeration it can be said that, with this finding Fritsch and Hitzig changed the rules of the game in studying the brain as well as its relationship with mind and behavior. Not only were Fritsch and Hitzig’s findings mobilized as further testimony to localizationism’s truth, their results also led, first, to the widespread application of electricity to brains in a medical context (Star 1989, p.13), and, second, to a concerted investigation of the cellular organization of the cerebral cortex.17 The theoretical supposition at work was that, given the fact that different areas of the brain were responsible for different functions,
these brain localities ought to differ from each other on a cellular level too.

Subsequent to the discoveries of Broca and Fritsch and Hitzig, a large number of neurologists and psychiatrists attached their names and fates to the attempt at furthering the localizationism that Broca had resuscitated. Among them are Wilhelm Griesinger (1817–1868), who was co-responsible for much of the necessary infrastructure requisite to the furtherance of the idea by founding a specialized journal and stimulating the institutionalization of “somaticist” psychiatry in Germany; Theodor Meynert (1833–1892), who worked in Vienna on cerebral localization by microscopically scrutinizing the neural tissue of neurological and psychiatric patients such as those suffering from Parkinson’s disease (Finger 1994, p.226); John Hughlings Jackson (1835–1911), who inferred from the patterns observable in the movements of patients during epileptic seizures a topographic organization of the cerebral cortex (Gazzaniga et al. 2009, p.3); Carl Wernicke, who picked up Broca’s theme, as he identified parts of the brain involved in language, only this time specifically having to do with the comprehension of spoken words (Gazzaniga et al. 2009, p.6); Paul Flechsig (1847–1929), who drew influential maps visualizing alleged cerebral functions in their localities and who, like Sigmund Freud (1856–1939), was a student of the influential Parisian neurologist Jean-Martin Charcot (1825–1893); and, last but not least, Sir David Ferrier (1843–1928), who created artificial lesions in monkeys such that they occasioned Charcot to comment, on seeing one of these monkeys hobbling around, “it’s a patient!” (Young 1970, p.240), and whose successive editions of The Functions of the Brain (1876, 1886) did much for the promotion of localizationism.20

The enormous efforts of this collective did so much good for localizationism that, by the beginning of the twentieth century, the idea of a functionally differentiated brain had become orthodoxy. Hence in 1902 the Encyclopedia Britannica records:

the principles of cerebral localization are, after all, only a scientific statement of matters that are of general belief. We are all more or less phrenologists

(Anon (1902), cited in Young 1970, p.240)

If anything, such a statement of general acceptance of localizationism in a medium such as the Encyclopedia Britannica is another clear symbolic landmark. We find another such landmark in the year of 1906. For in this year two new high points in articulating the findings of Broca and Fritsch and Hitzig were reached and localizationism’s victory over field theory seems to have been settled entirely. Firstly, the Nobel prize in Physiology or Medicine was awarded to the remarkable duo of Italian histologist Camillo Golgi (1843–1926) and Spaniard Santiago Ramón y Cajal (1852–1934), and secondly, Sir Charles Scott Sherrington (1857–1952) published his magnum opus, The Integrative Action of the Nervous System.
The story of Golgi and Ramón y Cajal is as dramatic as it is famous, and it has been narrated time after time in publications, not the least of which are neuroscience textbooks.\textsuperscript{21} For example, Michael Gazzaniga and colleagues (2009) relate a typical heroic account of scientific discovery in which its authors emphasize the harsh conditions under which Golgi had to work. Though Golgi had made a promising start to his academic career, in order to make a living he had to take up work as a physician in the small Lombardic town of Abbiategrasso. Not wanting to give up on his academic ambitions, Golgi persevered, and, working by the candlelight in his kitchen, he developed the most famous cell stain in the history of the world: the silver method for staining neurons—\textit{la reazione nera}, ‘the black reaction.’ (Gazzaniga et al. 2009, p.8)

The silver salt method of staining made it possible to visualize elements of the nervous system that until then had never been seen. Golgi himself, based on his observations using this method, believed the entire brain constituted one “syncytium,” a single cytoplasmic mass with several nuclei. Having improved upon Golgi’s staining method, however, Santiago Ramón y Cajal found that, in fact, the brain was made up of discrete entities (which we have come to know of as neurons), and furthermore showed that electrical transmission through neurons goes only in one direction—from dendrites to axons and never the other way around.\textsuperscript{22} This has come to be known as “the neuron doctrine,” and it is one of the most solid facts of neuroscience. Despite their ongoing disagreement, in 1906 Golgi and Ramón y Cajal shared a Nobel Prize, as mentioned above.\textsuperscript{23}

The second highpoint for localizationism was Sherrington’s 1906 publication of his \textit{The Integrative Action of the Nervous System}. Why this was such an important event for the further development of neuroscience is elaborated in Susan Leigh Star’s \textit{Regions of the Mind} (1989). Star directs attention to the existence of a historical puzzle concerning the battle between localizationism and field theory\textsuperscript{24} as it raged in the late nineteenth, early twentieth century. For instance, the work of the physician Charles-Edouard Brown-Séquard (1817–1894) raised many anomalies for localizationism, in particular anomalies of “inconstant correlation”: often the surmised relationship between some specific area of the brain and some function or trait was not supported by evidence. Such anomalies, so Star details, were acknowledged as such by the localizationist collective (see especially Star 1989, chapter 5). The work of Friedrich Goltz provides a case in point. Goltz removed areas of the brains of animals, which, for example, were responsible for walking or eating according to localizationists, and he showed that these animals kept on walking and eating without the parts of the brain “requisite” for those behaviors (Star 1989, p.122-123). Additionally, there were also more theoretical or philosophical objections to localizationism; for instance, those expressed by Mario Panizza and William James. They were of the opinion that localizationism simply could not be verified. For
the loss of a function [in absence of the allegedly responsible brain area] does not necessarily show that it is dependent on the part cut out; but its preservation does show that it is not dependent; and this is true though the loss should be observed ninety-nine times and the preservation only once in a hundred similar excisions. (James (1890, p.43), cited in Star 1989, p.124)

To illustrate the point James makes, think of a car from which the wheels are removed. The observation that a car without wheels cannot be driven, does not permit the conclusion that the “driving power” of cars is located in their wheels. And the same holds, so James argued, with respect to the relationship between mental powers and parts of the brain.

In response to the logical problems associated with localizationism and the empirical anomalies it encountered, however, localizationists did not turn to field theory. Instead, “they crafted localizationist explanations for the phenomena” (Star 1989, p.6). As Star clarifies:

Over the 36 years [from 1870 to 1906] examined here, the main direction of effort was overwhelmingly addressed not to the problem of whether the brain could be divided into areas but which areas could be located where. By concentrating on such areas, researchers often sidestepped the question of whether there were areas at all. (Star 1989, p.195)

Exemplifying this, Sherrington in his magnum opus formulated an explicitly localizationist view which was logically and theoretically indistinguishable from the diffusionist model developed by Brown-Séquard. That localizationists secured a victory, then, did not have much to do with the contents of their theoretical arguments or their accompanying models of the brain. Rather, localizationism won the day because by 1906 it was so entrenched that overthrowing it would have taken something more nearly catastrophic. No counterprogram could have found a strong institutional base because no need for it was perceived (Star 1989, pp.196-197)

Moreover, the localizationists had structured their work such that it was tightly linked to many professionals in various fields of expertise, including not only medical doctors but also (evolutionary) biologists, neurologists, psychiatrists, physiologists and neuroanatomists (Star 1989, pp.195-196). Localizationists, in other words, found (almost) all who could be either a potent allies or opponents to be on their side. This history of localizationism, then, confirms a lesson concerning the issue of how—rather than why—research programs become progressive; as Latour states in his Pasteurization of France (1988), it is by assembling a strong-enough network of allies. In this case the network included anything from model organisms to human brains, from evolutionary thought to psychiatry, as well as everything else discussed above. As Star states, “the theory of localization of function was embedded in, and indistinguishable from, scientific practice” (Star 1989, p.7). And as we saw in the previous chapter, much the same appears to be the case in neuroeconomics,
young as the field of study may be.

3.5 From proto-idea to scientific reality: or, What the above is a pre-history of

The developments of ideas and phenomenotechnologies relating to localizationism over the period from Descartes, Willis and Locke via Herder to Ferrier and Sherrington that I have recounted all properly belong to the prehistory of neuroscience (cf. Foucault 2002). I say the prehistory rather than the history, because neuroscience did not really exist as anything resembling an independent scientific discipline until the 1960s. Similarly, I traced the history of a proto-idea of localizationism. As will become apparent in later chapters, many elements of the style at work in neuroeconomics have been in the process of developing over a long period of time, though others have been firmly in place for a decades or longer.

For instance, I have shown that localizationism goes back to an eighteenth-century proto-idea, and we likewise saw that it was already realized scientifically in neurology and psychiatry early in the twentieth century. A similar development holds true for the role of evolution with regard to scientific experimentation. Even if Flourens himself believed in the fixed nature of species and was a fierce critic of Darwin’s theory of evolution (Hodge 1874), in the late nineteenth century, evolution became an ever stronger ally in actively linking conclusions drawn from lesion studies, such as Flourens performed in animal models, to man (Young 1970). This history, then, instantiates a more general lesson concerning the dynamics of science, which has been verbalized by Rheinberger: “Phenomenon and instrument, object and experience, concept and method are all engaged in a running process of mutual instruction” (Rheinberger 2010a, p.31).

In writing such history it is essential to be clear on the relevant units of analysis. From an epistemological and ontological point of view, then, the distinction between style on the one hand and discipline on the other proves vital: Stylistic elements, still very much at work today, have been long in the making, even if the discipline of neuroscience emerged relatively suddenly in the 1960s—and that of neuroeconomics only about a decade ago. Since the 1960s, significant investments have been made towards integrating the many dispersed fields of investigation that touch on the structure, anatomy and function of the brain, including its diseases and the powers, behaviors and mental disorders it is implicated in or affected by. Only since this time has the conventional disciplinary infrastructure been put in place—including journals, professional associations and university programs. And only since this time did researchers start thinking of themselves as neuroscientists—no matter how internally diverse was (and still is) the work done by the group of people who
call themselves neuroscientists.26 Despite all this, various organizing stylistic elements pre-
ceded the socio-material conditions that allowed their firm disciplinary institutionalization
in neuroscience, and later in such subfields as neuroeconomics.

This distinction between style and discipline is not always sufficiently taken into account
in reflections on neuroscience. For instance, Rose and Abi-Rached describe the emergence
of neuroscience in the 1960s as the emergence of a novel style, which they call the “neuro-
molecular style of thought” (2013, p.43).27 They describe the installment of this new style
as “an event in epistemology and ontology, in the nature of the object of the neurosciences
and the forms of knowledge that can render it in thought”—echoing Fleck’s formulation
that style can be defined as readiness for “directed perception, with corresponding mental and
objective assimilation of what has been so perceived” (Fleck 1979, p.99). Moreover, with their
claim that this is an event in epistemology and ontology, they contrast their account with
what would be (merely) an account of discipline formation (Rose & Abi-Rached 2013,
p.42). Despite their good intentions, however, Rose and Abi-Rached miss the mark here.

Apart from the fact that their characterization of the style at issue is lacking, an issue I
will address shortly, the idea of style that Rose and Abi-Rached implicitly adhere to in-
sufficiently captures the distinction between style and discipline. For when Rose and Abi-
Rached suggest that with the emergence of neuroscience in the 1960s we saw an ontological
and epistemological shift occurring, too much is made of discipline and too little of style.

Typical aspects of disciplines are confused with what is of epistemological and ontological
significance. In fact developments on the stylistic level paved the way for the emergence
of the discipline. Thus, the gradual emergence of a new style of scientific thinking and doing,
including some critical events or symbolic landmarks that occurred in the process, as I have
discussed above, constitutes a condition for the possibility of the new institutional reality
that emerged in the 1960s.

Rose and Abi-Rached deconstruct the neuromolecular style they see emerging in the
1960s into eight different “structuring principles” (2013, p.43). In regard to the elements
into which they analyze the style of neuroscience, there is not much I can find fault with.28

In brief:

1. The brain is an organ like any other.
2. Evolution has shaped the brain, and therefore one can find out about the human
   brain using animal models.
3. Basic neural processes are to be understood at the molecular level.
4. This is the level of neurotransmitters (and neuroendocrine hormones such as oxy-
  tocin) and electrical transmission of information along and between neurons.
5. The level of neurotransmitters also brings into view a host of other entities—from
ion channels to neurotransmitter metabolizing enzymes.

6. There is “localizationism”—the idea that the brain knows various functionally specialized regions, each with its own evolutionary history.

7. Everything mental takes place in/ is correlated with/ is underpinned by/ is based in/...the brain. Or, put succinctly, the mind is what the brain does.

8. Any behavior associated with mental states or processes is intrinsically related to the functioning of the brain.

Indeed many, if not all, of these elements of the style of neuroscience were in place before neuroscience was institutionalized. The institutionalization in a loose disciplinary mold that took place in the 1960s is of ontological and epistemological significance mostly because it contributed to tying these stylistic features together, “disciplining” them and those working on their behalf. What remains to be done here and now, is to relate the above genealogy to the dominant localizationist phenomenotechnology of today—fMRI. This will add to our understanding, in chapters of how in the context of neuroeconomics, localizationism actively links up with facts about trust.

**Functional magnetic resonance imaging**

To many researchers in neuroscience, fMRI is little more than a black-boxed research-enabling technology. In their eyes it is something that allows them to test and develop hypotheses regarding the localization of all sorts of mental states and processes. What exactly happens inside the box, they consider as being of little importance. However, because fMRI is probably localizationism’s strongest ally in our day, we would do better to try and understand it. That is to say, if we are to have any idea of how it is that we can come to know something about the brain by using fMRI, we at least have to lift up the lid of that black box a bit. To do that, I will briefly describe each of the component terms of the acronym fMRI: functional, magnetic, resonance and imaging.

**Functional**

The difference between magnetic resonance imaging (MRI) and functional magnetic resonance imaging (fMRI) is that, whereas with the former only static structural images are produced, the latter is used to detect almost real-time changes in metabolic values known to be correlated with brain activity; with fMRI, a temporal resolution can be achieved close to separating out neural activity at one second (though two to three seconds is more common), during experiments can last several tens of minutes. These metabolic changes occur in different areas of the brain and correspond to the (cognitive, perceptual, social) tasks performed by experimental subjects during such paradigms.
The f of fMRI is sometimes said to stand for fast (see e.g. Kevles 1997, p.197). This makes sense, because the capacity of MRI scanners to detect the changes in metabolic values, which are a major aspect of investigations of brain function, has everything to do with being able to make many different measurements in very fast succession. With fMRI 30–100 frames are captured per second, which makes it faster than traditional motion pictures and oftentimes even faster than the 60 frames per second maximally provided by today’s high-definition television standard.

Not surprisingly, even though the f comes first in fMRI, it was one of the latest extensions added to a technology that developed over a period that started around 1924. This was when Wolfgang Pauli suggested that, under specified conditions, either or both of the protons and neutrons inside atomic nuclei could “move with angular momentum, or ‘spin,’ and become magnetic” (Kevles 1997, p.176). The high point was not reached until Seiji Ogawa developed the so-called BOLD (blood oxygen level dependent) contrast imaging in the early 1990s. I will explain the distinctive traits of fMRI more thoroughly after first discussing the basics of MRI technology.

**Magnetic**

The first thing one will notice when closely approaching an MRI scanner, is the power of its magnetic field—not surprising, given that most of what you see of an MRI scanner is indeed a huge hollow, cylindrical magnet in which people can be “inserted” via a moving table. If you stand close to the opening of an MRI scanner and wear a belt with a metal buckle, you will feel as if you are being pulled towards the inside of the cylinder.

Today the strength of the magnetic fields MRI scanners produce is often 1.5 or 3 Tesla (T), but it can even be as much as 7 T, which means that the strength of the magnetic field in an MRI scanner is between 23,000 and 280,000 times the magnitude of the earth’s magnetic field. In order to make such strong magnets, liquid helium is used to cool down enormous coils, thus giving them superconducting properties. In terms of annual energy use, an MRI scanner consumes around 150 MWh of electric power, which produces 90 tons of CO₂ and costs about €33,000—even when it is not being used.

Vital to the functioning of (f)MRI is that the magnetic field produced by a scanner is completely stable and static. Why this is so, will become clear in my discussion of resonance.

**Resonance**

As mentioned, fMRI measures metabolic values, and in discussing the meaning of resonance, how it does this will become clear. To get there, however, I must first explain why MRI was previously called—in fact, is still alternatively called—NMRI, or nuclear magnetic
resonance imaging. And in explaining this, it will also become clear that MRI involves quantum mechanics.

The strong magnetic field central to MRI, and hence also to fMRI, serves to align the protons of hydrogen atoms—atoms that are part of water molecules and, thus, abound in blood. MRI therefore exploits (1) the property of human bodily tissue, including blood, to contain a lot of water and (2) the property of protons in hydrogen atoms to align with the magnetic field MRI scanners produce. Protons are subatomic particles with a positive electric charge, and belong to the collection of entities studied in nuclear physics; hence, nuclear magnetic resonance imaging.

Sociologist of science Kelly Joyce tells us that the change of names from NMRI to MRI was due to the bad reputation everything nuclear had at the time NMRI was widely introduced to (American) hospitals in the early 1980s. As John Mallard, NMRI pioneer and one of Joyce’s informants, put it: “Nuclear was associated with bombs and wars and God knows what” (Joyce 2008, pp. 41–42). Those who pushed for a change in terminology in response to this—dropping the word nuclear from the name of the technology—did not meet much resistance.

To come back to the phenomenotechnology of (f)MRI: it is the spin property of protons that is critical to (f)MRI. Spin \( s \) is an angular momentum which elementary particles, atomic nuclei and composite particles have. In brief, protons belong to the composite particles and in particle physics’ contemporary standard model they are classified as hadrons made up of three quarks, with a spin of \( \frac{1}{2} \).

When only affected by the earth’s (minimal) electromagnetic field, the orientation with which protons spin around their axis is randomly distributed. For a significant proportion of the protons in hydrogen this changes when the protons in someone’s blood enter the magnetic field of an MRI scanner. Thus, the moment characterized by the spin of protons is affected by the strong magnetic field of MRI scanners. The protons in the hydrogen atoms in the blood of subjects inside the scanner get aligned with the magnetic fields these scanners produce.

What happens in imaging experiments is that, in addition to the static magnetic field continuously produced by the MRI scanner and responsible for the gross alignment of protons, additional radio frequency magnetic fields are turned on very briefly. This is done by running pulses of electricity through the gradient coils, causing strong vibrations in those coils. As an effect of these vibrations, subjects inside the scanner experience an extremely loud pulsating noise.

The frequency of the electromagnetic field caused by the gradient coils is called the “resonance frequency”; thus, we have arrived at the point where resonance enters the story. At the roots of this part of the story are findings by Isidor Rabi (who coined the term “nu-
clear magnetic resonance”) concerning the absorption and release of energy by nuclei in response to such frequencies occurring in a static magnetic field (cf. Joyce 2008, p.25). The briefly activated fields generated in MRI are of exactly the right frequency to be absorbed by the protons, causing them to “flip” their spin. When the brief electromagnetic field is turned off again, the spins of the protons return to thermodynamic equilibrium and they are realigned with the scanner’s static magnetic field, something known through the work of Felix Bloch and Edward Purcell (Kevles 1997, p.176). This “relaxation” is accompanied by a radio frequency signal which is then measured using receiver coils. These measurements, finally, provide the data on the basis of which three-dimensional (3D) images can be constructed.

**Once more: functional**

So far, what I have described is the technique of nuclear magnetic resonance imaging, or rather, nuclear magnetic resonance. Building on the phenomenotechnique of NMR, then, MRI scanners can be tuned to measure all kinds of properties of tissues. What is distinctive of the use of MRI most common in cognitive neuroscience is that it builds on a finding by Linus Pauling and Charles Coryell from the 1930s concerning the differential paramagnetic properties of hemoglobin—that is, hemoglobin’s sensitivity to being magnetized. Hemoglobin is the oxygen carrier in the blood and therefore present in large quantities in living humans. As Pauling and Coryell found out, hemoglobin’s paramagnetic properties depend on whether it is oxygenated or deoxygenated, meaning that it is more or less magnetizable depending on whether it is carrying oxygen (to cells requiring it) or has delivered it already.16

MRI scanners, in the context of (cognitive) neuroscience, most commonly use a technique that goes by the name of BOLD-fMRI, or blood oxygen level dependent-fMRI. BOLD-fMRI builds on a finding by Seiji Ogawa from AT&T Bell Laboratories, who discovered a contrast mechanism reflecting the blood oxygen level. With remarkable insight he realized the potential importance of its application, concluding that ‘BOLD contrast adds an additional feature to magnetic resonance imaging and complements other techniques that are attempting to provide positron emission tomography-like measurements related to regional neural activity’ (Ogawa et al., 1990b). (Logothetis 2003, p.3963)

The crux of the matter is that hemoglobin carries oxygen through the bloodstream and neuronal activity entails oxygen uptake. When a certain area of the brain is active, the vascular system sends more blood to that region, thus increasing the concentration of oxygenated hemoglobin. After the blood has arrived at the right spot, the hemoglobin will be deoxygenated. Deoxygenated hemoglobin is more paramagnetic than is oxygenated
hemoglobin, so that the ratio between oxygenated and deoxygenated hemoglobin can be detected with fMRI. This is known as the BOLD effect. In general, fMRI results report the increase in the ratio of oxygenated to deoxygenated hemoglobin related to the higher amount of blood directed to active areas of the brain. However, fMRI does not so much convey the decreased ratio of oxygenated hemoglobin to deoxygenated hemoglobin due to oxygen uptake, as the neural tissue will generally not be able to absorb all the oxygen around.

This regulative process involving the demand and absorption of oxygen takes a certain amount of time. However, the time course of the blood flow is much slower than that of the neural events triggering the blood flow or “commissioning” oxygenated blood. Whereas neural events happen in milliseconds, the modulation of blood flow takes place at such a pace that the first rise is not visible until some seconds after neuronal activation and only peaks after 6 to 10 seconds. But by continually making measurements, fMRI scanners detect signals that enable the construction of three dimensional maps of changes in blood flow. And from these maps, neuronal activity is inferred.

With this we have reached the last letter of our acronym and that aspect of fMRI which the technology indubitably owes its fame to: the I for imaging.

**Imaging**

Unlike with fMRI, it is more conventional to speak of NMR than of NMRI, leaving out the I for imaging. This makes sense if one considers that the data captured through NMR are simply time series of data sets concerning three-dimensional measurements of brain tissues, based on the magnetic properties of the tissues one is interested in. In the case of fMRI these can be turned into 3D images representing the BOLD-signal, where the 3D images are composed of so-called voxels (the 3D equivalent of pixels). fMRI is commonly praised in the neuroscientific community for its relatively high spatiotemporal resolution, even if “[a] typical unfiltered fMRI voxel of 55 µl in size [...] contains 5.5 million neurons, $2.2 - 5.5 \times 10^{10}$ synapses, 22 km of dendrites and 220 km of axons” (Logothetis 2008, p.875).

More importantly though, there is nothing obvious in visualizing the data from NMR or fMR that approaches a form that may be recognized as “realistic” images of brain slices; that is, the form of images that allegedly relate to what one might optically see if one were to look with one’s eyes what underlies the data captured with NMR or (f)MRI. Reflecting on the nature of fMRI and comparable images such as those from PET, SPECT (single photon emission computed tomography) or Ultrasound, art historian Jonathan Crary described the nature of all such technologically mediated and data-processing-intensive images acutely:
the visual images no longer have any reference to the position of an observer in a ‘real,’ optically perceived world. If these images can be said to refer to anything, it is to millions of bits of electronic mathematical data. (Crary 1992, p2)

Obviously, the visual images produced with NMR or MRI do not relate to something that is also visible without NMR—these images bare no optical relation to anything whatsoever and in that sense have more the character of charts or diagrams than of pictures.  

Moreover, such values as constitute NMR primary output can, in principle, be represented in various different ways, and it took a lot of effort to develop the (mathematical) tools to turn data such as those springing from NMR into images of (brain) slices as we now know them. This can for example be read in the history of medical visualizing technologies by Bettyan Kevles (1997) or in the history and ethnography of fMRI in by Joyce (2008). As for the mathematical tools at issue, these were to a large extent made possible by a variety of scientific contributions made in the context of the development of another visualizing technology, namely computerized tomography, or CT (Kevles 1997, pp.147-153). This does not mean that, before these innovations regarding the visualization of large data sets had reached NMR research, there was no use for NMR. At that time it was used mainly by chemists, who were interested in the technology’s capacity to identify different types of molecules present in complex substances. Chemists could well do without imaging, and they tended to represent “the information [...] as lines with peaks on a graph” (Joyce 2008, p.27).

In fact, at the time when NMR was used mainly by chemists for chemical analysis and when researchers such as fMRI visionary Raymond Damadian were first experimenting with medical applications of the technology, it was not yet even possible to tell where exactly an NMR signal came from. It was only with Paul Lauterbur’s discovery that an image could be extracted from NMR data using single-line projection data [...]; Richard Ernst’s implementation of the mathematics of Fourier transforms that brought in data from two dimensions; and Peter Mansfield’s echoplanar application that a decade later [i.e., in the 1990s] led to functional, or fast, magnetic resonance. (Kevles 1997, p.186)

In the period before fMRI had been developed, but NMR had already come to be used more regularly in a medical context, both numerical representations and images existed alongside each other, because many believed that information would get lost if only the images were provided (Joyce 2008, p.35). It did not become more of a convention to represent NMR data only in the visual form that is familiar from today’s well-known brain images until NMR became more important in the medical context in the 1980s.

Even in this respect history shows some remarkable metamorphoses. Today when one
Figures 3.2: A typical fMRI image from a neuroscientific experiment.

Brain responses in the Trust Game, as depicted in Krueger et al. (2008, p. 3869).

thinks of fMRI images, one thinks of brain shapes with bright red or green blobs that represent activity levels (see figure 3.2 for an example), but even the coloring conventions have gone through changes. While at first the images were colored, radiologists insisted successfully on using greyscale images, mainly because they had been used to working with grayscales since the introduction of x-rays early in the twentieth century and CAT (or CT) scanning in the 1970s (Joyce 2008, Kevles 1997). Eventually, though, the use of colors returned.

3.6 Conclusion

Fleck suggested that style inhibits, constrains and determines what questions scientists are interested in and where scientists look for answers using which means and on the basis of which certainties. He also suggested that scientific ideas and practices are best understood when seen in the light of history. Moreover, according to Fleck, the development of or toward science often goes through several characteristic stages, including one in which “ideas appear far in advance of their rationale and independently of it” (Fleck 1979, p. 9). This is the case when what at some point is recognized as scientific fact can be traced back to a so-called pre- or proto-idea. This chapter goes a long way in confirming Fleck’s
view on this matter, using a case study of the contemporary fact that mental states and characteristics can be localized in the brain.

In my examination of the history of the general idea of localizationism, it became clear that this idea preceded any and all of the phenomenotechnologies that ostensibly provide empirical evidence in its favor. This history started with an idea, most clearly discernible in the philosophy of Locke, concerning the seat of consciousness and the essentiality of the brain to who and how we humans are—an idea captured with the term “brainhood.” I found a more precisely localizationist idea, but still very much a proto-idea in the Fleckian sense of the word, in Herder’s philosophy. This proved to be key to the work of Gall, who used it in building the renowned localizationist system known by the name of phrenology. Gall’s work, then, constitutes the first thorough attempt at developing an empirical research program along localizationist lines. However, it took until the 1990s, when brain images from PET and fMRI started circulating widely, for the idea to be realized scientifically and on a large scale.

After having discussed the pre-history of neuroscience and several phenomenotechnologies crucial to the emergence of neuroscience as we now know it, I focused attention specifically on fMRI. The epistemological significance of my detour into the details of fMRI lies in the articulation of the intricacy of this phenomenotechnology. We saw that, all by itself, fMRI technology recruits a large number of research traditions and strata of reality, including of course that of quantum mechanics. In addition, the physiology of oxygen transport, such that BOLD-fMRI becomes meaningful and neuronal activity can be (indirectly) investigated, proves to be crucial to neuroscientific localizationism, as does, as we will see in chapter 6, the whole mathematical and statistical armory requisite for turning the data fMRI delivers into standardized readable, comparable and transportable images.

Now that I have introduced the object and nature of the project at hand (chapter 1), explained (chapter 2) and illustrated the philosophical framework of this undertaking, I can now direct attention specifically toward the neuroeconomics of trust. As will become clear, investigating this along Fleckian lines only adds to one’s appreciation of how important it has been for neuroeconomics to be entangled in a very rich field of scientific practices. As I will show, the piles underlying neuroeconomics are many and various, and they provide a significant contribution to the strength of the neuroeconomic structure they support.