Self-Consciousness and "Split" Brains

The Minds' I

Elizabeth Schechter

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

First Edition published in 2018

Impression: 1

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data

Data available

Library of Congress Control Number: 2017961518

ISBN 978–0–19–880965–4

Printed and bound by
CPI Group (UK) Ltd, Croydon, CR0 4YY

Contents

The prudent carries a revolver,
He bolts the door,
O'erlooking a superior spectre
More near.

Emily Dickinson, 1862

1

The Unity Puzzle

> It is worth remembering that the main reason everybody—really, just about everybody—is fascinated with, and troubled by, work in cognitive science is that it so manifestly promises or threatens to introduce alien substitutes for the everyday terms *in which we conduct our moral lives.*
>
> (Dennett, 2009, p. 235; original emphasis)

1. The Dual Brain

At least twice so far—first in the nineteenth century and then again in the twentieth—debates about our nature and identities as embodied psychological beings have taken as a primary source of puzzlement and inspiration what may be the most readily visible fact about the human brain: its dual structure.

Like our limbs and faces, but unlike many of our internal organs, our central nervous systems are bilaterally symmetric: they look the same on the right as on the left, at least with respect to their gross morphology, that is, their large-scale shapes and structures, visible without magnification. Most of the structures that make up the brain are *paired*, again just like our eyes and ears and arms: there is a caudate nucleus on the left and a caudate nucleus on the right, a left cerebellar peduncle and a right cerebellar peduncle, and so on. As one would expect, these paired structures are in general roughly symmetric in function: there is a left cortical visual system and a right cortical visual system, a left thalamic relay system and a right thalamic relay system, and so on.

The duality of the brain is nowhere more evident than at the cortical level, since the cortex is both the most outward-lying and the largest portion of the cerebrum, itself the most outward-lying and largest portion of the brain. Its size is one reason that for so long "higher" cognitive functions (self-awareness, reason, language, moral sensitivity) were believed to be basically cortical phenomena. These were the capacities that supposedly made human beings unique among animals, and the cortex was known to be unusually large in us, relative both to the size of our bodies and to the rest of the human brain. (Its relative size has since come to be seen as less psychologically and evolutionarily significant; see Herculano-Houzel, 2009.) And the unusually large cortex of the human brain is visibly doubled—that, or divided—into two half-globes of gray convolutions: the right cerebral hemisphere and the left cerebral hemisphere—the "great" cerebral hemispheres, as they were sometimes called.

The duality of the brain was what led the seventeenth-century mathematician, philosopher, and natural scientist René Descartes to propose that the point of interchange between the mind and its material organ might be the pineal gland, one of the few unpaired structures in the brain and thus one of the few that seemed a viable candidate for interacting with the essentially unified mind (see Shapiro, 2011, for discussion). In a fascinating scientific and intellectual history, Harrington (1987) notes that a number of nineteenth-century thinkers argued that the duality of the brain undermined the materialist view of the mind (for a neuropsychologist's perspective on the same history, see Bogen, 1969, 1979). Their reasoning was simple: the brain is dual, so if the mind were the brain, the mind would be dual too, but in fact the mind is unitary, so the mind is not the brain.

Materialists did not respond to this argument in one voice. One option was to accept the first two premises of the anti-materialist's argument: the brain is dual, so if the mind is the brain, then the mind is dual, too. Wedded with materialism, the conclusion is that each brain is the seat of two minds, a right hemisphere mind and a left. Wigan (1844) defended just this position in his provocatively titled tome, *The Duality of Mind: Proved by the Structure, Function, and Diseases of the Brain and by the Phenomena of Mental Derangement and Shown to Be Essential to Moral Responsibility*.

Wigan's was an argument for *universal mental duality*: each of us has two minds apiece. As his title makes clear, Wigan did not appeal solely to neuroanatomical considerations. Among other things, he was much impressed by case reports of persons who had lost almost an entire cerebral hemisphere and still carried on reasonably well. No one could carry on reasonably well after losing half a mind—whatever that would even mean. What, then, are we to conclude, but that a brain is ordinarily the site of two minds? Once proposed, the dual mind hypothesis seemed poised to explain so much: there was evidence almost everywhere, in robust mental faculties and psychic instability alike. Over the course of the next century, the notion that we all have two minds or, better, two selves, would be used, in some hands, to explain everything from hypnosis to homosexuality.

Most of Wigan's materialist contemporaries did not accept his conclusion, however, but instead rejected one of the anti-materialist's premises. The mind is unified, they said, because the brain itself is unified, via the great white band of fibers that visibly connect the two cerebral hemispheres (see Figure 1.1).

Different structures and areas of the brain are connected by white matter tracts. Some of these pathways are physically direct, running from one structure to another without passing through and synapsing at an intermediate third. Others are quite indirect. Structures on opposite sides of the brain are connected via white matter pathways that themselves cross the midline of the brain. Some of these crossings occur via *decussations*, crossing the midline while ascending or descending to another level. Others are *commissures*, crossing the midline without ascent or descent, and generally connecting homologous (equivalent) structures. The *corpus callosum* is the largest of these commissures. It is in fact the largest fiber tract in the human brain.

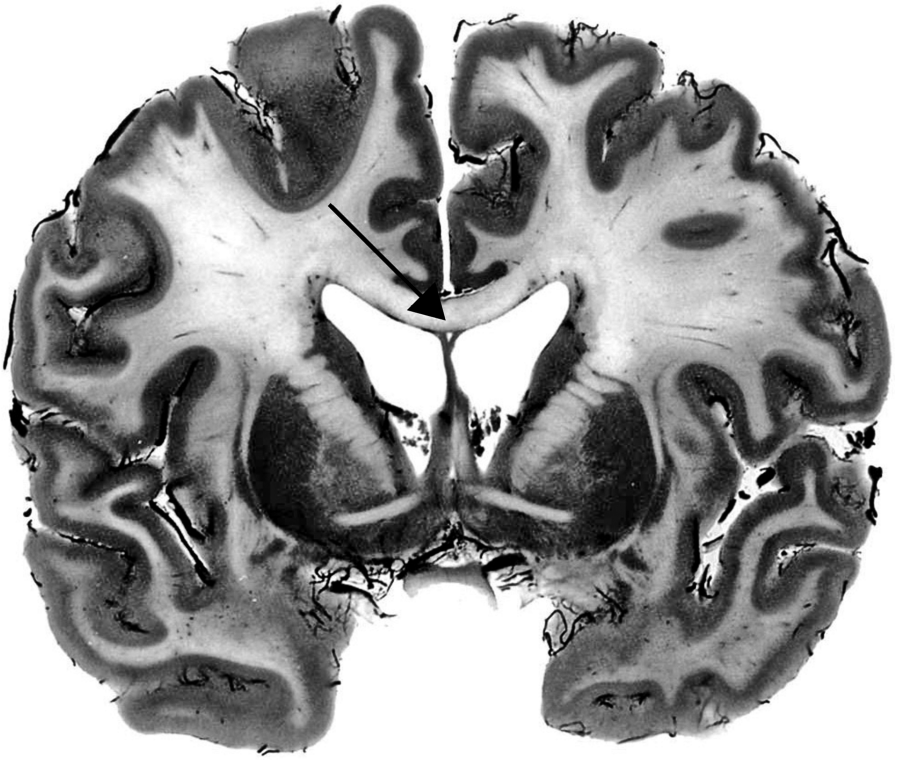

Figure 1.1 The corpus callosum. The arrow indicates the corpus callosum in a coronal section of the brain.

The corpus callosum connects mostly homologous areas of the cerebral hemispheres. It is readily visible to the naked eye and made some impression upon nineteenth-century materialists and anti-materialists alike (see e.g. Holland, 1852, p. 175). The great nineteenth-century psychologist Gustav Fechner was led to predict that if this commissure were cut, then the human soul or self would itself become two, in the same way that a worm, cut in half, simply becomes two living organisms (1860, p. 537). As Corballis notes, this hypothesis "is essentially a mechanistic one: dividing the brain should also divide the mind" (Corballis, 1998, p. 1084).

That one worm becomes two living beings when cut had struck Augustine as well: could *souls* be so divisible? (Augustine, 1947, pp. 128–9.) But worms are a special case; you can't double a herd by vivisection. Who is to say that you could divide the brain without destroying thought and consciousness altogether?

There were reported instances in which the corpus callosum had been partially destroyed by organic causes, that is, by illness or injury, and while in some cases this damage was accompanied by "psychic dysfunction," there were also a "disturbing number" of cases that seemed to show the opposite (Harrington, 1985, p. 627). Organic

brain injury makes for messy evidence, though—especially at the time, since neither the location nor the extent of the damage could be confidently determined until postmortem autopsy, however many years later. It was also possible to section the corpus callosum in non-human animals for purely experimental purposes, but they couldn't speak, making their mental lives relatively opaque to the theorist, and anyway, their minds might simply be different, more material than ours, somehow.

What was really needed to settle the matter was a controlled surgery on a sane and intelligent person.

At least one man offered to make the ultimate sacrifice for human understanding (and for his own long-standing scholarly defense of anti-materialism): the British psychologist William McDougall, author of the expansive and impassioned tract *Body and Mind* (1911), who

> tried to bargain with Sherrington . . . that if he should ever be smitten with an incurable disease, Sherrington should cut through his corpus callosum . . . "If I am right," he said, "my consciousness will remain a unitary consciousness." And he seemed to regard that as the most convincing [possible] proof of the existence of something like a soul. (quoted in Zangwill, 1974, p. 265)

No one took McDougall up on his offer, however; then as now, there was a great reluctance to cut open a living person just to settle a metaphysical bet. The great commissure connecting the great cerebral hemispheres would not be deliberately destroyed by a surgeon's knife short of medical necessity.

2. The Split-Brain Surgeries

Epilepsy is a chronic condition characterized by recurring seizures that have no other known cause. A seizure is a sudden excess of electrical activity, like a short circuit in the brain. When the disruption is confined to a localized area, it is called a *focal* seizure. In a *generalized* seizure, the activity spreads.

A *tonic-clonic* seizure is a generalized seizure. In the first phase the body clenches, becomes rigid. In the second phase it begins to jerk violently. Epilepsy characterized by this type of seizure was classically called *grand mal*—French for "great illness." A person experiencing a grand mal seizure may lose consciousness. She may be halfway across a busy intersection when it happens. She may only have a few seconds of warning, one that comes in the form of an *aura*: a visual disturbance, an auditory or olfactory hallucination, a feeling of numbness or tingling on one side of her face, or a wave of nausea, anxiety, or fear.

Seizures most often originate within a single hemisphere, but in a grand mal seizure this electrical wildfire quickly spreads to the other side, too. The corpus callosum physically bridges the two hemispheres. What if this was how the seizures crossed? And then what if you destroyed the bridge? Sparing even a single hemisphere from seizures might at least allow patients to retain consciousness during the episode, to avoid collapsing where they were.

2.1. The first split-brain surgeries

The first human beings to receive "split-brain" surgeries will not receive much attention in this book. They didn't receive much attention at the time, either.

I call them "pre-modern" split-brain subjects because we don't have their complete history. According to some reports, we don't even know how many of them there were: one epilepsy researcher first capped the number of such surgeries at twenty-seven and later clarified (or revised) this to mean that there had been twenty-seven done at least in one particular year (Williamson, 1985, 1995). Akelaitis (1945), writing at the time they were performed, reported thirty such surgeries while citing a paper (van Wagenen and Herren, 1940) that mentions only ten. Careful analysis of all the papers concerning this first set of split-brain surgeries led Bogen (1995) to believe that there were thirty-four callosal surgeries but on only twenty-eight patients.

The surgeon was Dr. van Wagenen. Years later, the sitting Chief of Neurosurgery at the University of Rochester would mention in a letter to another neuropsychologist, Joseph Bogen, that "Dr. Van Wagenen always was sorry about what he did to those patients" (cited in Bogen, 2000, p. 93). Certainly, the surgery didn't catch on, though the results van Wagenen himself reported appear to have been fairly positive: the patients suffered some irritating new symptoms, but also enjoyed some relief from their far-beyond-irritating seizures. They were extensively examined post-operatively by Akelaitis and colleagues, who published a short series of papers presenting the findings of their investigations into the psychobehavioral consequences of callosal section.

They found very little (Akelaitis, 1941). Putting a positive spin on his anti-climactic conclusions, Akelaitis wrote only that "evidence has been presented showing that interhemispheric neuronal pathways exist other than those found in the corpus callosum and anterior commissure" (1944, p. 101). No candidate pathways were proposed.

And that, presumably, was that: as van Wagenen himself said (van Wagenen and Herren, 1940, p. 741), so much for the "hallowed ground" of the corpus callosum as the "seat of the soul"! In fact, by the mid twentieth century the functional insignificance of the corpus callosum was arguably "common knowledge" (see e.g. Fessard, 1954). As one neurophysiologist joked, "I have laughingly said that, so far as I can see . . . the only demonstrable function of the corpus callosum, [is] to spread seizures from one side to the other" (McCulloch, 1949, p. 21).

2.2. The discovery of the split-brain phenomenon

In retrospect, it seems likely that van Wagenen's surgeries did not fully section the corpus callosum, which would partially explain Akelaitis's negative results. On the other hand, Akelaitis's findings were not entirely negative; he did, for instance, mention one "remarkable type of behavior which has not [yet] been described in patients with callosal lesions as far as I can discover. It consists of an apparent conflict between the desired act and the actually performed act" (1945, p. 594). For that matter, van Wagenen himself commented upon a patient who, after surgery, briefly began referring to herself in the third person (van Wagenen and Herren, 1940).

No *systematic* behavioral consequences of callosal section were revealed by clinical testing, however (van Wagenen and Herren, 1940; Akelaitis, 1944), and this was partly due to the testing methods themselves.

Akelaitis was a neurologist, trained to discern behavioral signs of mental disorder or impairment following brain injury. He looked for the signs that neuropsychologists still consider hallmarks of the *callosal disconnection syndrome*: impaired *stereognosis* (impaired capacity to recognize objects by touch), *apraxia* (difficulty with motor planning), and *anomic aphasia* (impaired ability to retrieve the words for things). The right way of looking for these signs in this particular patient population, however, had not yet been developed. In fact, you would never think to look in what turned out to be the right way unless you came armed with a specific hypothesis about what you might find:

> Experienced split-brain experimenters are often surprised when noted and unquestionably competent neuropsychologists… show themselves initially unequal to the task of testing the commissurotomy patient. The chronic disconnection syndrome is dramatic, widely known, and readily explicable. However, the arsenal developed to assess it is complex, sometimes subtle, and often based on implicit assumptions. (Zaidel et al., 1990, p. 147)

Developing and employing this "arsenal" requires an unusual conceptual shift, one that is illustrated by the first "proto" split-brain experiment ever performed on a human being. The man was Bill Jenkins, the first of the "modern" split-brain subjects, who underwent commissurotomy at the hands of Dr. Vogel in 1962. Not least because Jenkins was the first person to undergo split-brain surgery in about twenty years, Vogel's colleague Joseph Bogen was eager to run formal tests of the patient's cognitive abilities following the operation, and Jenkins was apparently "eager" to oblige (Bogen, 2006, p. 93).

Bogen writes, "Vogel shrugged off the idea. I supposed it was a result of his having learned surgery at a time when, if a brain-operated patient left the hospital speaking sensibly and walking without assistance, the operation was a success" (2006, p. 93). Even Roger W. Sperry, who had already been investigating the psychobehavioral consequences of split-brain surgeries in non-human animals, "shrugged: what theoretical preconception would be falsified?" (Bogen, 2006, p. 93). But a graduate student of Sperry's named Michael Gazzaniga was also interested, and he and Bogen eventually found a psychologist to help them administer to Bill a variety of tests measuring different aspects of intelligence.

Dr. Edwards had experience testing clinical patients, but unlike Bogen and Gazzaniga, he did not know of the split-brain animal experiments (see Section 2.4) and did not enter the room armed with a hypothesis about the psychological significance of the corpus callosum. Careful clinical testing of a split-brain subject thus once again threatened to turn up nothing of interest, until Dr. Edwards began to administer the Block Design Subtest.

The Block Design Subtest is first and foremost a test of visuospatial reasoning. The subject is given a collection of colored blocks and asked to arrange them to match a sample geometrical pattern. At first:

> Bill [Jenkins] pushed the blocks around somewhat ineffectually. Meanwhile Edwards was timing in his usual fashion and ended up [giving Bill] a zero score. I suggested that he [Bill] use one hand at a time. Dr. Edwards objected because it was customary for subjects to use both hands. However, he was persuaded to try this momentarily, so we asked Bill to use only his right hand while sitting on his left hand. *Then we asked him to do it with [only] his left hand. He had considerable success.* Mike [Gazzaniga] and I looked at each other as if we had caught a glimpse of the Holy Grail. "Now try it with just your left hand," I asked. He was quite successful! "Now try the next pattern." With his left hand he did the next one quite quickly.
>
> "No!" Edwards said, "He is supposed to use both hands." It was getting a little tense, because he insisted on doing it the standard way and we were anxious to further pursue our Grail. Dr. Edwards quietly prevailed and he finished the tests. We thanked him, and he replied, "Yes, it was interesting. We should test 20 or 30 more of these patients with various lesions." … We realized Edwards was in the dark as to what had happened and what sort of patient Bill was, or why we were so wreathed in smiles. (Bogen, 2006, p. 93; emphasis added)

Dr. Edwards saw no rationale for instructing the patient to use only one hand. He was trying to test the visuospatial abilities of Bill Jenkins, and Bill Jenkins had two hands. Joseph Bogen and Michael Gazzaniga were trying to test the visuospatial abilities of *Bill Jenkins's right hemisphere*. And the right hemisphere enjoys greater control of the *left* hand, specifically.

With Bill Jenkins, split-brain surgery was revived as a treatment for severe cases of epilepsy in the second half of the twentieth century. The split-brain subjects I will focus on in this book belong to three different patient groups. The first group (including Bill Jenkins) was operated upon by Dr. Vogel in the 1960s (the "West Coast series"); the second group was operated upon by Dr. Rayport (the "Midwest series"); the third and largest group was operated upon by Dr. Wilson in the 1970s (the "East Coast series").

There were substantial differences between the surgeries both within and across groups. Patients in the West Coast series underwent *cerebral commissurotomy*. This surgery sections the cortical or cerebral commissures: the corpus callosum, the anterior commissure, the posterior commissure, the hippocampal commissure. Patients from the Midwest series underwent *full callosotomy*, which sections the entirety of the corpus callosum (and also the hippocampal commissure; see Gazzaniga, 2000, p. 1311) but leaves intact the anterior and posterior commissures. Patients from the later East Coast series underwent either full callosotomy or only *partial callosotomy*, since it was discovered that in many patients, significant seizure reduction with fewer side effects could be achieved by a surgery that left intact not only the anterior commissure but also the posterior portion of the corpus callosum (the *splenium*). Adults in the United States now undergo at most partial callosotomy, although full callosotomy is

sometimes still performed on children in the United States and on adults elsewhere in the world. I will refer to human beings who underwent either commissurotomy or full callosotomy as *split-brain subjects*.

The term "subject" is meant to be as neutral as possible on all philosophical questions of personal identity. I'm using it, as Rupert says, with a nod to "the dry language of methodology sections… [in which] 'subject' means little more than 'warm body that can be reidentified over time'" (Rupert, 2009, p. 51). The term "split-brain patient" is more often used, but split-brain experiments are not medical treatment; the split-brain men and women we know much about are those who generously agreed to act as participants in ongoing scientific research into the psychological consequences of "splitting" the brain. I'll generally use the present tense to refer to them, though by now, many of them are deceased.

In light of the sheer size of the corpus callosum, it isn't surprising that its removal should produce the *acute callosal disconnection syndrome*, lasting days or weeks or months. Immediately after surgery, a split-brain subject may be entirely mute, then incapable of spontaneous speech, and lastly incapable of obeying verbal commands using her left hand. Or she may have trouble using her two hands cooperatively. Or she may say that she cannot control her left hand, that it won't do what she wants or that it does what she doesn't want. Still, she knows that the hand is hers (that is, she isn't delusional) and she is fully aware of her impairments (that is, she isn't *anosognosic*). Once the acute phase of the syndrome has passed, the subjects generally report that they are doing well, and their friends and family agree. Improvements in their medical condition usually outweigh the seriousness of any new, post-surgical difficulties.

Most of these difficulties are not immediately striking: a diminishment of emotional expressivity in speech, some memory loss around the edges. The only striking sign that subjects often complain of is some ongoing "trouble with the left hand": the left hand begins unbuttoning the shirt the subject is attempting to button, or the left hand pushes away the paper the subject is trying to read or grabs items from the supermarket shelf that the subject has no intention of purchasing.

On the face of it, these signs constitute a motley crew. The *systematic* psychic consequences of split-brain surgery emerge clearly only under experimental conditions.

2.3. The first split-brain experiments

Split-brain experiments have two paradigmatic features. Their basic rationale is to allow experimenters to interact with each hemisphere independently.

The first feature of the split-brain experimental paradigm is *stimulus lateralization*: sensory-perceptual information is carefully directed to only a single hemisphere at a time, to the extent possible. This feature exploits the fact that the two hemispheres are not totally symmetric with respect to their access to incoming sensory-perceptual information. In the realm of tactile perception (fine-grained touch, as could be used to identify an object by feeling it), each hemisphere receives information from only the

contralateral, that is, opposite-side hand: left hemisphere (LH) from right hand, right hemisphere (RH) from left hand. In the realm of olfaction, each hemisphere receives olfactory information from mainly the *ipsilateral*, that is, same-side, nostril: if you hold a rose before a split-brain subject's nose after sealing her right nostril, leaving only her left nostril open, olfactory information about the rose will be sent to only her LH. Each hemisphere receives auditory information from both ears, but more strongly from the contralateral ear (that is, RH from the left ear, LH from the right ear); if you simultaneously present to the two ears different complex auditory stimuli—such as a different stream of speech to each ear—each stream will be suppressed in the ipsilateral (same-side) hemisphere, which will consciously process only the speech from the contralateral ear. And there are similar stories to tell for perceptual information from other modalities, which can be lateralized differently and to different degrees in each case.

The second feature of the split-brain experimental paradigm is *response lateralization*: split-brain subjects are directed to respond to stimuli or requests either vocally or using a specific hand. This feature exploits the fact that the two hemispheres are not totally symmetric with respect to the control of movement. Each hemisphere exerts dominant motor control over the contralateral hand, so that the left hand is especially controlled by the RH and the right hand is especially controlled by the LH. In subjects in whom the left hemisphere is the so-called *dominant hemisphere*—which is most subjects—the left hemisphere is dominant for spoken language and for many (though not all) linguistic functions; indeed, split-brain subjects often cannot generate speech out of their right hemispheres at all.

The first successful split-brain experiments were performed on cats and monkeys (see e.g. Myers, 1956; Myers and Sperry, 1953; Sperry et al., 1956; Sperry, 1958). This was only after a considerable period of unsuccessful attempts. For more than a decade prior, callosotomy had been performed on non-human animals for experimental purposes and yet those animals hadn't seemed much changed after recovering from the trauma of the surgery itself: they could still see, walk, eat, learn from experience.

These first experimental callosotomies had left intact the *optic chiasm*: a midline crossing (a decussation) at which some axons from the right optic nerve pass over to the left to join the left optic tract while some axons from the left optic nerve pass over to the right to join the right optic tract (see Figure 1.2). The optic chiasm had also been experimentally sectioned in non-human animals, but in those animals the corpus callosum had been left intact. Following either kind of surgery, animals trained to make a perceptual discrimination using only one eye (the second eye forcibly closed) could nonetheless, as soon as given the opportunity, perform it perfectly using only the other eye (the first eye now forcibly closed). In other words, there was no evidence that such surgery disrupted or divided the unity of perceptual experience, learning, or action. However, as Morgan and Stellar noted (1943, pp. 449–50), no one had yet sectioned both the corpus callosum and the optic chiasm in one animal.

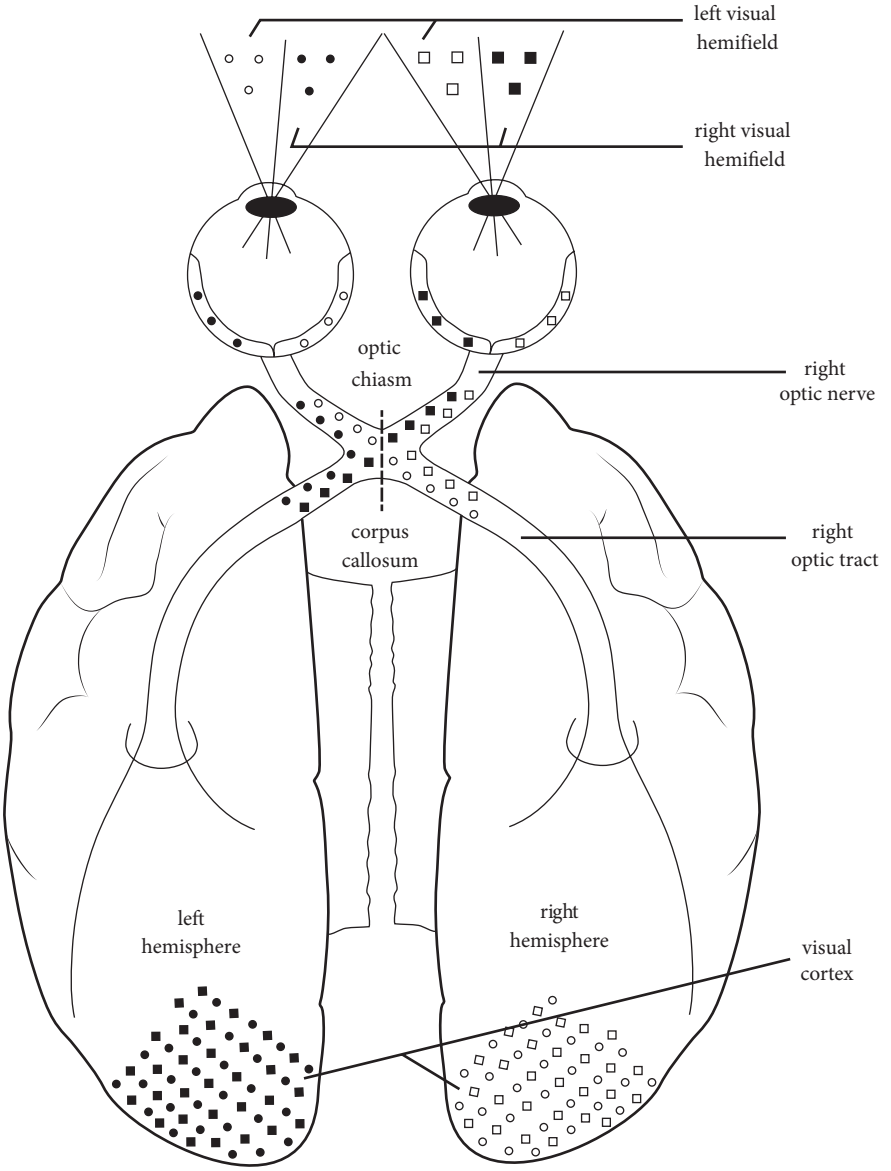

Figure 1.2 Primary visual pathways to the hemispheres. A subject's visual field is the entire region of space that they can perceive, holding constant their direction of gaze. The point at which the subject's gaze is focused, at a given moment, defines the visual midline. A subject gets the richest and most detailed visual information from the very small region of space located around the midline. Everything to the right of the midline falls within the subject's right visual hemifield; everything to the left of the midline falls within the left visual hemifield. In this figure, information from the right visual hemifield is represented by shaded symbols (squares and circles), and information from the left visual hemifield is represented by unshaded symbols (squares and circles). Information from both visual hemifields enters each eye, and information from both eyes enters each hemisphere, but information from only a single visual hemifield enters each hemisphere (right visual hemifield to the left hemisphere, left visual hemifield to the right hemisphere). Normally, the hemispheres share this information via the corpus callosum, shown sectioned here. Were the optic chiasm to be sectioned as well, each hemisphere would still receive visual information from only a single visual hemifield but would also receive visual information from only a single eye.

Sperry suggested to Myers that he do just this: perform a double surgery, a callosotomy-chiasmotomy (Evarts, 1990, p. xxi). Myers went to some trouble perfecting the technique of sectioning the (very small) optic chiasm (see Gazzaniga, 2015, pp. 42–3). It was this double surgery that finally allowed experimenters to present different visual information to the two hemispheres.

After the optic chiasm is sectioned, each hemisphere receives visual information directly from only the ipsilateral (same-side) eye. After the corpus callosum is sectioned, neither hemisphere receives visual information indirectly from the other hemisphere. With both the optic chiasm and the corpus callosum sectioned, visual information from the left eye is received by only the LH, and visual information from the right eye by only the RH. So, in an animal who has undergone sectioning of both the optic chiasm and the corpus callosum, a single hemisphere can be rendered effectively blind just by placing an eye patch over the ipsilateral eye.

Now the observed "disconnection effects" were striking. Among other things, an animal could be trained to make competing visual discriminations with each eye (e.g. selecting *O*s rather than *X*s with the left eye open, and *X*s rather than *O*s with the right eye open), with no interference effects, as if each eye had its own memory and volitional system. But of course no one thought that a mere *eye* had its own memories, much less an independent will. It was rather as if each *hemisphere* of these split-brain animals *had its own eye*.

Sperry would later be awarded the Nobel Prize for several aspects of his research, including his research with split-brain animals and then with human split-brain subjects. The latter research he initially collaborated on, at California Institute of Technology, with his then PhD student Michael Gazzaniga and with former Caltech postdoc and local neurosurgery resident Joseph Bogen. Relatively early on, Sperry—already eminent as an experimentalist—seems to have begun focusing on the more philosophical issues raised by the split-brain phenomenon (Bogen once called him "among the deepest, the most profound, neurothinkers" (2006, p. 87)). By comparison, the younger Gazzaniga published a mass of experimental works on various aspects of split-brain psychology, with Bogen perhaps intermediate between the two in his mix of energies and attentions.

2.4. Working with split-brain human beings

The experimental paradigm developed by Myers and Sperry to test split-brain non-human animals needed to be modified to test human split-brain subjects. (Much of the credit for these modifications seems to belong to Gazzaniga, though this is made clear not by Gazzaniga himself, in his own recent autobiography (see e.g. Gazzaniga, 2015, pp. 42–4), but rather by others, including Bogen, 2006, and Corballis, 2015.)

Epilepsy presents no medical need to section the optic chiasm, and thus a mere eye patch will not suffice to lateralize visual input, since each hemisphere receives visual information from both eyes. Instead, the various methods that were developed to lateralize visual information for human subjects exploit the fact that each hemisphere

directly receives visual information from only one side of the visual field (see Figure 1.2). The RH receives visual information from only the *left visual hemifield*, and the LH receives it from only the *right visual hemifield* (I am simplifying; see Chapter 5).

One interesting requirement of working with human split-brain subjects is the extra attention needed to prevent them from engaging in *cross-cueing*. Cross-cueing is the interhemispheric communication of information subserved by the subjects' own movements. Many cross-cueing behaviors are automatically deployed by both human split-brain subjects and non-human split-brain animals: for instance, moving one's eyes to fixate on something glanced suddenly to one's left, thus bringing a portion of it into the right visual field also. On the other hand, some cross-cueing behaviors look like strategic attempts to circumvent the lateralization of stimuli. Suppose for instance that an experimenter wants to learn whether semantic meaning can still transfer between the two hemispheres, via some kind of intact non-cortical route. So she presents a word, say, "CAT," in a split-brain subject's left hemifield, so that it is sent to only his RH, and then asks him to indicate what he saw by naming it, thus using his LH. If the experimenter is not careful, the subject might try to use the finger of his left hand, controlled by his right hemisphere, to trace the letters "C-A-T" on the back of his right hand—which the left hemisphere can feel (Sperry et al., 1979).

The special features of human split-brain experiments are intended to permit or preserve stimulus lateralization and response lateralization. Such lateralization allows neuropsychologists to *ask questions of* and *receive answers from* each hemisphere individually.

Imagine a split-brain subject, S, visually focusing on some point on a screen marked by a dot or a cross. Anything visually presented to the left of this *central fixation point* will be sent directly to his right hemisphere (RH), and anything visually presented to the right of this point will be sent directly to his left hemisphere (LH). Now, this is also true of subjects who have two cerebral hemispheres *and* a corpus callosum, such as me and (I assume) you—but the corpus callosum of a "non-split" subject allows either hemisphere to gain indirect access to the information that only a single hemisphere initially received. In a split-brain subject, however, the information remains largely confined to the hemisphere that initially received it.

Suppose for instance that (as described in Sperry, 1968a) two symbols are presented onscreen simultaneously: a dollar sign in S's left visual hemifield and a question mark in his right visual hemifield (see Figure 1.3). Once the screen goes blank, you ask S what he saw, and he says he saw a question mark on the right. You ask S what was on the left and he says he didn't see anything or that he doesn't know. But now suppose you ask him to close his eyes and you give him a sheet of paper and a pencil to hold in his left hand and ask him to draw what he saw on the screen using that hand. S now draws the dollar sign, that is, the left-sided stimulus that he said he didn't see. If his eyes are still closed and you ask him what he drew, he says "without hesitation that the figure he drew was the question mark, or whatever appeared in the right half of the field" (Sperry, 1968a, p. 726). S is not crazy or blind, and once you allow him to open his eyes

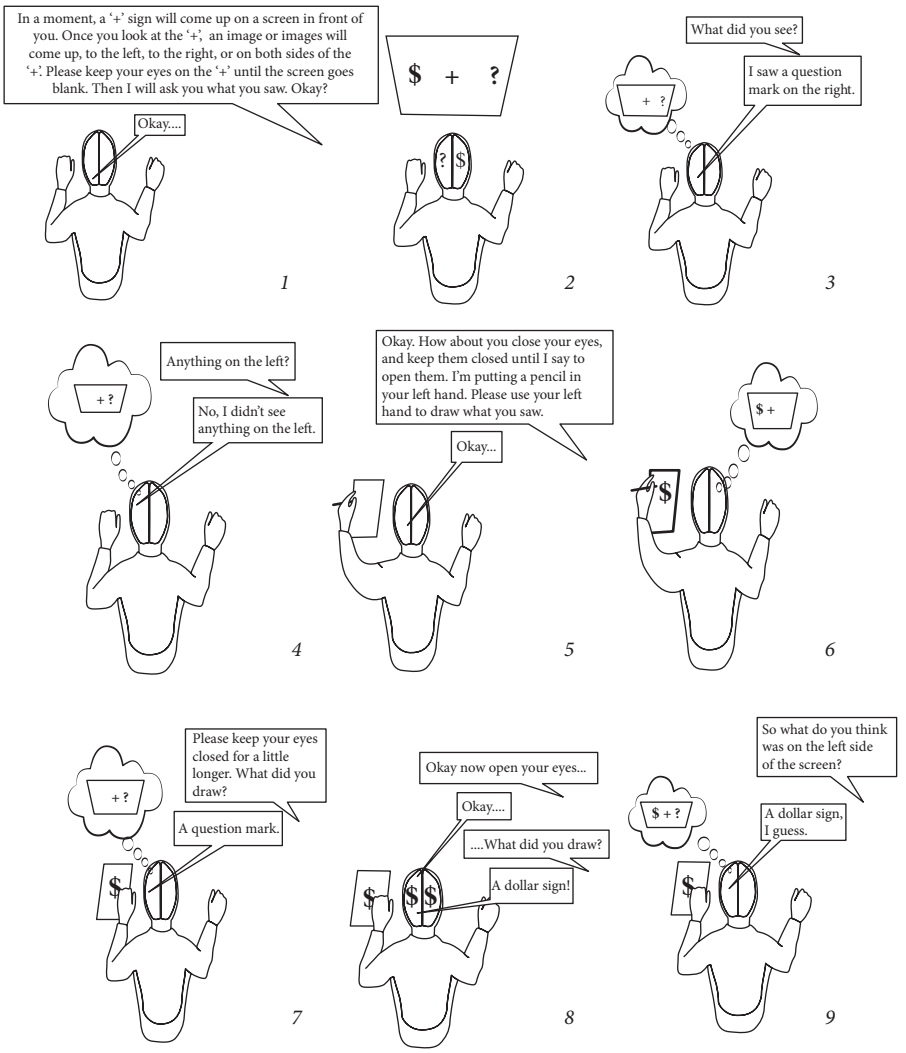

Figure 1.3 A split-brain experiment. Figure depicting a common split-brain experiment design type, with prototypical results (illustrating an experiment described in Sperry, 1968a). The second panel shows the subject seated in front of a screen. The "+" sign represents the central fixation point on which the subject's gaze is focused, defining the midline of the subject's visual field. In this experiment a "$" is presented in the left visual hemifield and a "?" in the right visual hemifield. When the subject is asked to indicate verbally what he saw, he names the "?" that was presented in the visual hemifield of the LH, which in this subject also controls speech. When the subject is asked to draw what he saw using the left hand, the subject draws the "$," which was presented in the visual hemifield of the RH, whose dominant hand is the left hand.
Source: illustration by Pippin Schupbach, with permission.

he quickly corrects himself: it's a dollar sign on the sheet of paper, not a question mark. But he may say that he doesn't know why he drew it, or that he drew it by mistake—though if you ask him whether it might thus have been a dollar sign on the left side of the screen, he admits that it might.

Here is as uncontroversial an explanation of the subject's behavior as I can give. The dollar sign presented on the left was received by S's RH exclusively, while the question mark presented on the right was received by S's LH exclusively. Very often in split-brain subjects—and let's assume so in S—spoken language cannot be generated out of the RH. So when S stated what he'd seen, production of the statement originated in the LH, which had received no visual information about the dollar sign, so S mentioned only the question mark. But the RH receives visual information from the left side of visual space and (typically) exerts dominant motor control over the left hand. So when S used his left hand to draw what he'd seen, production of the drawing behavior originated in his RH, which had received visual information about the dollar sign. But the LH apparently lacks access to the RH intentions causally responsible for the drawing, so, prior to being allowed to look at the drawing, S's statement about what he drew originated in a hemisphere that still hadn't received any information about a dollar sign: S's left hemisphere. Once S was allowed to look at the drawing, however, his LH finally received visual information about the dollar sign on the sheet of paper, and so S could now provide a correct verbal identification of the symbol originally presented on the left side of the screen.

This explanation traces the causal flow of information into and through and back out of the brain more than once: from onscreen stimulus into the RH and from the RH to an action that produces a new stimulus in the form of the dollar sign on the sheet of paper, and then back into the brain, this time the LH. This kind of *information flow story* crucially relies upon hypotheses about the architecture of perception and the control of action in the brain, hypotheses supported by a range of evidence, including, at this point, the split-brain literature itself.

Here is a second explanation of the same behavior. S's right hemisphere saw the dollar sign, but the RH can't speak, and the left hemisphere, which can speak, saw only the question mark, and so that is all the LH at first mentioned. The RH can use the left hand, however, and so drew the dollar sign that it saw. But the LH didn't at first know what the RH had drawn and therefore understandably assumed that the drawing was of the question mark the LH had seen. When finally allowed to see the sheet of paper, however, the subject was startled. Or at least, his left hemisphere was startled; his right hemisphere presumably knew what *it* had drawn.

The first story sounds unobjectionable from a philosophical standpoint. The second is highly controversial. What I will call the *unity debates* about the split-brain cases center on the question of whether the second explanation offers an accurate way of conceptualizing the activities of the two hemisphere systems after split-brain surgery, or whether it is just a shamelessly anthropomorphic shorthand for a much subtler psychological reality.

In other words: can a mere hemisphere really see? Speak? Feel startled? Know what it has done?

3. The Unity Debates

No other neurological condition has attracted the same multidisciplinary attention. By the mid 1990s (at which time Ornstein, 1997, estimated that there had already been about 45,000 psychological and biomedical papers published on the split-brain phenomenon), "split-brain ideas [had] spilled over into academic fields in both the sciences and the humanities, from cognitive and developmental psychology, clinical psychology, anthropology, archeology, and even paleontology, to education, philosophy, religious studies, and theology" (Wald, 2008, p. 36).

Why? Well, for one thing, the plain isomorphism between putative mental and neural structure in the split-brain case offered a particularly vivid symbol of materialism: it's one thing to know that taking a hammer to a man's head damages him psychologically and another to believe that dividing his brain at the cortical level somehow divides *him* into two psychological beings. By this point, the secular academy had achieved confident consensus on materialism anyway, but anti-materialist theologians and Christian philosophers still took on the split-brain phenomenon as a "hard case." Michael Gazzaniga once relayed this anecdote: Sperry went to the Vatican to give an impassioned speech on the necessity of integrating science and the study of values for the survival and flourishing of humankind—and yet it seemed (to Sperry's slight frustration) that all the questions from the holy audience concerned his relatively brief mention of the split-brain cases.[1]

Within philosophy, the so-called "split-brain phenomenon" would ultimately be invoked to argue for everything from the virtues of a Kantian moral framework (Korsgaard, 1989) to the coherence of the Christian doctrine of the Trinity (Merricks, 2006). But a main reason for the philosophical interest of split-brain cases is that they appeared to raise the old specter of *self-blindness* in a contemporary context: that of the growing predominance of scientific perspectives on human beings. In the first philosophical article devoted to the split-brain cases (though they figure also in Parfit, 1971), Nagel used them to highlight what he saw as a growing tension between our ordinary ways of understanding people and the understanding of human beings "as physical systems" emerging from the developing sciences of the mind/brain (Nagel, 1971, p. 396).

This was in part a tension between disciplines. Despite the diversity within each group, philosophers and neuropsychologists generally tended toward different interpretations of the split-brain phenomenon. Many neuropsychologists seemed to accept, implicitly or explicitly, what I will call the *2-thinkers claim*. Philosophers, as always, showed little convergence, but perhaps most expressed skepticism. Some more or less

[1] M. Gazzaniga, personal communication, January 10, 2013.

explicitly accused neuropsychologists of running away with the data (Rigterink, 1980, pp. 449–50). To this, some neuropsychologists retorted that they, at least, had *had* "direct contact with the data" (Bogen, 1977, p. 282)—that is, presumably, with the split-brain subjects themselves.

At a deeper level, however, the tension the split-brain cases evoked was intrapsychic: they served as a focal point for a felt tension between two divergent and potentially incompatible perspectives on human life and human nature (Erdmann and Stover, 1991).

They evoked this tension first and foremost because the subjects themselves did not *feel* psychically split; neuropsychologists, with their ingenious experiments, seemed to adopt or assume a perspective on the subjects that the subjects themselves lacked. Worse yet, it was said that outside the lab, split-brain people were indistinguishable from anyone else. If scientific work could nonetheless reveal that each of *them* is merely a composite of multiple psychological beings, whose identity might the lab coats come for next?

Philosophical discussion of the split-brain cases has overwhelmingly concerned the topic of *psychological identity*, that is, of *personal identity* broadly construed, since philosophers have debated not just how many persons there are within or associated with a split-brain subject, but also how many minds and subjects of experience. Presumably, split-brain subjects are unitary qua animals (even if animals' psychological capacities are essential to their identities as animals; Maiese, 2016). But human beings are not unitary qua animals only; we are also, we suppose, unitary qua persons, qua responsible agents, minds or thinkers, self-knowers, subjects of experience, and so on. What I will call the "*how many?*" questions are questions such as: How many subjects of experience is a split-brain subject? How many intentional agents? How many persons?

As a group, philosophers and neuropsychologists would come to defend every one of the possible answers that Nagel (1971) laid out: that a split-brain subject has one mind vulnerable to conscious dissociations, or two conscious minds, or two minds at some times and one mind at others, and so on. Nagel himself ultimately concluded only that all such conclusions were "unacceptable for one reason or another" (1971, p. 403).

In other words, the split-brain phenomenon presents a *puzzle*: a question to which "every imaginable answer seems wrong" (Pietroski, 2000, p. 13). The question is something like this: how many of *us* is a split-brain subject? If the question seems vague or indeterminate—who and what are we, anyway?—then that is just a part of the problem.

I will call this the *unity puzzle*. It arises from two intuitions about the split-brain cases in particular plus a background assumption about persons generally. The first intuition is that a split-brain subject has two minds, or rather that each hemisphere has its own mind. The second intuition is that a split-brain subject is nonetheless a unitary (single) person.

These intuitions are in tension only against the background of some idea or assumption linking the unity of the person to the unity of mind. I will call this the "one mind per person" (1-mind/person) rule. It's in the *background* insofar as it is not

explicit in ordinary understanding or social interaction but nonetheless presupposed by much or all of it. When someone voices their decision, you don't typically expect or ask to hear a second, different decision voiced from the same mouth, and if you do hear one, you assume that the first decision has been reversed, that *one thinker* changed *his mind or her mind*, singular. Or if, at time *t1*, you attribute an intention to someone and then at *t2* see them act inconsistently with that intention, you assume either that you were wrong at *t1* about what they even then intended to do or that although you were right at *t1*, something changed by *t2*. You do not assume—as another possibility, on equal footing with either of these—that one thinker or agent within the human being did indeed intend at *t1* what you believed them to intend and that the *t2* action was the action of a second thinker or agent who had different intentions even at *t1*. Or if a friend promises something, you make a global judgment about how likely she is to honor that promise, and you don't need to make separate judgments about two beings somehow within or associated with your friend, one of whom always keeps her word and the other of whom never does. Or you delight that your friend is deservedly proud of his recent accomplishment, without worrying that—as he has some deep-seated insecurities—he may now become jealous and resentful *of himself*.

The unity puzzle thus consists of this inconsistent triad:

The duality intuition: A split-brain subject has two minds.
The unity intuition: A split-brain subject is one person.
The one mind per person (1-mind/person) rule: Each person has exactly one mind.

On the one hand, each member of the triad is independently appealing. On the other hand, they can't all be true.

Philosophers have usually responded by offering *illusion accounts* of the split-brain phenomenon, which preserve the *1-mind/person rule* by rejecting either the unity intuition or the duality intuition as illusory. Relatively few philosophers have offered a *reconciliation account*, according to which both the duality intuition (*two minds*) and the unity intuition (*one person*) are true. It is this latter sort of account that is offered in this book.

Reconciliation accounts are prima facie less appealing. Although reconciling apparently conflicting intuitions is nice when you can get it, it often comes at too high a price, since it requires making changes to the larger background against which such intuitions are in conflict. Why not think that the 1-mind/person rule is the worst possible place to make changes? At least the unity and duality intuitions only explicitly concern split-brain subjects—a special and a small population—while the 1-mind/person rule governs our understanding of persons generally. As Pietroski points out, "I want not just any coherent set of beliefs, but a coherent set of beliefs that is recognizably *mine*" (Pietroski, 2000, p. 17; original emphasis). There is no point in resolving the split-brain unity puzzle only by offering an account of personhood that is so radically unfamiliar that I myself cannot accept it.

My approach therefore won't be to develop a new way for us to think about persons but to try to bring to light certain things that we already on some level accept. I will in part appeal to certain empirical facts about split-brain psychology that have gone under-examined in the philosophical literature on the phenomenon and that I hope will explain how the 1-mind/person rule admits of at least this one exception. While potentially unique, the exception is nonetheless made possible by more general features of persons and of our psychologies: by complex social practices designed to meet needs under perpetual threat, by the resources that enable these practices, and by the assumptions that undergird them. The strategy, then, isn't simply to reject the 1-mind/person rule but to contextualize it, thereby showing that its violation in the case of some rare individuals does not threaten the substance of our basic ways of thinking about human beings as psychological beings.

4. Outline of Book

Here is how the rest of the book will go.

Chapter 2 presents the case for thinking that a split-brain subject has two subjective perspectives. This *2-perspectives claim* is entailed by three others: that there are elements of experience associated with both the right hemisphere and the left hemisphere, that these elements are unified *intra*hemispherically, and that they are disunified *inter*hemispherically. Chapter 2 also rejects challenges to the *interhemispheric disunity claim*, and explains why the truth of the 2-perspectives claim suggests that the two hemispheres are associated with distinct subjects of experience, whom I call *R* and *L*.

Because of the close connections between consciousness and agency, the truth of this latter *2-subjects claim* ultimately depends upon it being the case that R and L are distinct intentional agents. Chapter 3 defends this *2-agents claim*, and explains why it is consistent with what we know about split-brain subjects' behavior even outside of experimental conditions, in their daily lives.

Chapter 4 generalizes the defense of the 2-subjects and 2-agents claims, explaining who R and L are, and how there could be two distinct psychological beings, two *thinkers*, co-embodied as one human being. What is puzzling about the *2-thinkers claim* is that R's mental states interact with L's richly and perpetually. The argument for the 2-thinkers claim therefore appeals to different kinds of psychic interaction and independence and especially to what I call the distinction between *direct* versus *indirect* mental state interaction. What evokes the various duality intuitions are cases in which R's mental states seem able to interact with L's only *indirectly*. Chapter 4 explains why this distinction matters to the individuation of thinkers.

Chapter 5 concerns the major objection to the 2-thinkers claim, the *objection from sub-cortical structures*, according to which split-brain psychology is ultimately not substantially divided by split-brain surgery. Chapter 5 argues that remaining direct interhemispheric interaction is neither substantial nor the primary source of what

unity and normalcy we see in split-brain subjects' behavior. The best account of split-brain psychology is thus still one according to which a split-brain subject is two distinct psychological beings, albeit beings that are neither physically nor psychologically *discrete*.

Several philosophers have presented arguments to the effect that if the duality claims are true of split-brain subjects, then they must be true of non-split subjects also, since we are all unitary human beings with two cerebral hemispheres. Chapter 6 responds to and rejects these arguments, and then presents the kind of psychological commonality between non-split and split-brain subjects that is secured by the *unity of the body*.

Chapter 7 turns to self-consciousness in split-brain subjects, arguing that R and L are not just distinct thinkers but distinct thinkers of *I-thoughts* in particular. In this sense, self-consciousness is dual in split-brain subjects. On the other hand, split-brain self-consciousness operates in such a way as to make R and L very different from other pairs of self-conscious thinkers. This will become the basis of Chapter 8's argument that R and L are not distinct persons but are instead mere parts of one person. Chapter 9 turns again to the relationship between the split-brain and the non-split case, and in particular to the question of why the split-brain cases garnered so much interest and attention even from non-academic quarters. There I offer a different explanation from the one standardly offered—including in Section 3 of this chapter.

5. On Paths Not Taken

Before continuing on to defend the duality intuition in the next chapters, I want to say something about the book's central framing device: as concerning psychological identity, the individuation of psychological beings. This approach is traditional, but some might view it as outdated, or naïve—perhaps willfully so, since I myself will acknowledge in Chapter 5 that RH and LH mental states interact in some ways that make it prima facie problematic to think of them as belonging to distinct conscious minds.

So why bother?

After all, we have other options. We could pursue a *1-thinker account*, and see what sense we can make of a split-brain subject, S, having one mind that incorporates both RH and LH activities, even though they don't seem to operate, interhemispherically, in the ways that the activities of a single mind ordinarily do. Or we could seek the nearest exit, as it were, and take a *deflationary attitude* towards the "how many?" questions, one that assumes that the psychological identities of split-brain subjects are metaphysically indeterminate.

Maintaining that a split-brain subject has one mind has the advantage of conservativism—a virtue in a philosophical account of psychological identity, as I have said. It could be viewed as a curious and open-minded conservatism at that, one ready to be enlightened about the potentially *sui generis* form that the unity of mind takes in this particular instance.

It would be worthwhile to try to develop a 1-thinker account that was fully consistent with the empirical extent of mental division in split-brain subjects. Perhaps such an account would appeal to some kind of unified meta-cognitive access that wasn't access to specific or detailed contents, an account according to which S remains a single thinker in respect of having a preserved sense of how to orient herself and respond to goings-on in the world despite limited information, or in respect of being able to continually mine the contents of her own mind for increasingly elaborated information, slowly exposing both RH and LH contents to the air of full-blown, cognitively accessible conscious experience.

The ultimate possibilities for such an account are not clear in advance. The basic risk is that no matter how intellectually honest its development, it will inevitably mislead on the feature of split-brain psychology that makes it uniquely puzzling: that it involves not disunity or even simply partial unity but *two unities*. Granted, these two unities are themselves partially unified with each other. But this is a slim pass between two mountains. The picture of S as mining the contents of her own mind for increasingly elaborated information is always only one half of the picture, for the elaborated information is in fact already present to *another* consciousness. A 2-thinkers account is in this way more faithful to the large-scale structure of split-brain psychology.

The deflationary account meanwhile gets one big thing right: no split-brain subject offers a perfectly clear-cut instance of "two minds in one body." The subjects rather occupy points intermediate between the ideal case of two thinkers in one body and the ordinary case of one thinking animal.

I say that they occupy *points*, plural, because in fact split-brain subjects differ from each other with respect to their degrees of interhemispheric dissociation and integration.[2] The duality claims defended in the first half of the book may thus not be true of N.G., for instance, in whom hemispheric independence has always been less striking, though she is still far from a non-split subject in this respect, insofar as she does show certain systematic dissociations characteristic of the callosal disconnection syndrome. And there are a number of split-brain subjects about whom we simply know very little. I therefore make no claims about how representative different split-brain subjects are of either the non-split population or of the split-brain population (Myers, 1984). I'm interested in the *possibility* of one person with multiple minds. Even a single actual case would suffice. As it happens, however, I believe that the 2-thinkers claim is true of at least P.S., V.P. (also known as P.O.V. or C.Z.), J.W., L.B., and most likely C.K. Of course, even they differ from each other; for instance, L.B. is probably not as "dual" qua

[2] Many of these differences are believed to be the consequence of brain injuries suffered prior to surgery; some may be the consequence of unintentional but inevitable differences between the surgeries themselves; for the most part, they are unexplained. Interestingly, the differences extend to the bilateral coherence of resting state activity; compare Johnston et al. (2008) and Brázdil et al. (1997) to Uddin et al. (2008). Note that Uddin et al. used the split-brain subject N.G., of whom the 2-thinkers claim does not seem to be true. (See also Paul and Tyszka, 2012, and O'Reilly et al., 2012.)

thinker as were P.S. and J.W. (as noted by Seymour et al., 1994). Like them, however, he is closer to the "two minds" end of the spectrum.

Still, if there is any sort of a spectrum whatsoever, doesn't this show that the "how many minds?" question requires a false presupposition? A deflationary attitude toward the unity debates may seem particularly appealing. It seems to be the most intellectually brave, forsaking the childish comfort of whole-number-countable minds. It also seems to be the least dogmatic, most willing to accept (surgically modified) nature as it comes (Lockwood, 1989). There is plenty to debate even in a purely empirical study of split-brain psychology; however those debates turn out, why not simply leave it at "one human being that is in some ways psychologically unified and in other ways psychologically divided or dual"?

One reason is that I'm not sure that we can leave it there. We can of course refrain from saying anything further, but that does not mean we could maintain this neutrality in our own minds. It may be that we cannot help but assign psychological phenomena to psychological beings, conceived robustly enough to make a difference to our downstream understanding. If this suspicion is right, then even if we ceased to *speak* of how many minds a split-brain subject has— preferring vaguer formulations of the sort that I myself will employ in Chapters 2 and 3, like that of experiences "associated with" or actions "originating in" one hemisphere or the other—we would nonetheless inevitably *think either* of S *or* of the two hemisphere systems as minds, and which way we thought of them would color how we understood split-brain psychology. In other words, an ostensibly deflationary answer to the "how many minds?" question risks implicitly collapsing into either a 1-thinker or a 2-thinkers account after all. If this is right, then it may be better to be explicit and self-conscious about how we are conceptualizing the domain, while acknowledging that we do so provisionally and fallibly.

A more basic concern about the deflationary position is that it may discourage us from taking a hard and detailed look at the empirical literature. This may seem implausible, and it would certainly be ironic: the possibility of metaphysically indeterminate cases is entailed by reductionism about psychological beings, and reductionism ought to encourage focus on particular forms and mechanisms and degrees of unity and disunity, since at the end of the day that is all there is, if reductionism is true. And maybe at the end of looking closely we will decide that the "how many?" questions have no answers. It is not a very inspiring place to start, however. There may be lumpers and splitters, but even splitters need a unifying picture, if only to have something to pick apart.

This is a concern about pursuing a deflationary account, not a criticism of the position itself. A deflationary attitude toward the "how many minds?" question is not equivalent to an "it's all just a giant mess" dismissiveness. My own sense, however, has been that too often the latter is what the former seems to inspire. And there is so much not to dismiss here that it is worth trying to forge an order from it, whether or not the attempt succeeds.

After all, imagination outstrips reality in some ways—but sometimes it's the other way around. Even if my account of split-brain psychological identity should turn out to be a castle in the air, it will have been built from real and important materials, ones that could be made to serve other philosophical purposes. The materials are different kinds of psychic unity and disunity, and, as the reductionist says, they are what we are.