

ABSTRACT

Title of Document: HOW MANY MINDS? INDIVIDUATING
 MENTAL TOKENS IN THE SPLIT-BRAIN
 SUBJECT

Elizabeth Schechter, Ph.D., 2009

Directed By: Professor Peter Carruthers, Department of
 Philosophy

The “split-brain” cases raise numerous difficult and fascinating questions: questions about our own self-knowledge, about the limits of introspection and phenomenology, about personal identity, and about the nature of consciousness and of mind. While the phenomenon is therefore of relevance to many areas of psychological inquiry, my dissertation explores the split-brain studies from the perspective of theoretical psychology.

The dissertation uses the split-brain cases to develop criteria for individuating mental tokens, and then applies those criteria back to the split-brain subjects themselves, ultimately arguing that split-brain subjects have two minds and two streams of consciousness apiece. Because the dissertation defends a particular account of the constitutive conditions for mental tokens against competing functionalist accounts, it also ends up being about the proper form for functionalist theories of the mental to take.

I argue throughout that psychofunctionalists who are realists about mental phenomena should accept that the constitutive conditions for mental tokens are partly neural. In particular I argue that, within an organism, multiple neural events that sustain mental phenomena causally independently of each other in some relevant sense cannot be identified with a unique mental token, regardless of how unified that organism's behavior may seem.

HOW MANY MINDS?
INDIVIDUATING MENTAL TOKENS IN THE SPLIT-BRAIN SUBJECT

By

Elizabeth Schechter

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2009

Advisory Committee:

Professor Peter Carruthers, Chair

Professor Georges Rey

Distinguished University Professor Jerrold Levinson

Associate Professor Allen Stairs

Distinguished University Professor Emeritus William Hodos

© Copyright by
Elizabeth Schechter
2009

Acknowledgements

Thanks to my family, for providing every kind of life support, year after year. And for the grammatical advice!

Many thanks to friends who consented to read parts of this dissertation at one point or another (i.e. Ryan Millsap, Matt King, Brendan Ritchie, Mark Engelbert, and Bénédicte Veillet). Thank you, Mark, for the reassuringly bored certainty that the disaster-tation was anything but. Thank you, Bénédicte, for sharing all the sage wisdom—and for leading the way!

Thanks most of all to members of my committee:

to Allen Stairs, for helping me balance the demands of teaching and writing,

to Bill Hodos, for the patient, outside perspective, and for general calm, experience, and good humor,

to Jerry Levinson, for several kind pep-talks over the years—every time I most needed one, like magic,

to Georges Rey, for early enthusiasm, for continued inspiration, and for helping me see the big picture (even in my own work),

and especially to Peter Carruthers, who deserves more of my gratitude than I can ever express without embarrassing both of us.

Table of Contents

Acknowledgements.....	ii
Table of Contents.....	iii
List of Figures.....	v
Chapter 1: Introduction.....	1
1 A History of the “Split” Brain.....	1
2 Minds, Brains, and Unity.....	8
2.1 Brain Bisection and the Unity of Consciousness.....	10
2.2 Five Hypotheses.....	13
3 Conclusion.....	24
Chapter 2: On Individuating Minds and Other Mental Entities.....	27
1 Introduction.....	27
2 On Mental Entities.....	29
2.1 Types, Tokens, States, and Systems.....	30
2.2 Identity and Individuation.....	34
2.3 On Kinds of Questions.....	44
3 The Criteria for Mindedness.....	48
3.1 Modest Criteria for Mindhood.....	48
3.2 Purely Psychological versus Partly Neural Criteria for Mindhood.....	51
4 Individuating Minds.....	63
4.1 Inter- versus Intra-mind Interactions.....	63
4.2 Deliberate versus Non-deliberate Interactions.....	66
4.3 Direct versus Indirect Interactions between Mental States.....	69
5 Conclusion.....	82
Chapter 3: A Defense of the Mental Duality Model.....	84
1 Introduction.....	84
2 Isomorphism and the Relationship between Neural and Psychological Facts.....	87
3 Behavior and Psychological Explanation.....	103
4 Perfect Parallelism: Does Redundancy Matter?.....	121
4.1 STR: Meaning, Motivation, and Empirical Adequacy.....	123
4.2 Constituting a Single Token.....	127
4.3 Three Claims on Irrelevance.....	132
4.4 Mental Tokens.....	145
5 On Integrating Theory and Practice.....	153
5.1 Entities and Evidence.....	154
5.2 Minds and Persons.....	162
6 Duality versus Partial Co-mentality.....	173
6.1 Partial Co-Mentality (Or, an Indeterminate Number of Minds?).....	174
6.2 Subcortical Structures: Plausible Sources of (Interhemispheric) Mental Integration?.....	179
6.3 What Kind of Interhemispheric Relations Do Non-cortical Structures Afford?.....	191
7 Conclusion.....	208

Chapter 4: On Co-consciousness and Conscious Unity.....	210
1 Introduction.....	210
2 Puccetti.....	211
3 Some Claims and Distinctions.....	216
3.1 Right Hemisphere Consciousness.....	216
3.2 Access versus Phenomenal Consciousness.....	219
3.3 Co-consciousness.....	222
3.4 A Unified Consciousness versus a Stream of Consciousness.....	226
3.5 Consciousness and Subjects of Experience.....	233
3.6 What is Co-consciousness?.....	237
3.7 Preliminary Conclusions and Assumptions.....	244
4 The Conscious Duality Model and the Challenge from Unified Behavior...	245
4.1 Duplication of Contents.....	250
4.2 Attention.....	257
4.3 Cross-cuing.....	265
4.4 Affect... and Aura?.....	267
4.5 Same Body.....	269
4.6 Same Brain.....	272
4.7 Single Hemisphere Control.....	273
4.8 Owing Actions.....	275
4.9 Single Streams versus Unified Behavior.....	278
5 Conclusion.....	281
Chapter 5: Models of Consciousness in the Split-Brain Subject.....	282
1 Introduction.....	282
2 The Partial Co-consciousness Model.....	282
2.1 A Case of Partial Co-Consciousness?.....	284
2.2 Transitivity and the Partial Co-consciousness Model.....	288
2.3 Objections to the Partial Co-consciousness Model.....	296
3 The Dynamic Singularity Model.....	312
3.1 Conscious Experience: the Input-Output Picture versus the Intentions-Content thesis.....	315
3.2 The Intentions-Content Thesis and the Dynamic Singularity Model ...	320
3.3 Motor Intentions and the Dynamic Singularity Model: Trevarthen's Case	322
3.4 Motor Intentions and the Dynamic Singularity Model: Sergent's Case	334
3.5 A Qualified Defense of the Intentions-Content Thesis.....	339
3.6 The Conscious Duality Model, Defended.....	341
4 Conclusion.....	347
Chapter 6: Functionalism and the Boundaries of the Mental.....	349
1 Introduction.....	349
2 What are the Vehicles of Consciousness?.....	353
3 Functional Roles and Fineness of Grain.....	359
4 Conclusion.....	370
Conclusion.....	376
Bibliography.....	380

List of Figures

Figure 3.1: Gross Anatomical Divisions of the Brain.....	180
Figure 5.1: Some Kinds of Conscious Structure.....	297
Figure 5.2: The Dynamic Singularity Model: Trevarthen's Case.....	327

Chapter 1: Introduction

1 A History of the “Split” Brain

The corpus callosum is the main commissure connecting the two cerebral hemispheres in placental mammals, and the largest fiber tract in the human brain. Epileptic seizures start in one hemisphere or the other; in a single individual, seizures will tend to start in the same hemisphere (right or left) again and again. A grand mal seizure is one in which abnormal electrical activity spreads from the hemisphere of origin into the other hemisphere, affecting the whole brain. Subjects in the midst of a grand mal seizure will usually lose consciousness. Over time repeated seizures may cause damage to neural tissue. And people with severe epilepsy may experience multiple seizures per day.

A *callosotomy* is surgical procedure in which the corpus callosum is severed or sectioned. In a partial callosotomy, the anterior two thirds of the callosum are sectioned; in a full callosotomy, the entire corpus callosum as well as the hippocampal commissure is sectioned; a commissurotomy entails full callosotomy plus sectioning of the anterior commissure.¹ Doctors began performing callosotomies

¹ Unless otherwise specified, I do not use the terms “callosotomy” and “split-brain surgery” to refer to cases of only partial callosotomy. I use it to refer to those who have undergone both full callosotomy and commissurotomy. Patients operated on by Bogen and Vogel in the 1960’s (the “West Coast series”) underwent commissurotomy, and those operated on by Wilson in the 1970’s (the “East Coast series”) underwent full callosotomy (although in stages), but not commissurotomy (Churchland, 1986). The term “split-brain” is generally used to refer to both callosotomy and commissurotomy subjects, and again I use it in the same way. When potentially relevant to a claim, argument, or idea, I will distinguish between subjects with intact anterior commissures and subjects in whom the anterior commissure has been sectioned.

in the 1940's on subjects with debilitating, sometimes life-threatening epilepsy. The operation apparently proved quite successful at preventing seizures from spreading from one hemisphere to the other in a majority of subjects; not infrequently it eliminated or drastically reduced seizure occurrence altogether, for reasons not fully understood. A majority of those subjects who were not helped by a callosotomy were helped by a commissurotomy. Meanwhile clinical post-operative studies of these subjects turned up only few side effects after the initial recovery period, mostly involving motor coordination, and little if any significant cognitive decline as a result of the surgery. More commonly, in fact, "split-brain" subjects experienced cognitive improvement on many measures, presumably the result of at least one hemisphere being spared further seizures. These preliminary results, suggesting that sectioning the callosum—again, the largest fiber tract in the human brain—has remarkably few negative effects, led one neurophysiologist to joke, in 1949, "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: 21).

At about this same time Ronald Myers and Roger Sperry were studying the results of callosotomy on cats (e.g., Myers 1955; Myers and Sperry, 1953; Sperry, Stamm, and Miner, 1956). Callosotomy had of course been tested on animals prior to performing it on human subjects. Myers' and Sperry's experiments were novel because in addition to sectioning the callosum of each cat they also sectioned its optic chiasm, limiting each hemisphere's visual information to that it could receive from a single (ipsilateral) eye. A simple eye-patch could then easily be used to effectively (yet temporarily) "blind" an entire hemisphere. The observed "disconnection effects"

were fascinating. Among other things, the animals could be trained to make competing visual discriminations with each eye, with no evident interference effects—as if the two surgically separated cerebral hemispheres embodied functionally independent perceptual, learning, memory, and action-guidance systems.

Sperry and others formed two, non-competing hypotheses for why human callosotomized or “split-brain” subjects seemed relatively normal after such major brain surgery, despite the important cognitive role they now had strong reasons to believe the callosum should play in human (not just feline) subjects. First, in humans (and similarly in other mammals), what is seen with a single eye is sent to both hemispheres², what is heard with a single ear is sent to both hemispheres, what is felt with either hand is likely to be seen as well, and so forth. In other words, there is a great deal of *perceptual redundancy* in day-to-day life: the coil that feels hot glows red; the cashier who puts the five-dollar bill and two dimes in your hand says, “Five twenty is your change,” and the cold rain is felt and seen and heard and even smelled. This perceptual redundancy ensures that, in both daily life and under most experimental situations, the two halves of a brain have access to similar information about the environment (and the body), whether they receive this information *from each other* or not.

² Each hemisphere receives visual input from the contralateral eye via the optic chiasm and from the ipsilateral eye via the nerve. Although the hemispheres receive different visual information from the optic tract, visual information is lateralized not by eye but by visual field. The *left* visual field (LVF) projects to the *right* hemisphere (RH), and the *right* visual field (RVF) projects to the *left* hemisphere (LH). In other words, each hemisphere receives, via the optic tract, information concerning the contralateral visual field, the RH about the left side of space, the LH about the right side of space. In “normal” subjects—subjects who have not undergone partial or full callosotomy—each hemisphere also receives visual input from the ipsilateral visual field via the corpus callosum.

Myers' and Sperry's second hypothesis for why split-brain subjects seemed so unimpaired following sectioning of the callosum was that *human* split-brain subjects in particular were, albeit not necessarily consciously or deliberately, finding ways to compensate for the lack of cortical (callosal) communication between their two hemispheres, partly via various attention-directing and self-cuing mechanisms and behaviors.³ Some of this “cross-cuing” they couldn't help but engage in; for example, as Sperry pointed out, “the two retinal half-fields of the eyeball move as one, and eye movements are conjugate, so that when one hemisphere directs the gaze to a given target the other hemisphere is automatically locked in at all times on the same target” (1974: 7-8).

A second wave of split-brain studies—on human subjects—began in the sixties. On the basis of Myers' and Sperry's results, the new studies were designed first of all, to direct sensory information to only a single hemisphere at a time and, second of all, to prevent or limit *behavior* that would serve as a means of “communication” between the hemispheres. The results of these studies—the split-brain experiments, which are still ongoing—are famous, and famously fascinating. Just like those of split-brain cats, the two hemispheres of human split-brain subjects evidenced high degrees of functional independence.

There are several ways to limit the path of incoming sensory information to one hemisphere only. The simplest is in the case of tactile stimulation of the hand,

³ Throughout this dissertation, I use the term *subject* to refer to a single human animal. I mean this term to be *neutral* on all questions concerning personal identity; thus a subject may or may not be equivalent to a person, and *subjects* may or may not stand in a 1:1 relationship with *subjects of experience*. In other words, a split-brain *subject* may or may not be (or possess) multiple subjects of experience, and the reader should not mistake talk of subjects for talk of subjects of experience. This choice of terminology may be slightly confusing, but I prefer referring to split-brain *subjects* rather than to split-brain *patients* because many split-brain subjects are merely voluntary participants in scientific research, and are not actually seeking medical treatment or advice.

which is transmitted largely contralaterally⁴; a split-brain subject's left hemisphere receives tactile information primarily from her right hand, and vice versa. Smells are only slightly more difficult; smell is transmitted only ipsilaterally, so if the person's right nostril is sealed, only his left hemisphere knows what's cooking. More complicated set-ups are required to keep the left hemisphere ignorant of what the right hemisphere sees or hears, but this is still possible, and many of the most compelling split-brain experiments involved visual perception in particular.⁵

As mentioned previously the experiments also had to be designed and monitored with the goal of preventing the hemispheres from communicating via behavior: subjects had to be prevented from *peeking* at the object they'd been asked to explore manually, from talking out loud to themselves as they worked, and even from gesturing with their hands. Stimuli presented to either hemisphere that yielded an *emotional* response (often creating a change in facial expression) often permitted more than one form of communication between the hemispheres, as we shall see later.

During these carefully designed and controlled experiments, split-brain subjects were found to suffer some cognitive or perceptual deficits: put an object into

⁴ In general, the more proximal the body part, the greater degree of bilateral projection; the more distal the body part—particularly the hands—the more lateralized the projection. Even for the trunk, however, there may be a *greater* degree of contralateral than ipsilateral sensory projection. Most split-brain experiments that stimulate the body stimulate the hands. Information about pain, temperature, and “passive touch” is carried via the spinothalamic pathways, which have both contralateral and ipsilateral projections. “Active touch” and proprioceptive information travels along the dorsal-column-lemniscal pathways, which have no ipsilateral projections. But because it's *active touch* that's required for object-recognition in the tactile domain—i.e. for recognizing an object on the basis of having *felt*, and, generally, *manually explored* the object—and because most of the studies of split-brain perception in the tactile domain tested object recognition, tactile perception can generally be treated as contralateral only.

Sensory information from any point on the face and neck is available to both hemispheres.

⁵ In a dichotic listening experiment, each hemisphere of a split-brain subject will perceive acoustic signals from the contralateral ear, only. Visual information is lateralized by presenting visual stimuli not at fixation point but left or right of that point. (Or two distinct stimuli may be presented simultaneously, one in each visual field.) The RH perceives the LVF, the LH the RVF.

a split-brain subject's left hand, and, if this hand remains hidden from the subject's view, he will tell you he has no idea what he's holding. But the phenomenon didn't look like anomia exactly; after all, if the same object was placed in the subject's *right* hand, he could tell you immediately what it was. It was only at right hemisphere perceptual tasks, indeed, that the subjects appeared to suffer from some sort of linguistic impairment.

If this were the end of the data, we would most likely conclude that the subject simply didn't know what he was holding in his left hand, just as he said. We would conclude that for some reason, the tactile information wasn't registering at the cortical level or entering conscious awareness, or even that the subject had some more profound sensory deficit on the left side of his body; that the subject's left hemisphere escaped the surgery cognitively intact, while his right hemisphere for some reason did not. This would already be interesting, even though, in a way, we might have expected something like this. We might expect that people who undergo such major surgery should suffer some major cognitive impairment. We might not even be surprised to find that the two hemispheres are differently affected by the same surgical procedure. By the 1960's it had already been known for a hundred years that in humans the two hemispheres are not functionally identical. It was believed even at the time for example (and is still popularly believed) that the right hemisphere excels at certain kinds of spatial reasoning, and visual tasks requiring holistic processing, like face-recognition. Best known and best studied is the linguistic asymmetry between the two hemispheres. For most people, the left hemisphere contains Broca's

area (responsible for language production) and does most of the receptive language-processing as well.⁶

But what was fascinating about the split-brain subjects was not that they showed severe right hemisphere impairment, but that they seemed, in fact, less impaired than they *claimed* they were. Put a pipe in a split-brain subject's left hand, and he will say, as long as he doesn't have visual access to that hand, that he doesn't feel anything and can't even guess what he might be holding; then place that same hand in a box containing a pencil, a pipe, a paperweight, and an apple, ask him to select the object he was just holding, and he will select the pipe every time, all the while complaining bitterly that the task is pointless since he has no idea what he was just holding, as he has already told you. The subject's right hemisphere, in other words, seemed able to *feel* the pipe, to *identify* it, to *remember* it, and to *re-identify* it. But the subject—speaking through his left hemisphere, the hemisphere with spoken language, but which hadn't received the necessary perceptual information from the left hand—would claim ignorance.

Via their left hemispheres, however, split-brain subjects actually professed ignorance less often than they should have. In particular, when a subject engages in

⁶ For the sake of brevity and simplicity, I generally speak as if the left hemisphere were the “dominant,” linguistic hemisphere in all adults. This is in fact not the case: a small but significant percentage of the population is right hemisphere dominant. (It is estimated that 10% of the population is left-handed—i.e., right hemisphere-dominant for motor tasks—and that up to 20% of the left-handed population is right hemisphere-dominant for language as well. This means that perhaps 2% of the general population is right hemisphere-dominant for language.)

There is a very small group of subjects who have undergone either callosotomy or hemispherectomy (in which callosotomy is performed and then an entire hemisphere removed or functionally incapacitated). The subset of this group which is right hemisphere-dominant for language must of course be many times smaller. That said, it seems possible that the rates of right hemisphere-dominance are somewhat higher among those who developed epilepsy in early childhood than among the general population, because early left-hemisphere damage (for many subjects, seizures have a pattern of beginning in the same hemisphere, whether right or left, each time, and thus over time may particularly damage the tissue in that hemisphere), might lead the right hemisphere to take over linguistic structures and functions.

some right hemisphere-initiated or right hemisphere-controlled behavior, instead of admitting (via his left hemisphere) that he has no idea why he's just done X, a subject often immediately confabulates (via the same hemisphere) an explanation instead. For example: one subject, N.G., giggled when her left visual field (LVF), i.e. her right hemisphere (RH), was presented with a picture of a naked woman (Gazzaniga, LeDoux, 1978: 154). When asked why she was laughing, she seized upon the experimental equipment. "That's a funny machine," she said.

Callosotomy results in what is sometimes called the "callosal disconnection syndrome," in which the two hemispheres of the brain can be shown to function, at least in experimental situations, independently of each other, with respect to a good deal of their perceptual, cognitive, and motor processing. I will simply call this the *split-brain phenomenon*. Split-brain studies on humans raised a host of philosophical problems and questions that the studies on cats alone would not have raised, at least not as vexingly or as pressingly: *does sectioning a person's callosum create two streams of consciousness? Two subjects of experience? Two persons? Two minds?*

2 Minds, Brains, and Unity

The split-brain studies done by neuropsychologists like Sperry sparked a flurry of philosophical papers, written both by philosophers and by various scientists writing in a philosophic mode, on the import of these studies for our understanding of personal identity, consciousness, free will, and the mind-brain relationship. In fact the history of responses to the split-brain studies is interesting in and of itself. The split-brain cases were ahead of their time in terms of the degree of interdisciplinary attention

they almost immediately attracted. And when the philosopher Puccetti argued in 1981 that both split-brain and “normal” subjects have two minds and are two persons apiece, his article, published in *The Behavioral and Brain Sciences*, was followed by peer commentaries written by professionals in philosophy, in hospitals and medical centers, in academic psychology, in cognitive psychobiology laboratories, and in medical colleges.

Despite having attracted the attention of a broad range of thinkers early on, the results of the split-brain studies continues to be grappled with by psychologists and by philosophers. Despite the fact that callosotomies are performed less frequently as the years pass and with the advent and improvement of less invasive treatments for severe epilepsy, those old (and those few new) split-brain subjects who are still around are heavily sought after by neuropsychologists, and new split-brain studies are published each year.

Attempts to characterize these subjects’ mental lives occupy almost every possible position. Eccles (1965) has argued that the right hemisphere isn’t associated with any mind. Gazzaniga has argued that the corpus callosum is the mechanism for mental and conscious unity in “normal” subjects, and that split-brain subjects therefore have two minds, but views the left-hemisphere mind with the distinctively human (Gazzaniga, 1983a, 2000). Marks (1981) has argued that the split-brain cases show that a single mind need not have unified consciousness; Sperry has argued that split-brain subjects have two intelligent minds (1990); Bogen has argued that the split-brain cases show that a single person can have two minds (1990); Puccetti

(1981, 1989) has argued split-brain subjects have two minds and are two persons—just like everyone else with two cerebral hemispheres.

The split-brain cases have retained their ability to surprise and amaze for forty years partly because the philosophical and scientific questions they raised then are still open now. This work is about *individuating mental tokens* from a theoretical or scientific perspective, and about the insights that the split-brain studies yield into such individuation. It focuses on two questions about mental tokens in split-brain subjects in particular: how many minds they have, and how many streams of consciousness they have.

2.1 Brain Bisection and the Unity of Consciousness

It may be instructive to examine a work in the existing literature on individuating mental entities in split-brain subjects—Nagel’s (1979) “Brain bisection and the unity of consciousness”—for two reasons. First, Nagel provides a good introduction to many of the philosophical puzzles about the split-brain phenomenon, and neatly describes some of the difficulties associated with ascribing either mental duality (two minds apiece) or mental singularity (one mind apiece) to split-brain subjects. Nagel’s article also ends up being as much about individuation *questions* as it does about their answers: his own remarks on individuation issues raise further questions about the methodology of individuating mental tokens, and about what it is that we’re doing when we ask how many minds split-brain subjects have.

Split-brain subjects can easily be cast as walking, talking embodiments of the so-called *mind-body problem*, which is no doubt one source of their enduring fascination. To Nagel, the split-brain phenomenon reveals an acute tension, one that

isn't only, or perhaps even primarily, a tension between mental and physical properties, or mental and physical descriptions of phenomena—the tension referred to by the *explanatory gap* problem. Rather, for Nagel, the studies reveal, first, a tension between two different modes of reasoning and acquiring evidence about the structure of split-brain subjects' cognition. For the subjects don't *feel* as if they have two minds, or two streams of consciousness. And in their daily lives they don't *behave* as if they do, either. But during carefully controlled experimental situations, they *do* act as if they have two minds, and have two streams of consciousness, and so forth. Moreover, the scientific explanation for their more “disunified” behavior seems to require that if they have two minds and two streams of consciousness at their most “disunified” moments, they have them always. So the split-brain subjects present us with two conflicting sets of evidence: that acquired from ordinary experience and interaction with the subjects, and that revealed via careful experiment.

But for Nagel this points to a much deeper tension revealed by the split-brain phenomenon. The phenomenon reveals a tension between two different ways of understanding the human subject and the human mind, all human minds: scientific modes of understanding, and non-scientific, intuitive or folk psychological, subjective and inter-personal, modes of understanding. Nagel's fundamental concern is whether there is any scientific plausibility to our concepts of mental, personal, and conscious unity, and whether a scientific understanding of ourselves as physical systems is compatible with our pre-theoretic ways—personal, practical, moral, legal—of understanding ourselves and those around us, as mental beings, as persons, as the subjects of a unified consciousness and a coherent mental life. And Nagel ultimately

suspects not only that the two modes of understanding are incompatible—that they will yield incompatible claims about split-brain subjects, and everyone else—but that we won't be able to let go of either one of them. He despairs, that is, that our understanding of ourselves as persons, as experiencing beings, and so forth, *outside* the context of a scientific psychology, will prove incompatible with a growing scientific understanding of ourselves as physical systems.

This work is in part an attempt to meet at least some of the challenge posed by Nagel, the challenge of showing that there is a determinate answer to the question of how many minds and streams of consciousness split-brain subjects have. But that portion of the work is necessarily tentative, and it could turn out that the concept of a single unified stream of consciousness, for instance, really doesn't apply neatly to split-brain subjects. I am less interested (at least in this dissertation) in *integrating* the theoretical and the pre-theoretic modes of understanding, though I do argue at some points that the approach to individuating mental tokens I defend, and any answers it yields, will not be so incompatible with our ordinary ways of understanding split-brain subjects. That is, I not only argue for what I think is the proper scientific understanding of split-brain subjects' mental lives; I also suggest that the models of mind and of consciousness that I apply to split-brain subjects would not do the violence to our ordinary (practical, personal, social, moral) ways of understanding them that Nagel feared. Still, the most fundamental aim of the work is simply to defend a theoretical approach to individuating mental tokens in split-brain subjects (and, of course, in general), in part simply by attempting to take such an approach.

2.2 Five Hypotheses

Nagel formed five competing hypotheses as to how many minds and streams of consciousness a split-brain subject might have. The first was that such a subject has a single mind and a single stream of consciousness, both associated with only her left hemisphere, and the second was that, while the right hemisphere might be associated with some conscious experiences, those experiences don't belong to a mind (and perhaps aren't part of a genuine stream of consciousness, either). He rejected these first two hypotheses quickly. The right hemisphere-controlled behavior that subjects engage in during the split-brain experiments reflects the presence of a right hemisphere mental system that perceives and remembers, that learns, that reasons, that experiences human emotions: love, humor, embarrassment. Right hemisphere mental processes, in other words, reveal the functional organization of a mind.⁷

Whether the right hemisphere of a split-brain subject is associated with *conscious* mental phenomena is somewhat more controversial. It is unlikely that it would have the very *same* kinds of conscious experience as the left hemisphere; it is unlikely that a split-brain subject's right hemisphere, for instance, would generate a stream of inner speech, given its non-dominance for the production of propositional

⁷ It should be noted that there are significant individual differences between split-brain subjects, with respect to all the "normal" mental properties (of personality, intelligence, etc.), but also with respect to the abilities of their right hemispheres in particular. At one extreme, some split-brain subjects don't appear to engage in any interesting right hemisphere-controlled behavior; at the other extreme some subjects even learn to *speak* out of the right hemisphere. (Subject P.S. began to provide verbal descriptions of LVF/RH phenomena about two years following callosotomy, for instance (Gazzaniga et al. 1979). See Gazzaniga, 1983a for one perspective on RH linguistic capacity in split-brain subjects.) Split-brain experiments are of course run using the subjects from whose right hemispheres interesting behavior *can* be elicited. These are the subjects I talk about in this dissertation. Of course there are individual differences among those subjects as well, and so some of my specific claims about the split-brain phenomenon will apply more to some of these subjects than to others (or perhaps only to some subjects and not to others), and perhaps nothing I say about the phenomenon will apply to *all* split-brain subjects. It is possible to make the philosophically important points, though, while abstracting away from these differences.

language. And right hemisphere consciousness may be depressed in many subjects (as Sperry (1990) speculates), due to left hemisphere dominance and draining of attentional resources. But right hemisphere-controlled behaviors are sometimes intelligent and *deliberate*-looking to a degree that strongly suggests consciousness. On some theories of consciousness, the right hemisphere's conscious status will be less certain than on others'. But Nagel, at least, concludes, and I accept in this work, that the right hemisphere of a split-brain subject is associated with *some* stream of consciousness, and some mind, or other. The question is: which one?

Nagel's next three hypotheses all closely concern *individuating* minds and streams of consciousness. Hypothesis three is that each hemisphere of a split-brain subject is associated with a unique mind and a unique stream of consciousness. I will call these the claims of mental and conscious duality. Hypothesis four is that the two hemispheres jointly constitute a single mind, and perhaps even generate a single inter-hemispheric stream of consciousness; at some moments, however, the hemispheres may each be associated with a distinct stream of consciousness, and, more generally, the contents of the two hemispheres are occasionally "dissociated." These hemispheres still always jointly constitute a single mind, however. (Because this hypothesis attributes to split-brain subjects a single mind at all times, and because it can attribute to them a single stream of consciousness at most times, I refer to it as making mental and conscious *singularity* claims.) And hypothesis five is that split-brain subjects *normally* have a single mind, and a single stream of consciousness, but

that sometimes—especially under experimental conditions—they have two minds and two streams of consciousness.⁸

According to the last hypothesis, callosotomy alone is insufficient to divide a single mind into two. Rather it is only the combination of callosotomy and a calculated directing of sensory information to one hemisphere or another and steps taken to prevent split-brain subjects from engaging in cross-cuing that divides one mind into two. The similarity between split-brain subjects' behavior, outside of experimental situations, and our behavior, owes to the fact that they like us have single minds during these times. The dissimilarity between split-brain subjects' behavior, inside of experimental situations, and our behavior, owes to the fact that they unlike us have two minds during these times.

This hypothesis most explicitly raises the issue of what sorts of processes and capacities are sufficient to yoke potentially distinct mental systems into a single mind. The primary strength of this hypothesis is its midway position between attributing either one mind always or two minds always to split-brain subjects—either of which makes a commitment to one popular intuition (one stream of consciousness per mind and one mind per body, respectively) at the cost of abandoning another (one mind per body and one stream of consciousness per mind, respectively). And the last hypothesis seems to fit particularly neatly with *all* the data on the subjects—at least

⁸ Note that in all these hypotheses it is the mental status of the right hemisphere that is considered in need of determination. The association of the left hemisphere with a mind, and the identification of the subject with at least one mind, is not questioned. It is clear why: pre-surgery, the subjects had minds, and post-surgery, they talk, act, and claim to feel much the same. Through their left hemispheres, the subjects speak of their families, their recovery from the operation, their moods. There must be a mind in that brain somewhere, in other words, and the left hemisphere can avow its own conscious mentality. The split-brain puzzle arises because one can't easily see how the right hemisphere could be associated with no mind at all (because right hemisphere-controlled behavior seems mindful) or with the left hemisphere's mind (due to the disconnection syndrome) or with its own mind (because then the subjects would have two minds each, which seems extraordinary).

on the surface of things. For it can attribute any odd behavior to duality, and any “normal” seeming behavior to singularity.

The problem with the last hypothesis, as Nagel notes, is that it is “entirely *ad hoc*” (Nagel 1979: 161); it achieves a “neat fit” with the behavioral data, or a kind of superficial explanatory plausibility, at a steep cost. For now how should we explain the significant change in mental structure a split-brain subject undergoes between entering and exiting an experimental situation? The split-brain experimental paradigm is certainly artificial. . . But still mainly just involves controlling the kinds of perceptual input that the subjects (or each of their hemispheres) have access to. How could merely *sealing a nostril* create a second mind? As Nagel says, “So unusual an event as a mind’s popping in and out of existence would have to be explained by something more than its explanatory convenience” (ibid).

Of course, this objection to this hypothesis isn’t decisive, either. We should have to know more about what kinds of things minds are to know how easily they can “pop” in and out of existence. And from one perspective, functionalist commitments may make the hypothesis look not entirely unreasonable. Perhaps whether two mental systems (understood broadly to include everything from whole minds to visual systems and so forth) jointly constitute a mind or instead each constitute a distinct mind is a matter of the kinds of communication (or lack thereof) between them. And perhaps the split-brain experimental paradigm alters that communication in a manner sufficient to break down that bi-hemispherically constituted mind into two minds. More searching questions about what minds are, and about the kinds of interactions

between mental systems that suffice to make them one mind, will be considered in Chapters Two and Three.

Nagel is soon left with hypothesis three (according to which split-brain subjects have two minds and two streams of consciousness apiece) and hypothesis four (according to which they have a single mind apiece, with sometimes dissociated contents conscious contents). A version of this latter hypothesis has been defended both by Marks (1980) and by Tye (2003): according to these philosophers, a split-brain subject always has a single mind, and usually has a single stream of consciousness, but occasionally—especially during the split-brain experiment, when the conscious contents of the two hemispheres (most) diverge—such a subject has two streams of consciousness.

Nagel believes that there are very strong conceptual connections between unity of mind, unity of consciousness—and unity of the person. In fact at some points in his 1979 he refers seemingly interchangeably to *being* a single person, *having* a single stream of consciousness, and *having* a single mind. It is in part because the conceptual connections between these things are so tight, for Nagel, that he cannot accept hypothesis four:

Lack of interaction at the level of a preconscious control system would be comprehensible. But lack of interaction in the domain of visual experience and conscious intention threatens assumptions about the unity of consciousness which are basic to our understanding of another individual as a person. These assumptions are associated with our conception of ourselves, which to a considerable extent constrains our understanding of others.
(Nagel, 1979: 160)

So for Nagel it doesn't appear *possible* to discuss how many minds or streams of consciousness split-brain subjects have without fairly definitive implications for personal identity. And we have deep commitments in the realm of personal identity, such that how we think about persons, including ourselves, and presumably in all sorts of non-scientific contexts, effectively constrains what we can accept about their minds and consciousness. For Nagel, it seems, attributing two streams of consciousness to a single person (simultaneously) is impossible primarily because we can't "conceive what it is like to *be* one of these people" (ibid; original emphasis). Thus a person is a subject of experience, for Nagel, also, and a single person cannot be associated with two subjects of experience or two streams of consciousness.

As we shall see in Chapter Three, Marks (1981) and Tye (2003), in contrast, again both accept versions of this hypothesis. And I think they are right to suspect more flexibility in our *concepts* of mental and conscious singularity than Nagel allows.⁹ Certainly scientific concepts and ideas appear amenable to a great deal of revision in light of new evidence and revised understanding, before they must be scrapped entirely. Even our folk psychological concept of a single mind may not absolutely *require* that a single mind has a single stream of consciousness (at a time).

For the folk appear comfortable ascribing various sorts of mental and conscious

⁹ Although most philosophers writing about individuating minds and streams of consciousness in split-brain subjects have contrasted the mental and conscious duality or disunity positions with the mental and conscious *unity* positions, I contrast mental and conscious duality with mental and conscious singularity. I prefer this terminology because talk of unity and disunity connotes harmony and discord, respectively, while talk of singularity and duality has fewer of these connotations. Split-brain subjects don't seem subject to a great degree of inter-hemispheric *conflict*; I argue that they may nonetheless have two minds and two streams of consciousness—perhaps just two cooperative minds, or two streams of consciousness with consistent contents. So I will generally contrast *singularity* with *duality* (or, sometimes, with *multiplicity*). There may be times however at which I refer to unity and disunity, either because I am saying something about mental or conscious harmony or discord, or because I am describing someone else's views, and those views aren't easily or accurately translated into talk of singularity and duality.

duality and disunity to single individuals: consider stories like that of Dr. Jekyll and Mr. Hyde, fear of demonic possession. . . . (Or even a recent mediocre film in the comedy genre, in which a mortal who becomes God for a week is shown with the prayers of however many billions of the world's inhabitants—or at least the world's believers—running through his head simultaneously). So maybe Nagel's conceptual framework is unusually fixed; maybe neither the scientists nor the folk would have as so much trouble admitting of a single mind with two streams of consciousness.

The hypothesis faces other challenges, however. Most fundamentally, a proponent would have to offer an account of the sort of mental architecture that would sustain one stream of consciousness most of the time and then two streams of consciousness occasionally—and one would have to offer an account of consciousness and conscious unity that permitted this also. But this objection obviously requires explanation and argumentation, and I won't attempt to provide those things in this introductory chapter. But moreover, if we *weren't* sure whether the right hemisphere of a split-brain subject is associated with any conscious phenomena at all, this wouldn't dissolve any worries that we had about how many minds the subject has. The subjects appear, at least at some moments, to have not just two sets of conscious experiences, but two sets of memories and beliefs, two sets of goals and desires and intentions—even two sets of dispositions, likes and dislikes. They really do appear to have two minds, and not just two streams of consciousness. Or, as Nagel says:

The experimental situation reveals a variety of dissociation or conflict that is unusual not only because of the simplicity of its

anatomical basis, but because such a wide *range* of functions is split into two noncommunicating branches. It is not as though two conflicting volitional centers shared a common perceptual and reasoning apparatus. The split is much deeper than that.
(Nagel, 1979: 159; original emphasis)

This objection isn't definitive, of course. But it at least threatens to undercut the main motivation behind hypothesis four, whose main advantage seems to be (this is certainly why Nagel proposes it) that it purports to obviate the need to attribute two *minds* to a split-brain subject. Except that it wouldn't.

What, then, of hypothesis three—that split-brain subjects have two minds and two streams of consciousness at all times? The mental and conscious duality claims for split-brain subjects face several objections, some more and some less compelling; since I defend these claims, objections to them will be dealt with in some detail later on in the work. The most obvious (and most often cited) objection to these positions concerns the generally “normal” and “integrated” character of the subjects' behavior, at least outside of experimental situations. In fact, as Nagel notes, they behave in a fairly integrated fashion even when they are at their most “disunified,” i.e. during the split-brain experiment. The challenge from split-brain subjects' generally integrated behavior is significant, and I will deal with it, again, at several points throughout this work, though especially in Chapters Three and Four. But for now, simply note that the challenge may not be quite as great as some believe. Nagel, for instance, is ultimately unwilling to accept his third hypothesis in part because of the “highly integrated character of the subjects' relations to the world in ordinary circumstances” (Nagel, 1979; 159). But surely this is explained in significant part, at least, by the fact that the hemispheres are both hooked up to the same world—in the same place at the

same time at every time. By achieving a fit with the world, they would achieve some degree of fit with each other—even without interacting with each other in any direct world. He also says that it seems “strange to suggest that we are not in a position to ascribe all those experiences to the same person, just because of some peculiarities about how the integration is achieved. The people who know these subjects find it natural to relate to them as single individuals” (ibid). But again, this first of all just assumes that a *person* cannot possess more than one mind. Second, how *else* would “the people who know these subjects find it natural to relate to them,” than as single individuals? Of course is natural to relate to the walking, talking, familiar body of a loved one as if he were a single person with a single mind—particularly when he seems, to you, unchanged by the operation (assuming you aren’t performing experiments on him yourself), before which you also related to him as a single person with a single mind. Though Nagel is right that it would indeed *feel* strange, and probably uncomfortable, to ascribe to a single person two minds—or to claim that a single body were home to two persons (if we want to keep a one-to-one relationship between persons and minds)—this strangeness and even our discomfort do not mean that such an ascription wouldn’t in fact be the most correct. For whatever it matters, making this ascription could in time come to feel quite natural, if it did in fact fit well with the data, and if this fit is important to our concepts of mind and person.

Nagel also makes a sort of regress objection to the duality position: if we allow that at least some human beings have two minds, why not allow that they have four or six or seven? Perhaps there is one mind associated with the visual cortex, one associated with the auditory cortex. . . . He suggests in other words that if we allow

that a human being can have two minds, there will be no non-arbitrary way of limiting him to just two, either because each hemisphere may be itself be composed of many somewhat functionally distinct systems, or, perhaps, because he thinks distinct centers of consciousness might supervene on individual sensory processing areas. I don't think most people who advocate a mental duality in split-brain subjects are concerned about this though. For many of those who defend the mental duality model of split-brain subjects may take minds to be things that have, for example, memories and motor systems and perceptual systems; Bogen for instance writes that, "If the 'mind' includes (at least) the abilities to perceive, discriminate, compare with previously stored information, consider alternative actions, choose among them, and then to act, it is doubtful that any smaller subdivision of a mammalian brain than one hemisphere can support a mind" (1990: 216). A simple distinction between minds and modules can probably be drawn firmly just where Nagel fears that it can't.

Nagel's final objection to the hypotheses of mental and conscious duality in split-brain subjects takes the form of a *reductio*. If the mental duality claim offers the right way to think about split-brain subjects, he asks, then might it not offer the right way to think about all of us? If the fact that split-brain subjects seem like one-minded people every day is best explained not by saying they do in fact have single minds but by referring to perceptual redundancy, shared embodiment, indirect or virtual communication between the hemispheres, then perhaps the fact that we all seem like one-minded people is best explained not by saying that we have single minds but by referring to perceptual redundancy, shared embodiment, direct or neural and indirect

or virtual communication between the hemispheres. In our case, the communication is perfect; in split-brain subjects, the communication is simply less perfect.

But it is clear that this line of argument will get us nowhere. For if the idea of a single mind applies to anyone it applies to ordinary individuals with intact brains, and if it does not apply to them it ought to be scrapped, in which case there is no point in asking whether those with split brains have one mind or two.
(Nagel 1979: 162)

But, first of all, even if accepting a mental duality claim for split-brain subjects somehow required accepting such a claim regarding “normal” subjects’ mental lives, also, this need not, contra Nagel, constitute a *reductio* of the duality model. Why couldn’t “normal” subjects turn out to have two minds? If the empirical evidence and best theoretical models suggested that almost all of us *do* have two minds—hemispherectomized subjects constituting an exception—the concept of a single mind would hardly turn out to be useless or meaningless. The concept would still usefully pick out what “normal” subjects have two of—and what hemispherectomy subjects have one of. It is probably true that our concept of a single mind developed with average human beings in mind as exemplars. But this does not yet show that the meaning of the concept is exhausted by these exemplars. And the concept could survive the loss of them *as* exemplars.

Moreover, the adequacy of a mental duality model for split-brain subjects would not entail the adequacy of such a model for “normal” subjects. Again, it *could* turn out that both “normal” and split-brain subjects have two minds apiece. It could also turn out that the former have one mind apiece and that only the latter have two minds apiece. Which of these is in fact the case depends both upon what it is exactly

that the corpus callosum *does* when it is present and functioning normally, and on what kinds of integration typify the inner workings of a single mind.

Still, Nagel, at least, was unable or unwilling to accept a mental duality claim for split-brain subjects, even if he doesn't reject it in the disdainful language he reserves for some of the other hypotheses. He in fact ended his paper not by answering the "how many minds?" question but by finally considering the question itself. What kind of a question is it; from what standpoint would it be answerable? Nagel confesses that he sees no compelling reason to conclude that split-brain subjects have either one mind or two—or even that we ourselves have either one mind or two. He suggests that this is because our very idea of mind:

may resist the sort of coordination with an understanding of humans as physical systems, that would be necessary to yield anything describable as an understanding of the physical basis of mind.

(Nagel 1979: 147)

. . . and that our concept of mind will turn out to be "recalcitrant to integration with . . . data [from neuroscience, etc.]." Thus, he says, either our "idea of a single person will come to seem quaint some day" or else—and this is perhaps worse—"we shall be unable to abandon the idea no matter what" scientists discover about the human behavior, psychology, and the brain. Nagel himself seems to believe that this second prediction is more likely to be born out by time (1979: 164).

3 Conclusion

As some of the quotes above illustrate, questions about how many streams of consciousness split-brain subjects have and about how many minds they have seem to be inextricably linked, for Nagel, to questions about how many persons they are. This

dissertation isn't about personal identity, however; I focus on the individuation of minds and streams of consciousness and mental states instead. But my concerns in this work do overlap with Nagel's, to a significant degree, for this dissertation attempts to sort out a few matters essential to developing an understanding of human minds as physical systems. Unlike Nagel's "Brain bisection and the unity of consciousness," however, this dissertation focuses on the individuation of mental tokens from a theoretical perspective—from the standpoint, that is, of scientific psychology.

Chapter Two lays out some of the background necessary to individuating minds in split-brain subjects. It includes discussions on identity and individuation, a modest characterization of what minds are, and an attempt at sketching a distinction between what I call *inter-mind* and *intra-mind* interactions. In Chapter Three I defend the mental duality model for split-brain subjects, arguing that neuroanatomical evidence supports such a model, and criticizing some competing views on the role of neural—and behavioral—facts in individuating mental entities. Chapter Four then lays out some of the background necessary to individuating streams of consciousness, and describes some of the resources that the conscious duality model has to explain the generally integrated-seeming nature of split-brain subjects' behavior. Chapter Five defends the conscious duality model for split-brain subjects from some alternative models. Finally in Chapter Six I offer a defense of (moderate) *mind-brain supervenience*. This chapter offers the final word on *why* neural facts constitute a special kind of evidence for the individuation of mental entities; they occupy a place of privilege which some philosophers have argued is ill conceived. But emphasizing

neural evidence (while in no way neglecting behavioral evidence) in individuating mental entities does not amount to some kind of neglect of or lack of appreciation for the mental (functional). In fact, quite the reverse is true. *Realism* about the mental requires making various forms of commitments to the neural entities and events that sustain mental phenomena.

Chapter 2: On Individuating Minds and Other Mental Entities

1 Introduction

In this chapter, I will begin to talk about individuating mental tokens, and in particular, individuating minds. Some of the many issues that arise in the course of this venture are: What are minds? Do minds (as opposed to mental states) really exist? Are neural facts relevant to individuating minds—or should we look at psychological facts only? Is there a principled distinction between the sorts of interactions that can occur between minds—and the sorts of interactions that can occur only within a mind? And what is the relationship between the identity questions raised here and other identity puzzles in philosophy?

I do not yet, in this chapter, draw any conclusions about how many minds a split-brain subject has. This will wait for later chapters. This chapter provides the background necessary to doing that later work. Its main goals are to say something about what a mind is, what we're doing when we try to individuate minds, and how we should go about doing that.

In the next section I talk about the kinds of identity or individuation problems I will be concerned with in this work, and roughly locate them with respect to some other identity and individuation problems in the existing philosophical literature. I also talk a bit about the *entities* whose individuation concerns me in this work, and about the kinds of things I think that they are.

In that same section I also briefly discuss the nature of the “how many minds?” question specifically. I identify a possible deflationary response to the

question according to which it is merely a verbal one. According to this position, it is not only indeterminate how many minds a split-brain subject has, but also there aren't actually *minds*, in any interesting sense: there are mental states, and there are real and interesting questions about mental states, and so forth, but none about minds qua minds. I argue that at the very least the "how many minds?" question, which is a question about how to individuate minds (in general and in the split-brain subject), is not merely a verbal issue, but rather is interesting from the perspective of a scientific psychology, and no doubt from other perspectives as well.

Section Three is about the criteria for a system's constituting (one or more) minds. I sketch a very modest set of criteria for mindhood, setting aside a range of controversial questions about the details of the mental architecture that characterizes a mind. In this section I also introduce a distinction between purely functional and partly neural accounts of mindhood and criteria for individuating minds. For knowing the criteria that some systems must meet in order to qualify as at least one mind, does not yet tell you how many minds a system is or some systems are. One also needs to know the *criteria of individuation* for systems of that kind.

In the split-brain case, one wants to know whether the two hemispheres are associated with one and the same or with two distinct mental systems or minds. The answer to that question depends upon facts about the ways in which and the extent to which the hemispheres interact with each other—for they do interact with each other, in all sorts of ways; the question is whether those interactions are of an "inter-mind" or an "intra-mind" sort. A purely functional approach to individuating minds would want to characterize those interactions purely functionally, while a partly neural

approach would also look at facts about physical interaction. Either sort of approach, however, will need to make some kind of distinction between intra-mind and inter-mind interactions, since causal interactions between mental states occur both within and between minds. In Section Four, I propose one way of beginning to draw this distinction.

2 *On Mental Entities*

Most work in psychology probably asks causal rather than constitutive questions about, and seeks to develop causal rather than constitutive accounts of, mental phenomena. But constitutive rather than causal analyses of psychological phenomena are common in philosophy of mind, and they are the primary focus of this work, though obviously I rely on causal claims and analyses as well.

Within the philosophy of mind, meanwhile, most work providing constitutive analyses of psychological phenomena is probably work on characterizing mental types. But this dissertation is most directly concerned, instead, with *individuating mental tokens*: mental states, streams of consciousness, and minds. It does offer some characterizations of mental states, streams of consciousness, and minds, too, of course, since determining a thing's *token* identity depends upon the *type* of thing that it is. So in this chapter, for instance, I offer a (very minimal) characterization of what minds are; in Chapter Four, meanwhile, I will turn to the nature of a stream of consciousness.

Because there is a lot of ground to cover, and because the major questions about split-brain subjects that I focus on and the approach that I take to those

questions are somewhat different from those many philosophers have taken, I think it will be helpful to make some more introductory remarks and distinctions before really beginning. In the remainder of this section, then, I lay out some broad positions and assumptions that form the necessary background for the project of this dissertation.

Section 2.1 explains the major theoretical orientation or perspective I take in this work. I then say just a little about what I take mental tokens in general and minds in particular to be. These will just be *very* rough, quick characterizations, and I won't at that point provide any arguments for them or for the psychological accounts they presuppose. Much of this will instead be offered, gradually, over the course of the next several chapters (though of course like everyone else, I have to start with some fundamental assumptions, as well). In Section 2.2 I locate individuation questions about minds with respect to other sorts of identity and individuation problems. And in Section 2.3 I motivate the debate about how many minds split-brain subjects have, arguing that the "how many minds?" question for those subjects has both empirical and interesting conceptual aspects. These remaining introductory sections pave the way for the characterization of minds and the criteria for individuating minds that are developed in the rest of the chapter.

2.1 Types, Tokens, States, and Systems

To begin with, this dissertation is psychofunctionalist in its orientation.

Psychofunctionalism is a branch of functionalism that says that our best understanding of the constitutive conditions for various mental phenomena will come from a developed scientific psychology. In other words, the story providing the best analysis (by ramification) of mental states will be one of scientific psychology, and

the best answer to questions like, “What is it to be a thought?” and “What is a mind?” will come from a developed psychological theory.¹⁰ Of course, scientific psychology is, at this point, a work in progress, and not a complete and well-accepted theory or collection of theories. This admittedly gives the claims defended in this work a definite tentativeness, but many of the claims I defend deal with basics and broad pictures and thus are hopefully not entirely premature.

Although I believe that mental states are tokens of particular functionally defined types, these mental types needn’t be characterized entirely functionally, either. Some of the constitutive conditions for mental phenomena could turn out to be neural, or other physical conditions, as well. Functionalism involves abstracting away from some aspects of the physical, of course. But it is far too soon to say that a developed psychological theory will abstract away from the physical (particularly the neural) altogether. So a functionalist theory of the constitutive conditions for mental phenomena needn’t be given in *entirely* functional terms; it might refer to neural phenomena also, or to other physical (bodily or environmental) phenomena. (See on this issue Rey’s discussion of “anchored” functionalism, 1997, pp. 191-194.)

I refer to mental states, mental events, and mental representations, interchangeably, to refer to things like Alena’s desire for a sandwich, or the fear Devon experiences when the phlebotomist gets out the needle. The term “mental token” will be used in a broader way, sometimes to refer to a particular belief or emotion, for instance, but sometimes to refer to a (particular) stream of consciousness, or a (particular) visual system, or a (particular) mind. The term

¹⁰ Ramsification is a (logical) technique for defining theoretical terms, according to which the meaning of a theoretical term is entirely determined by the role that term plays within a theory. (The technique was developed by David Lewis (1972), inspired by an idea of Frank Ramsey.)

“mental token” will be used, that is, to refer to any *thing* that is mental or that possesses mental properties. (Context should make clear whether the mental token being referred to is a particular mental event or representation, or some “larger” mental token, such as a whole mind.)

There are different *types* of mental tokens, such as events of experiencing, of perceiving, of guessing, of remembering, of doubting, of believing, and so forth. Note that one or more of these psychological types just listed could turn out not to exist, however, since the best account of the constitutive conditions for mental entities will be given by a psychological theory or by theories that we don’t yet possess, and in which events like *guessing* could turn out to play no part. But this work is not really about mental types: it is really about what makes a particular mental token *that* particular mental token, as opposed to a different token of the *same* type, for instance. I do nonetheless require some basic assumptions about mental types generally, such as that mental states are intentional—they have content—and that to be in a mental state is to stand in a certain functional relationship to a representational content. (Something—a subject, or a system—may stand in different functional relationships to the same representational content, and something may also stand in the same functional relationship to different (i.e. to multiple) representational contents.) But while the dissertation assumes a functionalist/representationalist account of mental states generally, and therefore assumes that the identity of any mental state is always *in part* a matter of its functional role, nothing said here precludes the possibility that *some* mental states *additionally* possess qualia, or non-mental, non-representational properties, that also matter to their identity.

Finally, I take it that physicalism is true, and that every mental token has physical properties. This in and of itself will not be argued for. For most of this dissertation, however, I additionally assume that the mental tokens are neural tokens, but in Chapter Six this claim will be argued for explicitly. This work is also (as is probably clear by now) written from the standpoint of realism, rather than instrumentalism, about mental phenomena and psychological explanation. I argue in Chapter Three that the joint commitment to realism and to physicalism has implications for individuating mental tokens in split-brain subjects and in general.

Like a desire or an emotion, a mind is a kind of mental token—but it is a different kind of mental token. If mental states are states *of* a mental system, a mind *is* such a mental system. So while some might identify minds with streams of consciousness, for instance, and some with agents, I accept what I think is the most useful view of minds—and the one that does the most justice to current psychological theorizing on minds—which is that a mind is a *system*: a collection of parts organized so as to function as a whole, and as to yield operations that are not just the sum of any of the operations of its parts. (Of course, an agent can also be understood as a kind of system.) So the system must engage in operations or behaviors of characteristic sorts, and it must have an organization or architecture of a certain sort also: a mind it is composed of various subsystems interacting with each other in particular ways. Because these subsystems also generate and manipulate mental tokens, and have their own internal organization, they are also mental systems (or, as I will tend to call them to avoid confusion, mental *subsystems*). A *mind*, though, is a special kind of mental system. A mind is a mental system that does all the things that a mind must do,

whatever those things turn out to be, and again it does this by virtue of possessing a certain architecture, whatever that architecture turns out to be. I say more about these matters in Section 3.

Now that I have laid out, very broadly, the kind of perspective on the mental that is taken in this work, let me also make some basic remarks about identity and individuation, and roughly locate the identity and individuation questions that are the focus of this work with respect to other sorts of identity and individuation problems.

2.2 Identity and Individuation

As I have said, this dissertation's primary focus is the individuation of mental entities. There is, of course, a great philosophical literature on identity and individuation. Much of this literature concerns the identity of persons, a subject not specifically addressed in this work. That said, certain concepts that figure in personal identity debates, including those of *mind*, *stream of consciousness*, *subject of experience*, and *conscious unity*, figure very significantly in this dissertation as well.

A great deal of the philosophical literature on identity also concerns identity over *time*. This literature's fundamental question is: given an X at t1 and a Y at t2, what makes it the case that X (where X is the Self, or the Ship of Theseus, or the Vienna Circle) = Y? And indeed the split-brain phenomenon does obviously pose some interesting problems for accounts of diachronic identity, perhaps diachronic personal identity in particular. Those tempted to identify the split-brain subject with two persons, for instance, would have to say either that the subject was (constituted by or associated with) two persons even before the surgery, or else say that the subject was a single person prior to the surgery, but that that person has ceased to

exist. . . . Or else say that the subject was a single person prior to the surgery, and that that person continues to exist as *one* of the persons *currently* associated with the subject.

But it is synchronic, and not diachronic, identity that is the focus of this work. Granted, because a mind is a mind in virtue of its activities and functional organization, rather than because of some static property, I cannot separate the synchronic from the diachronic entirely; I cannot, and do not, confine myself to speaking of instantaneous mental properties in individuating minds. But it is nonetheless the case that much of the philosophical literature on identity concerns diachronic identity most directly, and that diachronic and synchronic identity questions involve a related and overlapping but still distinguishable set of concerns.

Diachronic identity puzzles typically arise because of the problem of *change*: what sorts of *changes* can a thing undergo and still remain *the same thing*? It is intuitively easy to see the difficulty of answering such questions. But one might think that *synchronic* identity or individuation problems would be easy to resolve—indeed, perhaps that they might not even arise. It is easy to know whether some stuff is one individual thing or several things, isn't it? And easy to tell whether, at any one moment, something is one thing not another (or part of another) thing of the same sort? (I don't mean to put the problem entirely epistemologically, since ultimately we are interested in what *makes* something an individual.)

I think that it is easy to think that individuation is easy because when we think of individual things, the first things we tend to think of are familiar physical objects: tables, chairs, sandwiches, alarm clocks, motorcycles. And it is relatively easy to

individuate such objects. At any given moment, even if one is not certain what has become of the Ship of Theseus, and whether the ship one is sailing on is *that* ship, it is at least clear that one is standing on *one* ship and not two. One might think that this clarity, this ease of synchronic individuation, generalizes. No matter what I think shall become of me tomorrow (or after my brain transplant, or after my teletransportation, or after my death), it is clear, at least, that I am a single person, with a single mind, now—isn't it?

But it isn't. Familiar physical objects are one thing, and other sorts of things are something else. Indeed, what is puzzling and difficult about some diachronic identity problems may be explained at least in part by questions about the synchronic individuation of the entities at issue.¹¹ A *ship* might be easy to at least *individuate* (at a time). The Vienna Circle is less so; if one knew the principle of individuation for such a group (if a group is what it was), then one might know—or at least stand a better chance of knowing—under which counterfactuals the Vienna Circle continued, or ceased, to exist.

Minds are more complex than ships. Additional problems arise for the individuation of minds that don't arise for the individuation of molecules, sporks, planets, and other physical objects. This is because, as I argue in this work, minds are *systems*. Because a system is a collection of parts with a certain, possibly quite

¹¹ Indeed, Nozick (1981) argues that what he calls “entification” (following Quine, 1964: 1) and what I call individuation “takes place in one fell swoop” (85); temporal considerations enter into even synchronic individuation, since causal connectedness is an element of synchronic unity or individuation. This is certainly true to some extent of individuating minds (and probably true of individuating mental systems generally) and the treatment of individuating minds and other mental tokens that is given in this work is of course sensitive to extended patterns of interaction and functional activity.

complex, functional, active organization, the *borders* and *boundaries* of systems are less clear than the borders and boundaries of, for example, sporks.

So there are no doubt individuation questions about lots of systems: is the U.S. fighting a single war against terror, on multiple fronts, or are we currently fighting multiple wars? What are the boundaries of an ecosystem, and how big or inclusive can one be? What makes something one storm rather than two? One root system or two? One legal system or two?

It may be difficult or even impossible to give precise criteria of identity or individuation for all sorts of entities, and the best criteria may be contextual, and sensitive to a wide range of (not necessarily obvious) considerations and interests. Thus I do not attempt, in this work, to give criteria for the individuation of minds that will work in all context. (Much less do I try to give general criteria for the individuation of all complex systems!) Instead I merely try to work out how to individuate minds within one particular group of subjects. And even there, I abstract and idealize to some extent; there are significant individual differences among “split-brain” subjects, in part because there is not a single kind of “split-brain surgery.”

In light of the no doubt great difficulty of giving individuating criteria for all sorts of complex systems, we should consider Rey’s (1997) “Fairness Maxim” for philosophy of mind:

DON’T BURDEN THE MIND WITH EVERYONE ELSE’S PROBLEMS. Always ask whether a problem is *peculiar* to the mind, or whether the issue could equally well be raised in other less problematic areas. If it can be, settle it for those areas *first*, and *then* assess the philosophy of mind.

(Rey 1997: 27; original emphasis)

The problems of identity and individuation raised in this work are, on the one hand, hardly unique, insofar as there will probably be some easy cases and numerous hard cases for the individuation of *any* sort of complex system. Moreover, some of the features that make something a hard case for individuating *minds* may make something a hard case for individuating other complex systems as well. One feature which makes for a hard case in the instance of individuating minds, for example, is *shared states*: i.e., some (mental) states of one putatively distinct system appear also to be states of another putatively distinct system. But two putatively distinct departments could share some members, or some administrative offices, and this could complicate those departments' individuation as well. More importantly, how to individuate different sorts of systems will hinge upon the same sorts of questions at a very abstract level. In particular, as I argue in this work, questions about individuating minds, and no doubt about individuating many other sorts of complex systems, are largely questions about functional and causal *independence and interaction*.

On the other hand, knowing how to individuate *one* sort of system, that is—for instance a root system—won't necessarily tell you how to individuate *another* sort of system—a legal system. To say that two sets of parts constitute distinct systems if those sets operate independently of each other in some requisite way and/or to some requisite degree, or that they are part of a single system if they interact in some requisite way and/or to some requisite degree, is to offer no more than a very abstract schema for individuation. For what counts as independence or interaction of the *requisite* sort or degree, will vary depending on the *type* of system in question.

Consider, for illustration, some of the individuation problems that arise for living organisms. The philosopher of biology Wilson writes:

The theories of individuation generated by considering only a narrow and conventional range of examples prove inadequate when applied to real living things whose normal modes of existence include complex metamorphoses, regeneration of lost parts, splitting apart and fusing together. A clonal population of the fungus *Armillaria bulbosa* occupies at least fifteen hectares in a Michigan forest. Some mycologists have called it the largest individual living thing on earth. What are the grounds for this claim? Some species of rhizocephalans, a group of parasitic barnacles, have several distinct developmental phases. Is each phase a separate individual or do they collectively compose an individual? Strawberries can reproduce through sexual or clonal reproduction. Is each clone an individual or does the entire set of clones compose an individual? Or are both individuals? Questions like these cannot be answered satisfactorily by a theory that treats the characteristics of a higher animal as the necessary and sufficient conditions of individuality. In fact, cases like these raise the question of whether there *are* necessary and sufficient conditions for individuality *simpliciter*.

(Wilson, 1999: 2, emphasis added)

Rather, a thing's individuality may be a matter of the type of thing that it is, and in some cases may be best determinable from within a science *of* that sort of thing. Wilson explains for instance that although the "true jellyfish" (a scyphozoan jellyfish) and the colonial siphonophore resemble each other physically, and have "essentially the same functional structure" (1999: 7). But from a biological perspective other things may matter also: developmental history, evolutionary origin, mode of reproduction, and so forth. How organisms such as colonial jellyfish are individuated may depend upon—or for that matter influence—the outcome of debate between the "gene's eye view" and the "organism's eye view" of evolution (Sterelny and Griffiths, 1991). Advocates of the "organism's eye view" of evolution will find it

important to determine exactly what qualifies as the life cycle of a single organism, and may determine this on the basis of factors which the “gene’s eye view” theorist finds irrelevant. Advocates of the “gene’s eye view” of evolution meanwhile take cases like these, in which it is difficult to determine what constitutes an individual organism, precisely as reasons for rejecting a view of evolution according to which phenotypic traits of *whole organisms* are the units of natural selection.

In other words, although individuation problems are hardly unique to psychology, there needn’t be a general problem of individuation admitting of a general solution. Rather, how to individuate organisms and the life-cycles of organisms may be an issue best resolved by a science of living organisms. That is why the work of this dissertation really is unique to philosophy of mind. Limited analogies between minds and corporations, between the two “disconnected” hemispheres and the circulatory systems of conjoined twins, or whatever, may be of some use. But ultimately how many minds split-brain subjects have really is a problem for psychology, and for philosophers of psychology, in particular. This dissertation explores the individuation of mental entities—in split-brain subjects in particular—from a scientific, theoretical perspective: that of cognitive psychology.

The analogy to individuation problems in biology above, however, should not be taken to imply that *all* individuation problems are scientific problems, problems best answered from a theoretical, scientific perspective. Some individuation issues may be proper to the legal domain, some to the philosophy of art. And arguably there are some individuation problems regarding split-brain subjects that are not best answerable, or are not determinable at all, from the standpoint of a scientific

psychology. In particular I believe that questions about the individuation of *persons* (in split-brain subjects and in general) are not proper to the domain of scientific psychology in the way that questions about the individuation of minds are. Of course many findings from scientific psychology may be *relevant* to the individuation of persons. But the concept of a person appears to be one with inextricably normative or as Locke (1690) once said “forensic” aspects, and of profound import to many non-scientific domains of human activity and understanding.

Of course, minds and persons are conceptually linked in various ways, and indeed questions about individuating minds in split-brain subjects have often been raised simultaneously with questions about personhood in split-brain subjects. In fact much of the philosophical and even neuropsychological literature on split-brain subjects has treated questions about individuating minds (and streams of consciousness) and questions about individuating persons interchangeably. That how many minds split-brain subjects have and how many persons they are have so often been treated as interchangeable questions is entirely understandable, as the concepts *mind* and *person* are again obviously deeply related. They are nonetheless distinct concepts; most obviously, a creature can have a mind without being a person.

There is therefore the potential for becoming badly confused by many of the claims that are defended in this work, if the distinction between minds and persons is ignored or forgotten. For this work is about the individuation of entities that arguably *do* fall within the proper province of scientific psychology: it is about minds, mental states, streams of consciousness, and so forth—but not about persons. I do have my own (not very definitive) views on personhood in split-brain subjects and in general,

and in Chapter Three I argue that the mental duality or “two minds” model for split-brain subjects may be compatible with multiple ways of understanding personhood in split-brain subjects. But the nature and identity of persons are excluded from this dissertation to the extent possible, because the concept of a person does not seem to be proper to scientific psychology. (Again, even if findings from scientific psychology contribute to our understanding of persons.) And I’ve tried to keep this work within theoretical psychology, not because non-scientific perspectives are less interesting, but simply to see what can be said within this one, and to provide a kind of unity to dissertation.

Finally I want to say something about the relationship between the classification, or type-identifying, and the individuation, or token-identifying, of entities. Intuitively these activities can be distinguished: we can ask for instance whether something is an artifact or a natural object, while still knowing that whatever it is, it is an individual; or, talking about what are uncontroversially meteorological phenomena of a certain sort, we can ask whether we’re looking at one cold front or two. But as the foregoing discussion implied, asking either one of these questions requires *some* sense of both classification and individuality. For in both cases, we pick out some portion of the world from the rest of it to ask about, and in both cases, we pick that portion to ask about on the basis of features that contribute to making it the type of thing that it is. This is in part to say again that the criteria for individuating a thing depends upon the kind of thing that it is.

Still, classification and individuation can be distinguished in part methodologically, for they differ with respect to what they take as given and what

they take as unknown. These correspond to two ways of approaching the “how many minds?” question for split-brain subjects, for example. One approach takes as given the token-identity of a *physical* system qua physical system, and asks questions about its mental identity; a second approach takes as given the criteria of individuation for a *mind*, and then asks what physical stuff meets those criteria:

Approach one: Given a split-brain subject’s *right hemisphere* (i.e. single cerebral hemisphere), and/or right hemisphere plus subcortical structures and so forth (but minus the subject’s left hemisphere), is *that single thing* a mind? (Or is it *not* a mind in and of itself—is it just *part* of a mind?)

Approach two: Given the criteria for being an individual mind, what meets those criteria in a split-brain subject? The whole organism? The organism’s nervous system? The organism’s left hemisphere? Left hemisphere and right hemisphere, individually? Left hemisphere plus subcortical structures? Left hemisphere plus right arm?

Either approach should get you to the same answer.¹² The former approach just puts the issue in terms of classification or type-identity of physical tokens, and the latter puts it in terms of individuation, or the entokening of mental types. But each approach presupposes an answer to the other approach’s problems. They simply differ in what they take as given versus what they do *not* take as given. The former approach presupposes the individuality of a physical thing qua physical thing and tries to determine its mental individuality; the latter presupposes an account of mental

¹² These are not the only two ways of approaching individuating minds in split-brain subjects. Another approach—one that probably has a longer, more venerable tradition—would be to ask of a set of token *mental states* whether or not they all belong to the same mind—whether or not they are all *co-mental*, that is. To some extent I take this third approach in this work, but I do so in a way that’s parasitic on the first kind of approach, because I ask whether *right hemisphere* and *left hemisphere* mental events belong to the same mind or not. *Some* way of picking out the mental stuff in question is needed, that is. If you can’t pick them out by reference to an already individuated subject of experience, for instance, or an already individuated mind, you refer to physical locatedness.

individuality and tries to find physical stuff that, relative to or under such an account, constitutes an individual.

If there seems to be a certain circularity here, there is. It is the same circularity, I think, that is necessarily involved in the construction of a functional, psychological theory of the brain (or a functional, psychological theory of whatever is the supervenience base of mind). And it is the same circularity that is necessarily involved in the construction of a realistic psychological theory—a psychological theory that is true. What makes an individual physical thing a mental thing is something—a functional role—it *shares* with many other physical, mental things. And what makes a kind of mental thing the very mental token that it is, is in part a matter of its physical properties. But this last is something I will have to argue for, as I do in Chapter Three. In the meantime, since the first approach—which asks about the mental identity of physically individuated hemispheres—is more familiar from the literature on the split-brain phenomenon, this is the one I start with as well. I shall move back and forth between the two approaches, however, as is inevitable, again, given the relationship between these approaches and the questions they ask.

2.3 On Kinds of Questions

I refer from time to time to the “how many minds?” question regarding split-brain subjects. That question asks whether a split-brain subject has one mind, or two. But someone might object, in the spirit of Parfit (1971, 1984, 1995, 2003a, 2003b), that though there are many facts about mental states in split-brain subjects, there needn’t be any *further* facts about *minds* in such subjects, and the subjects needn’t have any determinate number of minds. (If facts about minds reduce to facts about the things

that compose them, then the split-brain case may simply be one in which the lower-level facts just don't yield a determinate number of entities at the higher-level.) And someone might add, in that same spirit, that the "how many minds?" question is therefore a trivial or a merely verbal one—one that could only be answered by arbitrary decision.¹³

But while it is certainly possible that a split-brain subject ultimately does have an indeterminate number of minds, the debate about how many minds split-brain subjects have is still interesting and important, for several reasons. To begin with, it is in part empirical. There are many empirical facts about cognition in split-brain subjects relevant to resolving how many minds they have, such as the nature and extent of interhemispheric subcortical interaction in such subjects. For that matter, there are empirical questions about *all* subjects that are relevant to answering the question, too, questions about memory and the structure of emotional states and so forth. I hasten to add that of course many of the empirical facts relevant to determining how many minds split-brain subjects have are of course highly conceptual or theoretical at the same time. Still, it is no more true that the empirical facts are entirely in, than it is that the empirical facts can be determined by observation, alone.

The debate about how many minds split-brain subjects have is still admittedly largely a conceptual one. But conceptual questions can be quite interesting. Among other things, these are questions about our conceptual framework, and the role a given concept plays within that framework. Of course questions about *terminology* are

¹³ Parfit is concerned with the identity of persons rather than the individuation of minds, but many of the conclusions he draws about personal identity—and the arguments he makes to support those conclusions—have possible analogues in positions on individuating minds.

rarely so interesting. But it is wrong to equate conceptual with verbal questions (as Parfit sometimes seems to); questions about how to think about something, or about how we *do* think about something, are not like questions about what *word* to use to refer to something. Among other things, terminological questions can be resolved by decision to a greater extent than can questions about our concepts. We can and do change our concepts, of course, just as we can and do change our beliefs; but just as one can't simply decide to change one's beliefs and have done with it, one can't always change one's concepts simply by deciding to do so. (One can just decide to call brooks "creeks," regardless of how one thinks about such moving bodies of water.) And even when conceptual questions can't be resolved by anything but decision, the decision often needn't be *arbitrary* in the way many terminological decisions are.¹⁴

So, although there may be no absolute distinction between empirical and conceptual questions, even those conceptual debates that are most distant from the empirical, that seem least capable of being resolved by observation, may still be interesting conceptual questions. Furthermore they may matter, simply because they matter to us. For instance, although I can imagine the concept of a *person* not figuring implicitly or explicitly into any empirical or scientific theory at all, the concept is arguably still a central one to many domains of thought: central morally, legally, socially, and simply pragmatically. The concept of a mind could also be similarly important in some or all of these domains.

¹⁴ I don't mean to draw too firm of a line between verbal and conceptual questions or decisions, either, though; terminological disagreements often aren't wholly arbitrary, and when they aren't, it is generally *because* different terms have different conceptual connotations that are not easily divorced from them. What might appear to be *purely* terminological disputes, in other words, in fact often turn on contested concepts.

But many conceptual questions are also *scientific*. Indeed one of the things I do in this work is explore the significance of the concept of a *mind* (among other concepts) to psychology (and to neuropsychology). There is at a minimum one sense in which the “how many minds?” question appears to be an important conceptual question, from the standpoint of scientific psychology. As a conceptual question, it concerns, in part, the relationship between neural and mental properties, entities, and events, as will be made clear over the course of this work. One can already see in some of the literature on how many minds split-brain subjects have several attempts at answering that question whose general logic, if accepted, would have a negative impact on psychological theorizing.

And although it is too soon to say whether the term “mind” or some successor of it will have any role to play in future psychological theories, still, even if it didn’t, the concept of a mind might still play an implicit role in some such theories. Most simply, the concept of a mind might play the role of simplifying our thought about the smaller entities and activities out of which they’re composed. Because minds are systems, their behavior is “emergent” (in an unspooky sense) relative to the behavior of the things that compose them: their behavior, that is, is not just the sum of the behavior of its component parts, but also a function of the interactions between them. I suspect, moreover, that the concept of a mind plays an important, if always implicit, role in *functional* analyses of psychological phenomena: we need some sense of the boundaries of minds in order to correctly characterize the functional type of some token mental states. This point will be returned to and explained towards the end of Chapter Three.

In this section, I discuss some basic criteria for mindedness: minimal criteria that something must meet to qualify as (at least) a (single) mind. I also compare purely psychological to partly neural criteria for mindedness. In this section, though, I focus on the *type* of things minds are; only in the next section will criteria for individuating *token* minds begin to be developed.

3.1 Modest Criteria for Mindhood

Minds are systems: structurally complex entities that are composed of parts organized in such a fashion as to yield the operations characteristic of a mind. But what sorts of operations or activities? What sorts of properties must a mind, a complex mental system, possess?

There is quite a lot of dispute about this question. Some make the criteria for possessing a mind very stringent, believing that to possess a mind one must possess natural language (see for instance Davidson, 1975, 1982; Dennett, 1996). According to such philosophers (they are mostly philosophers) having a mind requires having a mind a lot like ours. At the other extreme there are those who argue that even insects have minds (Carruthers 2004a), even if minds fairly differently (functionally) designed from ours. As with most things, most philosophers and neuropsychologists probably occupy an intermediate position on the question of what criteria a mind must meet. But there are numerous positions occupying even this middle ground, and numerous disagreements about the importance of consciousness and self-consciousness, for instance, or about whether minds are massively modular or rather

characterized by a capacity for the “promiscuous” interaction of informational (or belief) and directional (or goal or desire) states.

Fortunately, most of these disputes are orthogonal to current purposes. For again, the primary focus of this work is the individuation of mental tokens, including minds—and criteria for mindhood are not equivalent to criteria for individuating minds. Note that one could be certain, for instance, when faced with two mental systems, that either one of them, in the other’s absence, would constitute a unique mind—and yet arguably not know whether either of them *did* constitute a unique mind. For whether multiple mental systems are part of a single mind on the one hand or constitute distinct minds on the other depends upon the functional relationship between them. What is most relevant to current concerns is what kinds of functional relationships *between* mental systems suffice for those systems to jointly constitute a single mind.

That said, the criteria for mindedness are of course not irrelevant to individuating minds. So not *all* of the disputes hinted at above, concerning the proper criteria for mindhood, are orthogonal to my purposes. The split-brain subjects constitute an interesting “hard” case for the individuation of minds primarily only if both hemispheres of a split-brain subject would meet the criteria for mindhood in the other’s absence. And this is the claim that I need to move forward: that *in the absence of the other hemisphere*, either hemisphere of a split-brain subject would constitute (possibly in conjunction with some non-cortical structures) a mind. The tenability of this claim requires the rejection of the most stringent criteria for mindhood—most

especially those requiring possession of full-blown (propositional) natural language capacities.

Fortunately, that the right hemisphere of a split-brain subject is at least a *candidate* for being *some* kind of mind, is not as controversial as the positions that I will be arguing for. Most philosophers and neuropsychologists do believe, that is, that the right hemisphere of a split-brain subject does function in such a fashion that it would, in the absence of the left hemisphere, support a mind “comfortably characterizable as human” (Marks, 1981: 47, fn. 18). This right hemisphere mind would not be capable of all the things of which a human mind (or a human left hemisphere mind?) is capable, again particularly productive propositional language. But the right hemisphere can certainly see, hear, smell, and feel, and learn and remember. Right-hemisphere controlled behavior reflects the presence of right hemisphere beliefs and desires, likes and dislikes—even the presence of a sense of humor, and of recognizable and appropriate emotions, from fear to pride to slightly embarrassed self-consciousness. Under control of their right hemispheres, split-brain subjects respond to experimental instructions, can draw objects from memory, can indicate how positively or negatively they feel about their lives and themselves (Schiffer et al., 1998), and can engage in abstract (non-verbal) reasoning (Zaidel, Zaidel, and Sperry, 1981). That the *left* hemisphere would support a mind in the absence of the right is meanwhile of course never called into question, in large part due to that hemisphere’s capacity to speak for itself.

In moving forward I will assume that in the absence of the left hemisphere, the right hemisphere of a split-brain subject would (again, possibly in conjunction

with non-cortical structures) constitute a mind. Partly in order to justify this assumption, I adopt a modest set of criteria for mindhood, that says that minds are information-integrating systems, mental architectures in which belief (or “informational”) and desire or goal (or “directional”) states interact to yield reasoning and decision-making, and that engage in various other paradigmatic mental processes, such as perceiving and learning. I do not enter any further into the disagreements between the various criteria for mindhood referred to above.¹⁵

3.2 Purely Psychological versus Partly Neural Criteria for Mindhood

The criteria for mindhood referred to up until this point, including the very modest one I just advocated, have thus far all been given in purely psychological or functional terms. As minimally described, these functional accounts of what it is to be a mind make no commitment as to what *implements* a mind’s functional architecture. But some constitutive accounts of mind would identify minds explicitly with brains or with nervous systems—or, rather, with *whatever* physical part of a creature implements its various mental processes, processes like perceiving, remembering, decision-making, and so forth.

Of course, a partly neural account of mindhood in and of itself might not seem to rule out much except dualism. (Compare this to an account of mindhood that views

¹⁵ See Rey 1997 on “modest mentalism” for a sense of the modest criteria for mindhood that I accept. Of course the modest criteria given above might be too modest in certain respects; it would be natural to object that additional constraints are needed e.g. temporal constraints: a mind must make decisions in real time, that is—quickly enough to *survive* in this world—and not just over the course of millennia. I’m sure the account *does* leave out some further constraints, but my intent isn’t to defend this one as adequate *for all purposes* as is. I think it is adequate for my purposes, however, for my intent is simply to say that on some modest, reasonable set of criteria for mindedness, the RH of a split-brain subject would meet it in the absence of any LH mind. Indeed this is an understatement: the RH meets the criteria for being a characteristically human mind, if not an entirely “normal” human mind.

natural language as central to genuine mentality; such an account rules out quite a lot of things as not viable candidates for mindhood.) And of course a partly neural account of mindhood is perfectly compatible with a very wide range of functional criteria for mindedness. Indeed a partly neural set of criteria for mindhood would still necessarily be *partly* functional or psychological. While one could adopt purely psychological set of criteria for mindhood, a purely neural set of criteria does not seem to be similarly possible. The closest one could come to providing a *purely* neural approach to individuating minds would be to say that a mind is a functioning nervous system (or a functioning brain). But, first of all, this might include too much; the *Aplysia* has a nervous system (probably the best understood nervous system of any living organism, currently), but I suspect many would balk at attributing a mind to it, because it is not clear that one needs a *psychological* theory (as opposed to a neuroscientific theory) to explain its behavior. If, on the other hand, one tries to insist on some more stringent reading of *functioning nervous system*, this is presumably an attempt to sneak in some reference to psychological capacities and behaviors after all. Finally, there could be some argument about how to individuate nervous systems. (In fact questions about how to individuate nervous systems inform the debate about how to individuate colonial jellyfish as mentioned earlier.) And surely that question depends *in part* on the special functions of nervous systems. And once the system reaches a certain level of complexity, its functions are psychological ones: nervous systems process, transform, and communicate information (representations); they remember, they perceive, and so forth.

Although a partly neural set of criteria for individuating minds would be compatible with a wide range of very different functional criteria for mindedness, the choice between a purely psychological and a partly neural set of criteria for mindhood is still significant, however, because the choice of criteria, here, influences or may even determine the choice of *individuating* criteria. One who adopted a purely functional account of mindhood, for instance, might similarly adopt a purely functional set of criteria for individuating minds. (*Which* functional criteria those should be will be discussed in the next section.) But one who adopted a partly neural account of mindhood would presumably adopt a partly neural method of individuating minds, as well. The difference between a purely functional and a partly neural set of criteria for mindhood is just that the former would acknowledge (at least the possibility of) non-physical minds, while the former wouldn't—a difference with no pragmatic significance, perhaps, whatever metaphysical (and theoretical) import it may have. The difference between a purely functional and a partly neural set of criteria for individuating minds is (at least *prima facie*) among other things of more methodological import, however. From the perspective of a purely functional set of criteria for individuating minds, the absence of a corpus callosum in a split-brain subject is not necessarily even relevant to determining how many minds such a subject has. From the perspective of a partly neural set of criteria, however, that absence, that neuroanatomical fact, has *prima facie* significance.

The partly neural account of mindhood is the one that neuropsychologists who've looked at the split-brain phenomenon tend towards, and the account probably reflects the sort of approach to psychological phenomena that they tend towards in

general, understandably, given their professional orientation. This sort of account is again perfectly compatible with the modest functional criteria for mindhood that I've said I will be assuming. The two sorts of account can easily be combined simply by saying that a mind is that physical part of a creature that bears its informational and directional states and sustains or implements their interacting, for example. Or, more generally, a mind is that physical part of a creature that implements its mental architecture.

Such a partly neural account of mindhood might seem to entail that a split-brain subject has two minds. For each hemisphere contains the mechanisms for memory and attention, each is subject to emotion, sensation, perceptual experiences of all kinds, and each sustains the interaction of informational states with each other and with directional ones, in order to guide behavior. Each hemisphere *thinks*. The activities each hemisphere engages in suggest that each hemisphere instantiates the mental architecture necessary for mindedness. If a mind is just any physical part of a creature that embodies or implements its essential mental capacities, then each hemisphere is probably associated with a distinct mind.

This partly neural account of mindhood, in other words, might be used to construct the following argument for mental duality in split-brain subjects. Call this the "argument from hemispherectomy":

P1: In the absence of the right hemisphere, the left hemisphere would constitute a unique mind; in the absence of the left hemisphere, the right hemisphere would be identified with a mind.

C: Therefore the right and left hemisphere are each associated with a mind.

The argument has a hidden premise, one that I think makes it, and the method of individuating minds that it assumes, unacceptable. Still there has been some temptation towards this first partly neural method of individuating minds; one can see it in Puccetti for instance (1981), and even Nagel flirts with it when he says:

If we decided that. . . [split-brain subjects] definitely had two minds, then it would be problematical why we did not conclude *on anatomical grounds* that everyone has two minds, but that we do not notice it except in these odd cases because most pairs of minds in a single body run in perfect parallel due to the direct communication between the hemispheres which provide their anatomical bases.

(Nagel 1979: 162; emphasis added)

In this vein, someone might suggest that, actually, if we're going to identify every physical part of a creature that *could* meet the criteria for mindedness with a distinct mind, then split-brain subjects, and "normal" subjects as well, probably have *three* minds apiece: one mind associated with the left hemisphere, one with the right, and one with the whole brain.¹⁶

¹⁶ Rey (1975) once made just this proposal (though for persons rather than minds), asking whether split-brain subjects, and "normal" subjects as well, might indeed each be associated with three persons: a left hemisphere person, a right hemisphere person, and a third person jointly constituted by the previous two. I think, however, that Rey was not using this first method of individuating minds being discussed here, but rather the one I will discuss in a moment, which identifies a mind with that physical part of a creature that in fact functions as a distinct mind. Rey suggested, that is, that both the right hemisphere, and the left hemisphere, and the brain as a whole, may all function as unique persons (or in this context minds).

Despite its initial apparent counter-intuitiveness, this position has a lot to recommend it, and accords well with the principles for individuating mental tokens that I advocate in this work. While I argue that the two hemispheres largely lack the kind of integration that characterizes a single mind (and thus adopt a mental duality model for split-brain subjects), I do this primarily in the context of arguing against a mental singularity claim for split-brain subjects. Depending on how the empirical facts come

The proper response to this suggestion, and to the argument from hemispherectomy, and to this first possible method of individuating minds as physical systems, is to insist on the distinction between what is counterfactually the case about a hemisphere, for instance, and what is actually the case about this hemisphere. That is, what *might* be a distinct mind given certain counterfactuals may nonetheless not *actually* constitute a unique mind, in and of itself, but may rather be only one of multiple mental systems that jointly constitute a single mind. It is no doubt true (and I am assuming) that *in the absence of the other cerebral hemisphere* (following hemispherectomy, say), either hemisphere of a split-brain subject would (perhaps in conjunction with some non-cortical structures) constitute a single mind. But this does not necessarily show that in the *presence* of the other hemisphere, a single hemisphere of a split-brain subject constitutes a unique mind.

Most obviously, in the absence of a second hemisphere, either hemisphere might undergo significant re-organization, physically and functionally, in the course of which it might come to take on roles that it did not play when there was another hemisphere to play them. But this says nothing about what either hemisphere is *actually* doing *without* such reorganization. Surely a mind is something that *does* (for instance) reason and learn and perceive and so forth, rather than something that merely *could* come to do all those things, given significant reorganization.¹⁷

in and are conceptualized, however, a three mind model (mental trinity?) *could* be best. But I don't sort out how such a model would work in this dissertation.

¹⁷ Note that I say that a mind is something that can reason and remember and so forth *without significant reorganization* to make sure that minds *asleep*, or minds severely intoxicated, or whatever, are still minds. For something needn't necessarily be *currently* reasoning or perceiving or remembering in order to constitute a mind. But it needs to be the case that the system in question *would* reason and perceive and remember and so forth given the *relevant* counterfactuals. Of course, giving the precisely relevant counterfactuals is challenging at this point, in part due to the absence of a developed psychological theory that would help specify them. But I take it that the relevant counterfactuals won't

But even if either hemisphere could support a unique mind in the absence of the other *without* significant reorganization, arguably this *still* does not show that the two hemispheres in split-brain subjects (or in “normal” subjects) are associated with distinct minds. For in each other’s presence, and with or without a corpus callosum between them, the two hemispheres might still *function* as one mind. (Compare: two people who are perfectly capable of acting as unique agents may nonetheless right now be acting as a single agent.) And, indeed, few neuropsychologists who have believed that split-brain subjects have two minds have believed that “normal” (non-split-brain) subjects do also. (Bogen (1990) is an exception, but Bogen’s reason for attributing two minds to a “normal” subject is not Nagel’s, above; Bogen is motivated by particular views about the *kind* of interhemispheric interaction the corpus callosum affords, and by particular views about the kind of interaction *necessary* to integrate what would otherwise be two minds into one.)

So it is too simplistic to say that any physical part of a creature that implements all the psychological processes a mind must implement in fact constitutes a unique mind. How that physical system relates to other physical systems implementing those same processes is also relevant. This is a crucial point: hemispherectomy data, for instance, while relevant to determining how many minds split-brain subjects have, isn’t decisive to that determination. One who wants to claim that a split-brain subject has two minds must show not only that either hemisphere *could* meet the criteria for mindhood independently of the other, but also that the two “disconnected” hemispheres in fact *function* as distinct minds.

be things like, “Given an additional six months of pre-natal development” or “given significant neural rewiring,” for instance.

This suggests a second partly neural method of individuating minds. First, determine which physical parts or systems in a creature implement the psychological processes and organization characteristic of a mind. Next, determine which functions are necessary (or at least sufficient) to integrate what would otherwise be multiple distinct minds into a single mind—a single goal-and-information-integrating system, for instance. Then look for the structures supporting or implementing those functions. If *those* structures are absent between the mental systems in question, those systems are each distinct minds.

Note that there is a circular relationship between individuating psychological entities and individuating neural entities within this partly neural approach. (This explains why there can be no *purely* neural or physical method of individuating minds.) True, the approach uses neural facts—facts about physical connections and their absence—to individuate mental entities. But the neural facts that matter, matter because of their psychological identity. So using facts about particular neural structures to individuate minds isn't justified by those structures' purely physical features (!), such as the shape or size or location or even physical connectivity of the structures. The role those structures play in individuating minds must be justified in part by their functional significance. So first, the neural is psychologized, or given a psychological taxonomy. Only *then* can purely physical features of the structures—such as, their being intact, versus their being severed—be used to individuate minds.

A defense of partly physical criteria for individuating mental things can be found in the account of persons taken by Wiggins (1976), who suggests that the concept, 'person' is defined both in terms of natural kinds and in terms of a functional

specification, along this sort of line: x is a person if x is an animal that perceives, feels, remembers, imagines, carries out projects, and so forth. So “On this account *person* is a non-biological qualification of *animal*” (161; original emphasis). Wiggins also insists that the individuating criteria for a person can’t be distinct from the individuating criteria for an animal (albeit an animal possessing certain psychological properties), although one might be able to argue with him on this point. (Perhaps a person is necessarily constituted by but not identical to an animal, for instance.)

Along similar lines, Rey (1977), citing Wiggins (1976) turns from a purely psychological criterion for survival in part because the central psychological criteria for survival require *causal links* of the right sort (between one’s mental states, and between one’s mental states and events in the world), and causal links in turn require (realizing) substance. But some substance matters more than others:

What brings the bodily and psychological criteria together is the fact that there are *certain* parts of our bodies that are crucial to important part of minds. . . Not all our embodiment is, that is, strictly personal.

(Rey 1977: 58; original emphasis)

In other words, mental and physical criteria for survival converge on the brain. So this too seems to be a philosopher’s analysis of a (partly) psychological kind, according to which the criteria for individuating tokens of that kind are partly physical (in this case neural). Yet the physical system in question must still meet certain psychological criteria.

Since few contemporary philosophers of mind are dualists (much less substance dualists), one might think that *all* who have looked at the split-brain

subjects and wondered how many minds they would have used a partly neural set of criteria for individuating minds in such subjects. This sort of account of mind, after all, still can't help but be deeply psychological. Yet my impression is that philosophers of mind as a group have been somewhat less likely to adopt this account of mindhood than have neuropsychologists. Or, at least, they have been somewhat less likely to adopt a partly neural set of criteria for individuating minds. (Tye (2003), for instance, explicitly states that a mind is (constituted by) a brain (or perhaps a brain's global organization, for one brain might have several such organizations). But his criteria for individuating minds, at least in split-brain subjects, are purely psychological.) Philosophers have had a greater tendency, in particular, to deny that split-brain subjects' neuroanatomy can be used to provide any support for a mental duality claim for those subjects. Marks (1981) and Tye (2003), as we will see, deny that lack of a direct causal relationship between two sets of neural events poses any obstacle to identifying them with a unique mental token; these facts are (for Marks and Tye) at most facts about how minds are implemented, and not about mental architecture in particular. These philosophers have argued for what I will call a *purely psychological* method of individuating minds. They have rejected the use of facts about neural connection and disconnection, neurally mediated interaction or lack of interaction, to determine how many minds a subject has.

Such arguments will be examined at various points throughout this work, but especially in Chapter Three. While the arguments in question make important contributions to the literature on the mind-brain relationship and on individuating mental tokens specifically, I reject them, in favor of the neuropsychologists' often

unspoken and usually undefended assumption that neural facts offer constraints on mental individuation. Again, they only do so in a particular way: it is easy to use physical independence to argue for the presence of distinct minds for example only if you have the *right sort* of physical independence—but the *right sort* of physical independence is partly *mental* to begin with.

Nonetheless, the claim that neural facts are *just* implementation facts, and not facts about mental architecture as well, is arguably poorly motivated in the end. Of course it is very difficult to see at what level of abstraction a developed psychological theory will be pitched, and therefore at what level of abstraction the constitutive conditions for minds (and other mental phenomena) will be given. (If all of psychology ultimately ends up even being pitched at a single level of abstraction.) Still, I believe that the claim that neural facts are just implementation facts ignores the important relationship between mental and physical architecture of which the split-brain studies constitute obvious evidence. For co-mentality—the relation that multiple mental tokens bear to each other when they either jointly constitute or all belong to one mind—requires, in fact consists in, a rich causal integration of cognitive states and processes. This causal integration has to be achieved via *some* physical medium—it has to *consist in* some kind of *physical* interaction—and that physical medium then poses additional constraints on identity. And as I ultimately argue in Chapter Six, neural stuff is, at least right now, the only physical medium capable of implementing the very subtle and intimate interactions that characterize mental processes.

If neural facts *are* relevant to individuating minds, then neuroanatomical facts about split-brain subjects provide some *prima facie* support for the identification of each hemisphere of a split-brain subject with distinct minds. This is not *just* because the hemispheres are not physically connected at the cortical level. It is because of the plausible psychological function of the corpus callosum: to integrate mental processing across the two hemispheres (at least to some extent). The split-brain experiments seem to provide some evidence that the corpus callosum is a major mechanism subserving interhemispheric mental integration in “normal” (non-split-brain) subjects. *If* that is correct, then, deprived of that mechanism, the hemispheres should each constitute a distinct mind.

Still, one might object that even if the corpus callosum is *a* mechanism of (for) interhemispheric co-mentality, one that a split-brain subject of course lacks, there are many other means by which the two hemispheres of a split-brain subject interact. For the two hemispheres of a split-brain subject are not causally independent or isolated from each other; there are all sorts of causal connections between right and left hemisphere mental states. Might these remaining mechanisms that sustain interhemispheric *interaction* in the split-brain subject constitute mechanisms for *co-mentality*? To answer this, we need to know what *sorts* of interhemispheric causal interactions remain in split-brain subjects—and whether they’re the kind that count.

Whether multiple mental systems, each of which would, on its own, meet the criteria for mindhood, do in fact constitute distinct minds, or whether they instead jointly constitute a single mind, is a matter of the *kind* and *degree* of causal integration between them. Given the right sort of sufficiently intimate (and regular)

interaction, those mental systems will plausibly constitute a single mind. And given the right sort of functional independence, those mental systems will each constitute a distinct mind. The questions to ask now, are: what *kind* of interaction? What *kind* of independence? For the two hemispheres of a split-brain subject function independently in some senses but not others, and they interact through a variety of means—but do they interact through the means relevant to individuating minds? These are the questions I turn to in the next section.

4 *Individuating Minds*

We have just concluded that whether multiple mental systems each constitute a distinct mind, or whether they instead jointly constitute a single mind, is a matter of whether and to what extent they functionally interact. It might seem that we are now in a position to determine how many minds a split-brain subject has. But we are not quite there yet. For we have not yet specified the kind and degree of independence and interaction that make the difference between co-mentality and multiple mindedness. For the human mind is far from an island: its states interact, causally, with the states of many other minds. In this section, therefore, I try to say more about the kinds of interactions between mental states that make (or at least tend to make) those states co-mental.

4.1 Inter- versus Intra-mind Interactions

In order to try to uncover some principled distinction between intra-mind and inter-mind interaction, let us start, in Gricean fashion (Grice, 1965), by examining what

seems to be a prototypical case of inter-mind interaction. Consider the intuition that I and my sister, for instance, do not share a single mind, however many minds each of us has individually. The “how many minds?” question cannot just be a question about behavioral unity and disunity, partly because my sister and I can be very psychologically similar, and very close, and behaving very cooperatively—let’s say we’re in business together, and working jointly on a project—and our behavior can therefore be very unified. Yet even if our behavior was extremely unified (each anticipating each other’s desires and objections and striving to meet them without being asked, perhaps we’re even finishing each other’s sentences), few would say that we share a single mind (or, relatedly, that the best, much less the only, way to explain our behavioral integration is to ascribe to us a single shared mind).

The reverse is also true: individual humans can behave in a disunified fashion, and yet we seem to have other explanatory resources at our disposal in such instances than just attributing to them multiple minds. It is true that scientists and philosophers may have a tendency to individuate entities of all sorts at the level at which its component parts behave cooperatively and at which there is competition between that entity and other entities identified at the same level. And of course any line drawn (non-arbitrarily) around a mind must be such that things (parts, states) within a mind are more closely related to each other, *in some respect*, than they are to things outside that mind (including things in other minds). But why think that the distance *between* minds must consist in competition or conflict? More precisely, why think that the relative distance of this relationship must manifest in *competitive or conflicted*

behavior between the agents whose minds they are?¹⁸ Certainly in our incredibly social and psychologically complex species, individuals often act to advance each others' interests and sabotage their own.

Since only the incredibly deprived human mind is an island—since the mental states of one individual are generally intimately causally related to those of many others—the difference between having one mind and having two can't be one of causal interaction versus independence in and of itself. The causal boundary between one mind and another, in other words, will not be absolute.

Answering the “how many minds?” question requires distinguishing what we might call *intra-mind* from *inter-mind* interactions. One possibility is that there is no distinction between intra- and inter-mind mental interactions as such, or at least no distinction we can draw in purely *psychological* terms. Perhaps the only thing that makes an interaction between mental states intra- as opposed to inter-mind is that it is an interaction between mental states of *one creature*, for instance, where a creature is individuated in terms of spatiotemporal location. But if there *were* a way to draw the distinction in purely psychological terms, we might be able to use this distinction to count how many minds a single creature has.

¹⁸ Note that there are probably several mundane senses in which the two hemispheres of a split-brain subject, or even a “normal” subject, might be said to compete, routinely, with each other. Consider for example the “horse-race” explanation for the “redundancy gain”: responses to bilaterally presented stimuli may be faster (on average) than responses to unilaterally presented stimuli because there are two mental channels—hemisphere right and hemisphere left—through either one of which the solution can be calculated and the response initiated—whichever mental system “finishes” more quickly can initiate the response, that is.

4.2 Deliberate versus Non-deliberate Interactions

Marks (1981) implies that the difference between inter- and intra-mind interaction (or “communication”) is the difference between deliberate versus unintentional or automatic communication, respectively. If permitted, split-brain subjects will engage in *cross-cuing* behaviors to integrate their behavior, even during experimental situations during which the hemispheres are denied their normal degree of perceptual overlap. Marks suggests that this cross-cuing represents a method of integrating what would otherwise be two minds into one, because there is no evidence that the cross-cuing behavior is *deliberate*. This behavior rather seems automatic and unintentional. So, Marks’ suggestion is that *intra-mind* interaction is automatic and unintentional; inter-mind interaction is via deliberate communication.

Cross-cuing behaviors in split-brain subjects are fascinating and extremely varied. The term “cross-cuing” itself has been used very broadly, to refer to anything from a split-brain subject using his left hand (under control of his RH) to spell out the answer to a question on the palm of his right hand, a question only his RH knows the answer to, to a subject frowning (generated by the RH) when her LH makes an incorrect guess. The left hemisphere shows an impressive (obviously not perfect) ability to “read” these cues and use them to help narrow in on the (roughly characterized) conscious contents of the RH. The cues are of very different sorts: spelling the answer on the back of one’s hand looks rather like a deliberate attempt at communication; frowning in response to knowing that one has just made a mistake (i.e. gave the wrong answer to a question) looks automatic and unintentional.

This particular way of marking the distinction between intra-mind and inter-mind interaction may be compatible with the individuation criteria and accounts of mind developed in the previous section. We can broadly define a *mental system* as a (functional and perhaps physical) thing organized in such a way as to allow the interaction of mental states. On this broad understanding, the human brain is a mental system, and so is a cerebral hemisphere, and so is an auditory cortex. My sister and myself (or at least our two brains) together constitute a mental system as well. Now suppose we are interested not just in mental systems broadly speaking but in whole minds in particular. So we say that a mind is a mental system that (among other things) integrates informational and directional states, that reasons and makes decisions, that learns and perceives, and so forth. These criteria for mindedness might serve to exclude my auditory cortex as a candidate for mindhood, but they *wouldn't* exclude my left hemisphere, my brain, or my brain *plus* my sister's brain. At least, they wouldn't necessarily exclude me and my sister together, for her mental states and mine can be integrated, and we can engage in joint decision-making, and so forth. Marks suggests, however, that the interaction of mental states *within* a mind must be automatic and unintentional, as opposed to taking the form of deliberate (and consciously engaged in?) communication. This, he must think, would prevent my brain and my sister's brain together from qualifying as *jointly* constituting a single mind. And he also thinks that this way of filling out the intra- versus inter-mind distinction shows that the two "disconnected" hemispheres of a split-brain subject *may* jointly constitute a single mind.

Imagine that a split-brain subject has just been asked a question, and that the subject's right hemisphere sustains the belief that the correct answer is "blue." The subject's left hemisphere sustains no belief about the correct answer, and so the patient, under control of her left hemisphere, verbally guesses, "Red?" The subject's right hemisphere hears the subject give the wrong answer, and initiates (or is somehow the origin of) a frown—not as a deliberate attempt of communicating that an error has been committed, but simply because it is displeased, or whatever, by the subject's error. Feeling the frown (and perhaps some negative affect), the subject, via her left hemisphere, immediately concludes that she (the subject) has made an error—and she corrects herself, "I mean—blue." Her left hemisphere now *shares* the previously solely right hemisphere belief that the correct answer is blue. So there has been an interaction between right and left hemisphere beliefs. Right and left hemisphere information has been integrated, we might say. And this interaction and this integration occurred via a wholly automatic and unintentional process. Accordingly, Marks says, the interaction was intra-mind: the subject's two hemispheres jointly constitute a single mind.

But the distinction between deliberate and unintentional communication is not sufficiently robust to support a distinction between inter- and intra-mind interaction. Most obviously, my mental states can causally affect my sister's mental states without my wanting them to do so; in fact, my disappointment with the gift she has given me may cause her to believe that she bought the wrong gift, and to feel sad, and to wish that she'd bought something else, and so forth—even if I am trying very hard

not to communicate my disappointment, or to engender in her any regretful beliefs or desires or feelings.

Only slightly less obviously, my own mental states may frequently interact with each other via a conscious, deliberate process, as when I practice deep breathing to calm myself, or play upbeat music to get myself psyched, or (most famously) talk to myself in order to keep myself on task, and so forth. In fact it is possible that the contents of my stream of consciousness are often determined by an intentional (deliberate) process.

So if there is a *principled* distinction between intra- and inter-mind interactions, it cannot be drawn in terms of intentional versus unintentional or automatic communication. (Note that it also, relatedly, can't be drawn in terms of "person-level" versus unconscious communication, for again I may talk to myself, and mental states of which I am unconscious can causally interact with my sister's.)

4.3 Direct versus Indirect Interactions between Mental States

Consider again the intuition that I and my sister do not share a single mind. It is interesting to note how strong this intuition seems *even though* the case would seem to present several significant challenges to that intuition. First of all, my sister and I are psychologically similar in many ways; we remember many of the same events; we have highly similar senses of humor; we share many of the same interests, and many of the same character and personality traits. Thus our minds are "type-similar" if not type identical, to a large extent. If we were conjoined twins they would presumably be even more similar (on the basis of an even greater number of shared

experiences)—and yet the intuition that we would still have two minds seems just as strong.

Secondly, it is not an accident that my sister and I are psychologically similar in many ways, the way my psychological similarity to my twin on twin earth would be accidental. It is not an accident that my sister and I remember having eaten the same foods this Thanksgiving, for instance. We remember having eaten the same things because we ate at the same table: many of our type-same mental states have their causal origins in the same external events. A similar point may hold for our mental *traits* and not just *contents*; if my sister and I have many of the same intellectual strengths and weaknesses, for example, this similarity may be explained in part by the fact that we made our way through the same (primary) school system—or by the fact that we carry similar genes. Thus not just our individual token mental states but our mental capacities may share a causal origin.

Third, again, my sister's mental states and my own frequently causally interact, and have done so for most of my life and all of hers. Thus my sister and I don't just have many of the same beliefs about our parents, and we don't just have many of the same beliefs about our parents because we in fact have the same parents; we also have many of the same beliefs about our parents because we've spent time sharing information about our parents. ("You will never believe what Mom said yesterday. Well, actually, *you* will.")

Yet despite these three challenges that this really very mundane example poses to the intuition that my sister and I have at least two minds between us, the intuition stands. Is there any way to defend this intuition using purely psychological

terms? (I.e., without simply referring to the fact that we are different organisms, or that we have different brains?) Or is it indefensible, at least in those terms?

One possibility is that even though my mind and my sister's mind are type-*similar*, they aren't type-*identical*. My sister tells me she has a crush on someone she calls "cookie boy"; I have never met "cookie boy," and therefore we have many different types of beliefs and emotions and desires towards this "cookie boy." My sister also dislikes horror movies and likes mushrooms, whereas I like horror movies and dislike mushrooms.

This is surely *not* what makes it the case that my sister and I have different minds, however, as is immediately obvious when we consider the fact that I and my twin on twin earth, with whom I am psychologically type-identical (and if you wish physically type-identical as well) do not share a single mind. For minds, like mental states, are *tokens*.

Note that, since I and my twin on twin earth are mentally type-identical, if you knew everything about my twin and had a complete psychological theory for her, you could use this knowledge to predict and explain all of my behaviors, because you would have a complete psychological theory for me also. But this would be sheer accident, for there would be no actual *causal* connection between my twin's staying up late last night and my feeling exhausted this morning. This is why we accept so easily that I and my twin have different minds: there are *no causal connections* between my mental states and hers. Or at least so we assume—but in fact even if some theory in physics said that there were such connections, there would be no causal connections between her mental states and mine that were *psychologically*

relevant to either of us. This lack of (at least psychologically relevant) causal connection can of course be expressed counterfactually: even if my twin *hadn't* woken up exhausted today, I still *would* have added an extra scoop of instant to my coffee this morning.

But again, there are many causal connections between my and my sister's mental states. For example, she likes mushrooms, and it is partly because she likes mushrooms that I believe that she likes them. Yet I and my sister still do not share a single mind, and while there are causal connections between my and my sister's mental states, they do not seem to be of the "intra-mind" sort.

We can begin to express this in terms of counterfactuals by noticing that even if it is true that I would not believe that my sister liked mushrooms *if* she did not in fact like them, I don't believe she likes them *just because* she likes them. My sister's fondness for mushrooms causes her to behave in various ways with regards to them, and *that behavior* causes me to believe that she likes mushrooms. But she could like them without acting as if she did, in which case I wouldn't believe that she liked them. Or she could dislike them and just act as if she liked them, and, my sister being a very good actress, I would believe that she liked them. Even if my sister's fondness for mushrooms *causes* me to believe that she likes them, it does not do so "directly," via a *wholly mental process*.

Contrast this with the way in which I come to believe that *I* do not like mushrooms. If you asked me how I knew that my sister likes mushrooms, I would say, "We've discussed mushrooms, and she's said she liked them. I've definitely seen her eat them. I think she even brined her own once." If you asked me how I knew that

I *don't* like mushrooms, however, I would blink (instead of shrug), and say, “Well—but I just don't like them.”¹⁹ Moreover, I have a good sense (or would have, if I thought about it), of exactly how much I do not like mushrooms; I could tell you, for instance, that they're better than okra, worse than cooked green beans, and about even with cauliflower.

Or contrast the way in which *my* perceptual states interact with (some of) my beliefs with the way in which my perceptual states interact with *your* beliefs. My *perceiving* a glass of water in front of me may be sufficient in some instances to generate in me the *belief* that there is a glass of water in front of me, due to facts about my cognitive architecture. (That it allows the generation or derivation of beliefs from a certain type of visual percept, for instance.) Thus there is what I will call a “*direct*” connection between my perceiving the glass of water in front of me and my believing that there is a glass of water in front of me. Whereas my perceiving a glass of water in front of me is *not* sufficient, in and of itself, to generate in *you* the belief that there is a glass of water in front of me. You will not come to believe this on the basis of my perception *unless* you follow my gaze in the direction of the glass, or

¹⁹ Those who are aware of the debate between introspectionists and interpretationists on the subject of self-knowledge, or knowledge of one's own non-perceptual states anyway, may protest that I am begging the question against the interpretationist by apparently assuming that I have some kind of direct, non-inferential access to my own propositional attitudes and so forth. This debate about self-knowledge is certainly relevant to forming the intra-mind/inter-mind distinction, and I mention it again in a moment, but here I don't think I am begging any questions. Assume that the interpretationist is correct, and my belief that I dislike mushrooms is the product of inferences generated by my mindreading system, just as my belief that my sister likes mushrooms is the product of inferences generated by my mindreading system. And suppose that to yield those inferences my mindreading system has access only to my perceptual states in both instances. The process by which I come to form these two beliefs (about my sister liking and my disliking mushrooms) still may differ because in the one case the formation of those perceptual states causally depends upon (my sister's) overt behavior, while in the other case it does not. Still, the interpretationist's claims do suggest caution in trying to specify too precisely the nature of the intra-mind/inter-mind distinction at this time, a point I return to shortly.

unless I say that there is a glass of water in front of me, etc. Note that this isn't because your mental architecture is any different from mine; it is not because we have different *types* of minds. It is simply because we have different token minds. And while "direct" interactions between mental states occur within each of our minds, "direct" interactions don't occur between them.

The distinction I am now trying to get at can be sharpened by noticing the different sorts of counterfactuals linking (any of) my mental states to (any of) my sister's, on the one hand, and at least some of my mental states to some of my own, on the other. We could begin to at least begin to list the sorts of counterfactuals linking my and my sister's mental states: my sister's taste for mushrooms will (*ceteris paribus*) make her desire some and if some are before her and she is hungry then she will (*ceteris paribus*) eat more than a little of it, and then I, observing her eating these mushrooms in the presence of delicious alternatives will come to believe (*ceteris paribus*) that she likes mushrooms, etc. (I am not suggesting that we could really list them all, and I'm sure there is much we don't understand about inter-mind interaction—though note that a great deal of what we don't understand about *inter-mind* interaction probably stems from ignorance about *intra-mind* processes. But again, we could at least begin to list some of the counterfactuals linking her mental states and my own.) But many of the counterfactuals linking my mental states to my own are comparatively subtle and complex; my mental states, that is, can causally interact with each other in particularly nuanced ways. (Just think of a Joycean stream of consciousness.) Folk psychology may tell us a lot of things about the counterfactuals linking my likes to my own beliefs, for instance—but on the whole, it

still tells us very little. (It took scientific psychology, for instance, to come up with the theory of cognitive dissonance—a theory describing at least *one* way that likes can interact with beliefs.)

While the distinction I am in the midst of drawing between “direct” and “indirect” interactions between mental states is (hopefully) somewhat intuitive, actually specifying the distinction in functional terms is not easy, or even possible at this point, given the uncertainty and the substantive (and great number of) disagreements about how mental states within a single mind *do* interact.²⁰ Someone, for instance, might read the mushrooms example above and eagerly add that of course I have *non-inferential* access to my *own* beliefs and desires concerning mushrooms, whereas any beliefs I form about my *sister’s* mushroom-concerning beliefs and desires will be formed only by an inferential process, presumably of my mindreading system. But in fact some believe (see Carruthers 2009 for a defense) that my beliefs about either my own dislike or my sister’s like of mushrooms will be equally formed by inferential processes of my mindreading system; and others, of course, quite strenuously disagree. And this is just *one* debate in contemporary psychology, concerning, in part, the kinds of interactions that do and don’t exist between various

²⁰ Of course, it isn’t difficult to specify in functional terms just this distinction, between intra-mind and inter-mind interaction! Specifying the functional distinctions between all sorts of psychological entities is immensely difficult. In fact the great difficulty of this task once drove Davidson (1973/1980) to despair that it couldn’t be done; he describes an anxious rock climber suddenly struck by the desire to cease to be imperiled by the weight of his partner, climbing up below him, and who knows that he could rid himself of this danger simply by relaxing his grip on his partner’s rope. And this desire, and this belief, unnerves him so much that the rope actually does slip from his hold in a sudden spasm (leaving the other climber to plummet to the earth presumably). The action seems caused by a conscious belief that letting go would relieve him of danger and a desire to be relieved of this danger—and yet in the example the first climber never wanted or chose for his partner to fall or for his own grip on his partner’s rope to relax. Davidson illustrates, in other words, the great difficulty of spelling out the criteria for intentional (deliberate) action—and by extension other psychological kinds. While conceding the difficulty of the task Rey (1997) proposes that it is one that a developed psychology will be up to (just as we expect that a developed biomedical science will be able to distinguish between different forms of cancer for instance), and I make this same assumption here.

sorts of mental states within a mind. Given this state of uncertainty about the kinds of *intra*-mind interactions that exist, it may be premature to attempt to formulate a principled distinction between intra- and inter-mind interactions.

Perhaps the wisest strategy to take at this point is the Gricean (1965) one of distinction by exemplar, followed by *deference to experts*. Grice, recall, was trying to offer a causal theory of perception: a theory of perception according to which a causal relationship between perceiver and thing perceived was in fact essential to the mental act in question constituting an act of actual perception. But when specifying the *right* causal relationship, the *right* functional role of perception, proved beyond his grasp, he proposed simply referring to prototypical *examples* of perceiving—and then agreeing to let the “specialist” explain, now or much later, what distinguishes perception from, say, hallucination:

I suggest that the best procedure for the Causal Theorist is to indicate the causal connection by examples: to say that, for an object to be perceived by X, it is sufficient that it should be causally involved in the generation of some sense-impression by X in the kind of way in which, for example, when I look at my hand in a good light, my hand is causally responsible for its looking to me as if there were a hand before me, or in which. . . . (and so on), *whatever that kind of way may be*; and to be enlightened on that question one must have recourse to a specialist. (Grice 1965: 463, original emphasis)

Similarly, perhaps the wisest move for me at this juncture would be to point to what we have some confidence are intra-mind interactions—such as a hemispherectomized subject’s decision to leave the house promptly on the basis of coming to believe that she can still make her train if she hurries—and then point to what we have confidence are inter-mind interactions—such as my coming to believe

that “cookie boy” is a jerk on the (partial) basis of my sister’s recent decision that he is a jerk—and then wait for the specialist to illuminate the different psychological mechanisms at play in the former that do not appear in the latter. *Then* we can call those psychological mechanisms that do come into play in the former case “direct,” and the others “indirect.”

I agree with these calls for conservatism to a large extent. Any distinction I could draw at this point would necessarily be not just tentative but quite rough, given the current lack of certainty, again, about the kinds of ways that mental states interact within a mind. The best distinction will be drawn out of a complete psychological theory that characterizes adequately and in detail the ways in which mental states interact within a mind. Note, for instance, that there will be a number of different kinds of “direct” interactions; beliefs might interact “directly” with other beliefs in a different way from the still “direct” way in which they interact with desires, or with perceptual states, and so forth. This will have consequences for the individuation of minds: whether two mental states that have interacted are co-mental or not will depend not just upon the form their interaction took (or could have taken), but also upon the *kinds* of mental states that they are, and the particular manner in which a *complete psychological theory* determines that those particular kinds of mental states should be able to interact. This point is another plea for caution.

Nonetheless I think that we can at least *begin* to say *something*, even if nothing very precise, about the difference between *some* “direct” and “indirect” interactions between mental states. (Perhaps what I will say here will be entirely wrong, and the experts will correct me later; hopefully, however, what I say will

simply be not entirely right... and quite incomplete.) For, assuming that we can currently choose some exemplars of intra-mind and inter-mind interactions, then we might be able to find *some* differences between *some* of these exemplars. And even if the distinction we draw now is ultimately untenable or refuted or insufficient, still, drawing it might help to understand the appeal that the mental duality position has had so far.

As a first, tentative attempt, I propose to define “direct” interactions negatively, in terms of “indirect” interactions between mental states. Two mental states interact with each other *indirectly* if the chain of events causally connecting them is composed partly of behavioral events, environmental events, and perhaps other nervous system events not categorizable as mental. (Events not carrying intentional content, perhaps.) Indirect interactions between mental events, then, are interactions that supervene on events on which no mental events supervene. And *direct* interactions between mental events are interactions that are not indirect.

Even if this distinction between “direct” and “indirect” interactions captures some important difference between the way my mental states interact with each other and the way they interact with my sister’s mental states, we have to be cautious in trying to link this distinction to the intra-mind/inter-mind distinction. We have to be cautious, in other words, in trying to use this distinction between direct and indirect interactions to individuate minds. For instance, the direct/indirect distinction clearly does not map perfectly onto the intra-mind/inter-mind distinction: it is not the case that all intra-mind interactions are direct, and all inter-mind interactions are indirect.

This bi-conditional cannot be correct because many of the interactions between my own mental states are not direct, as defined above. To take the most mundane sort of example, one of the reasons I *believe* I am sitting down right now is because I *am* sitting down right now. . . And I'm sitting down right now because I *decided* to sit down several minutes ago. Thus there is an “indirect” causal relationship between that decision and my current belief that I'm sitting. *This* kind of “indirect causal relationship” between one's own mental states is obviously incredibly common. (Of course, we're not assuming that I have a single mind, but the problem with saying that intra-mind interactions are always direct is that the mental states of a single mind will no doubt have to interact with each other indirectly. It seems very *unlikely*, to say the least, that there might be a creature—at least one who lasts very long in this world—for whom *all* its mental states are directly causally connected! For one thing, any behavior such a creature engaged in could play no role in generating further mental states.)

Nonetheless the distinction just drawn between “direct” and “indirect” interactions, while not equivalent to a distinction between intra- and inter-mind interactions, may still be useful in drawing that latter distinction. For even if my mental states frequently interact with each other indirectly, just as they frequently interact with my sister's mental states indirectly, it is still the case that *some* of my mental states interact with each other directly, whereas *none* of them interact with my sister's mental states directly. So perhaps we can instead say that if there are two sets of mental states, between which *no* direct interactions are possible, then those two sets of mental states belong to (at least) two minds, and that if there are two sets of

mental states between which *any* direct interactions are possible, then (if all the mental states within each set belong to a single mind) the two sets of mental states belong to one and the same mind. This seems to offer more secure grounds for making an intra-/inter-mind distinction. In particular, it draws the distinction on causal and counterfactual grounds.

But arguably this proposal should be modified further, in two ways. First, I said just above that if there are two sets of mental states, between which *no* direct interactions are (nomologically) possible, then those two sets of mental states belong to (at least) two minds. It might be better, however, to say that the two sets of mental states *do not* belong to *one* mind—for perhaps one set of mental states belongs to no mind at all. (Some may believe that there can be mental states without whole minds.) Second, I said just above that if there are two sets of mental states between which any direct interactions are possible, then, so long as all the mental states within each set belong to a single mind, the two sets of mental states belong to one and the same mind. But this might be too strong; perhaps a pair of conjoined twins, joined partly at the brain, might share *some* mental tokens, but still have two minds. Instead of making a binary distinction between *no* direct interaction and *any* direct interaction, the former signifying multiple minds and the latter signifying a single mind, to count minds, we might instead perhaps look at relative *degrees* of direct and indirect interaction. Between the mental states of a single mind there should be a rich web of direct interactions, while between multiple minds there should not be such a rich web of direct interactions—even if there are, perhaps, *some* direct interactions.

Note again that the direct/indirect distinction I've drawn here is very general; it gives only the barest sketch of the difference between the ways in which one of my likes can interact with one of my beliefs, and the ways in which one of my sister's likes can interact with one of my sister's beliefs. For it says nothing about the ways in which my likes and beliefs can interact and the mechanisms that allow this *other* than that they can interact in a way that my sister's likes can't interact with my beliefs! For any degree of detail, then, we still need to defer to the experts.

This distinction between direct and indirect mental state interactions, and the distinction between intra-mind and inter-mind interactions that we developed from it, can easily be integrated into either a purely psychological or a partly neural method of individuating minds. Those criteria both allow that whether multiple mental systems jointly constitute a single mind or separately constitute distinct minds depends upon the kind and degree of interaction possible between them. And the functional distinctions (direct vs. indirect, intra- versus inter-mind) developed in this section begin to fill out, just a little, the kind of interaction that matters.

Is it possible that we could go somewhat further towards refining the distinction between intra-mind and inter-mind interactions? The examples of direct and indirect interactions that I've discussed so far refer to interactions between likes and beliefs, or percepts and beliefs. Given the account of minds as information-integrating systems adopted earlier, however, we might want to place special emphasis on the necessity of direct interactions between informational states and informational and directional states, specifically. And indeed the distinction between direct and indirect interactions in and of itself might recommend a similarly tightened

focus on interactions between informational and directional states specifically. For the prototypical cases of *indirect* mental state interaction (or perhaps just inter-mind indirect interaction?) seem to involve *perception*. And it is meanwhile part of the functional role of informational and directional states, meanwhile, that they can interact with each other *absent* the mediation of any perceptual state.

I am hesitant to try to tighten or refine the distinction too much, however—hesitant to restrict the types of states to which it is best applied—because it is safer to commit myself less rather than more. So while I will note, where relevant, in the next chapter, the kinds of mental states involved in potentially direct inter-hemispheric interactions in the split-brain subject, in general I will simply speak about direct and indirect interactions between mental states, and not confine myself to talk of direct and indirect interaction between informational and directional states only.

5 Conclusion

This chapter was about the individuation criteria for minds. I have argued that a mind is a complex system with a particular mental architecture: one that allows informational and directional states to interact to yield reasoning and decision-making, and one that engages in learning and perception. I have also begun to explain and defend the role I think neuroanatomical facts can play in the individuation of mental entities, suggesting the plausibility of a partly neural account of minds, and of using neural facts to individuate minds. In Chapter Three I will argue that split-brain

subjects do indeed have two minds, and I do this in part on the basis of the distinction between intra-mind and inter-mind interactions developed in Section Four here.

But as I have cautioned several times already, that distinction was drawn only tentatively, and so far it is just a sketch. Ultimately the best distinction between intra-mind and inter-mind interaction will come from a developed psychological theory, one that provides laws or principles or models concerning the ways in which mental states of various types can and can't interact with each other within a mind.

But it is at least *conceivable* that there won't be any robust or principled distinction to be drawn. I can only hope to have made plausible that not just substances and organisms but *causal processes* can come in natural kinds, too, and that therefore psychological theories may ultimately vindicate the intuition that there is some deep psychological difference between all the ways my mental states can interact with my sister's and *some* of the ways my mental states can interact with themselves. And in the next chapter, which focuses on individuating minds in split-brain subjects in particular, I hope in describing the split-brain phenomenon to provide some indirect evidence that there is some important, deep psychological distinction between the intra-hemispheric and inter-hemispheric causal processes in those subjects.

Chapter 3: A Defense of the Mental Duality Model

1 Introduction

This chapter is organized around a defense of the *mental duality model* for split-brain subjects. In the next section I explain why both a purely psychological and a partly neural set of criteria for individuating minds and mental tokens will tend to yield the same answer to the “how many minds?” question, for any creature, and why they do so in this instance. The explanation is simply that there is an isomorphism between neural structure (neuroanatomy) and at least one element of cognitive architecture: mental connectedness, or integration. Section Two also describes Hurley’s objection to the “isomorphism thesis” and shows that her objection fails at least so long as we accept a general mind-brain supervenience claim. By the conclusion of that section, the presumption in favor of a “two minds” claim for split-brain subjects will be clear.

Most of the rest of the chapter defends the mental duality model against what I take to be the three major objections to the “two minds” conclusion. The first objection, dealt with in Section Three, is the argument from unified behavior. That argument says that split-brain subjects surely do not have two minds, since their behavior is generally so integrated—as integrated as is yours or mine. I argue that the integrated nature of split-brain subjects’ behavior provides compelling, but defeasible, evidence for the existence of a single mind.

A second objection to the “two minds” conclusion, presented in Section Four of this chapter, says that the two hemispheres jointly realize a single mind in virtue of

bearing similar psychological properties. I call this the “singularity-through-redundancy” (STR) model. While I believe that this model founders for split-brain subjects on purely empirical grounds, I mainly submit it to theoretical scrutiny. The “singularity-through-redundancy” position says that when a partly neural and purely psychological criteria for mental individuation *do* yield different answers to the “how many minds?” question, the best answer is that yielded by the purely psychological criteria. The view, that is, rejects facts about causal interaction and independence as criteria for individuating minds in some instances. Yet one central claim of this chapter and this work is that physical properties play an essential role in the individuation of mental tokens. As I note in Section Two, this commitment to the role of physical properties in individuating minds, at least in the form it takes in this work, bears strong resemblance to the “isomorphism” claim rejected by Hurley.

After discussing the STR model and before discussing a last objection to the mental duality model, I pause, in Section Five, to consider what might be motivating some philosophers’ unwillingness to accept the mental duality model. One possibility is that these philosophers have fallen prey to verificationism about psychological explanation. But another possibility is that they cannot see how to integrate the mental duality model for split-brain subjects with our understanding of those subjects in important, non-scientific contexts. I suggest, however, that the mental duality model is more compatible with our social, legal, moral, personal, and pragmatic ways of relating to and understanding split-brain subjects than one might suspect.

The last objection to the mental duality model for split-brain subjects, discussed in the last major section of this chapter, Section Six, is the most

compelling. Even though the two hemispheres of a split-brain subject are divided at the cortical level, they remain connected via their mutual connectedness to various intact, non-cortical brain structures, and thus causally linked. In light of this, both the purely psychological approaches to individuating minds that I've advocated *and* the commitment to isomorphism between mental and neural architecture may threaten to make the "two minds" position untenable with respect to split-brain subjects.

Furthermore, the commitment to isomorphism threatens to make the "how many minds?" debate *merely verbal*. For if the isomorphism view is correct, then just as a body of water might be intermediate between being one lake and two lakes, a mental system might be intermediate between being one mind and two minds. Once we acknowledge that, why should we think there is any non-arbitrary way of deciding whether split-brain subjects have one mind or two?

While conceding that the split-brain cases show that two minds are not always wholly distinct in the sense of distinctness relevant to individuating minds, I nonetheless maintain that a split-brain subject does have two minds. The interhemispheric interaction afforded by non-cortical structures in split-brain subjects does not serve to *integrate* the mental processing of the two hemispheres but rather to *coordinate* their perceptual inputs and motor outputs. While non-cortical structures engage in mental operations, these operations are incorporated into cortical ones in such a fashion as to mean that a split-brain subject still has two minds, each associated with one hemisphere *and* various non-cortical structures.

The nervous system consists of cells engaged in rapidly and continually receiving, transforming or operating over, and communicating information to other cells.

Understanding what the brain does is largely about understanding both the division of labor in the brain and how the fruits of this divided labor are integrated to yield sophisticated cognition, experience, and action. The nervous system, in other words, is about division and interaction. The “how many minds?” question, a question about mental architecture, is about this same thing, whether we approach the question armed with purely functional/psychological facts or with neural facts as well.

The answer to the “how many minds?” question, at least for split-brain subjects, but presumably in most other cases also, *may* not depend upon whether neural facts are or are not used to individuate minds. All other things being equal, purely psychological and the partly neural criteria for individuating minds should tend to yield the same answer to the question. For both sets of criteria (properly formulated) are fundamentally about the same thing: they both cast the question as one about the causal relationships holding between various types of mental entities. While the purely psychological approach individuating minds that I sketched in Chapter Two, involving the distinction between “direct” and “indirect” interactions, may seem able to ignore facts about neural architecture, neural stuff sustains the interactions that the approach identifies as “intra-mind”; without that stuff between two mental systems, only inter-mind interactions are possible. (At least, for the species we know of. This would not be true of creatures that engaged in telepathy, for instance.)

While the partly neural approach to individuating minds may appear to look only for the existence of direct physical interaction between two neural systems, which interactions *qualify* as direct is in fact a psychological matter. That is, the partly neural criteria for individuating minds will tend to yield the same answer to the “how many minds?” question for an almost trivial reason: the brain structures whose presence or absence it uses in determining whether two mental systems are identified with one and the same or with distinct minds is given a psychological taxonomy: it is itself identified as playing this role because of beliefs about the psychological properties with which it is associated. Certainly the partly neural approach to individuating minds yields the same answer to the “how many minds?” question *in this instance*, where the neural story and the (correct) psychological story seem so neatly to cohere. The relationship between severing the corpus callosum on the one hand and mental dissociation on the other is easy to understand, at least in broad strokes.

That the partly neural and purely psychological approaches to individuating minds converge on a response to the “how many minds?” question supports, and is supported by, what Hurley (1998) has called the “isomorphism” thesis. This is the claim that mental structure is isomorphic in some sense to neuroanatomical structure. Hurley discusses the appeal to isomorphism “with objective neuroanatomical structure to resolve the indeterminacy” between competing views on the structure of split-brain subjects’ consciousness—a topic that will not be my focus until later chapters. But of course isomorphism to neuroanatomical structure could be appealed to in order to resolve the “how many minds?” question as well. If mental architecture is isomorphic

to neural structure, then neuroanatomical structure can be appealed to in order to at least help determine mental structure. And indeed I believe that “split-brain” neuroanatomy does tell us something about “split-brain” cognitive architecture: the fact that split-brain subjects have two cortically divided hemispheres provides some reason to think that such subjects have two distinct minds. Thus I will say something in defense of a version of the isomorphism thesis—a thesis Hurley rejects—here, since my arguments throughout the rest of the chapter all concern the isomorphism thesis, implicitly or explicitly.

First I want to say the kind of isomorphism between neuroanatomical and cognitive structure that I think is plausible and that I will be defending, for it would clearly be easy to formulate an isomorphism claim that was too strong. The kind of isomorphism that I think is plausible is just this: functional relations between individual mental states or between mental systems—i.e. between mental tokens—are underlain by physical connections between those tokens. Of course the physical route connecting one to the other need not be direct (in an intuitive notion of direct). Then again, the functional relation need not be direct either. The activities of two mental subsystems could be coordinated with each other via a third subsystem. For that matter, two mental systems could also interact not with each other per se but rather with some further system that then coordinated their outputs. The first two systems could thus be functionally and causally connected, despite lack of a direct physical route connecting them—but even the functional and causal route between them would then be indirect insofar as it was mediated by some third system. That two tokens interacted with each other only via some third mental token would of course not in

and of itself mean that those tokens didn't belong to or weren't a part of the same mind. We would need to know more about the form their interaction took.

So all I mean when I speak of isomorphism between physical and psychological structure is this. First, functional relationships between various mental tokens (systems, subsystems, or individual mental states) supervene on physical relationships between those tokens. Second, facts about the physical routes between mental tokens tell us something about the functional relationship between those tokens. But what those facts tell us depends upon the psychological taxonomy of the physical routes in question. So at some level functional structure and neural structure must be isomorphic because the psychological is neural, and (some of) the neural is psychological. There will therefore be some level of description of the neural events implementing an accessibility relation, for instance, that makes this isomorphism clear. (See Revonsuo, 2000, for a similar discussion.) Granted, the level of description may be very abstract, so the isomorphism itself may be very abstract. For the most part we're just in the dark at this point about what the right level of description is or could be. But the kind of isomorphism that would support the mental duality model for split-brain subjects is intuitively easy to understand. Mental integration is achieved via causal interactions; causal interactions are physical interactions; the physical interactions in question are neural (or so I will eventually argue). So, mental connectedness is isomorphic to physical connectedness.

Again, this is not to say that the fact that split-brain subjects lack a corpus callosum *in and of itself* provides evidence that those subjects have two minds apiece. First the corpus callosum must be given a psychological (functional) taxonomy. So it

is the fact that split-brain subjects lack a corpus callosum in combination with the claim or hypothesis that the corpus callosum serves as an important means of interhemispheric interaction (and in combination with model of the hemispheres themselves as cognitive systems of course) that provides evidence that split-brain subjects have two minds apiece. This kind of isomorphism is in a way circular, then, though hopefully not viciously so, and so stated hopefully not absurdly strong, either. (In fact it arguably isn't very dramatic.)

But Hurley (1998) believes that dissociative identity disorder (DID) and agenesis of the corpus callosum (ACC) show that there is *no* necessary isomorphism mental and neural structure, or between conscious and neural structure in particular. People with DID have multiple streams of consciousness, she says, yet a single physically unified brain. People with ACC meanwhile have two physically disconnected cerebral hemispheres and yet a single stream of consciousness. I won't talk about DID here, partly for space constraints, and partly because DID remains a controversial diagnosis. But something should be said about agenesis of the corpus callosum, for individuals with this condition (ACC) are understandably seen as the "natural" version of split-brain subjects: individuals with ACC are born without a corpus callosum (or with only a partial corpus callosum). The condition is not rare; it has a prevalence of about .3% in the general population, and a prevalence of between 2-3% among developmentally disabled individuals. It is associated with other central nervous system abnormalities in about 80% of cases (Geoffroy, 1994), but some individuals with ACC seem otherwise "normal" and have IQs in the "normal" range.

Hurley is far from alone in thinking that individuals with ACC (“acallosal” subjects) each have a single stream of consciousness. Partly she believes this on behavioral grounds; acallosal subjects behave more like “normal” subjects than they do like split-brain subjects, she says. Thus while we have compelling behavioral evidence that split-brain subjects have a dual consciousness, there is no analogously compelling evidence that acallosal subjects do:

When the corpus callosum is severed or absent within one body, we may have either a commissurotomy patient, who seems to support separate centers of consciousness, or a callosal agenesis patient (someone born without a corpus callosum). Callosal agenesis patients, or acallosals, typically pass almost all the experimental tests of unity that commissurotomy patients fail, including under conditions involving fixation. Their actions argue for a unified consciousness, even in experimental conditions and despite their similarity in gross neuroanatomical structure to commissurotomy patients. (Hurley 1998: 189)

While acallosal subjects do exhibit far fewer *dissociation* effects than do split-brain subjects, however, they do not behave precisely like “normal” subjects. It is again hard to generalize here, since approximately eighty percent of such subjects have other brain and developmental abnormalities, but it appears that the lack of a corpus callosum alone does have some behavioral influences. At “cross-matching” tasks (a blindfolded subject has an object placed in her left hand, and now must select the same object using her right hand, for example), acallosal subjects are far superior to split-brain subjects, yet their performance time is significantly slower than that of IQ-matched controls (Vanasse, Forest, Lassonde, 1994), suggesting that they may use different neural and perhaps different cognitive means of performing such tasks. The authors suggest that such subjects may be relying upon “information provided by both

ipsi- and contralateral somesthetic pathways. . . [performing] a bimanual matching task using an *intra*-, rather than an *inter*-, hemispheric comparison process” (ibid., 199), and that their slower response times result from the fact that nerve transmission is slower in the ipsilateral lemniscal and spinothalamic pathways that would permit intra-hemispheric performance of the cross-matching problem.

Subjects with ACC exhibit somewhat unusual fine motor behavior. For example, by the time a “normal” subject’s fingers have reached a coffee cup, her fingers have already narrowed to the width appropriate for that object; in acallosal subjects, “the hand opens but remains in a gaping posture until it hits the object and only then do the fingers begin to close in an appropriate configuration on the object” (Silver, Jeeves, 1994: 216). These authors found that bimanual coordination “can be maintained by using visual and proprioceptive feedback, as long as performance is slow” (ibid 217). Performance breaks down at high speeds, however, and “Unlike normals, the acallosal subjects do not, with practice, become independent of visual information” (ibid). Like split-brain subjects, they exhibit alexithymia (O’Brien, 1994). They appear to experience some kind of (not yet well-understood) impairments at visual and visual memory tasks (Temple, Ilesley, 1994). (It is possible that due to the absence of a corpus callosum that would allow the hemispheres to both benefit from their developing separate specializations, the right hemisphere as well as the left becomes specialized for verbal processing (see Sperry, Gazzaniga, and Bogen, 1969), but I think that this hypothesis has waned in popularity (See Schmidt, 1994).) Interhemispheric transfer of visuo-motor learning is significantly impaired (Lassonde, 1994).

Nevertheless the claim that acallosal subjects have a single stream of consciousness is not implausible. Would the correctness of this claim require rejecting the isomorphism thesis—the thesis that the structure of mind (or, in this case, the structure of consciousness) is isomorphic to neuroanatomical structure, in certain respects or to some degree? It might appear so, for acallosal subjects have, relative to “normal” subjects, a *divided* brain; if both groups of subject nonetheless have a *unified* consciousness, then the structure of consciousness clearly can’t be isomorphic to neuroanatomical structure. And assuming this point generalizes—assuming that mental structure in general isn’t isomorphic to neural structure—then neuroanatomical facts about *split-brain* subjects also cannot be used to help determine how many minds or streams of consciousness they have. (Though Hurley in fact questions the last move in this argument, as we will see below.)

It is too quick, however, to say that acallosal subjects have a divided brain in the sense necessary to reject the isomorphism thesis. For on at least one understanding, the isomorphism thesis essentially just says that the the structure of consciousness is isomorphic to the neural structure that *supports* the structure of consciousness. (Other understandings of the isomorphism thesis are possible—but some of these theses are implausibly strong. The isomorphism thesis as *I* am defending it, is pretty modest: it just says that mental connectedness—co-mentality or co-consciousness—requires physical connectedness, because that’s what mental connectedness supervenes on. In fact, co-consciousness, for instance, consists in a certain type of physical connectedness. *Which* type of physical connectedness? Whatever type sustains co-consciousness.

While the isomorphism thesis, so understood, may not be tautological, it is not particularly dramatic, and there is certainly a kind of circularity involved in testing it. This kind of circularity is essentially just the bi-directional relationship between the individuation and classification of psychological and neural entities. To see whether a creature, C's, consciousness is isomorphic to his neuroanatomy, you have to find the *relevant* portion of C's neuroanatomy. (Suppose C has a single stream of consciousness—and two, disconnected pineal glands. This couldn't weigh against the isomorphism thesis if the mechanism for co-consciousness in C had nothing to do with C's pineal glands.) But *which* portion of his neuroanatomy is relevant to this investigation is a *psychological* matter. Thus the investigation into the structure of C's consciousness and the structure of C's brain, for the purpose of testing the isomorphism thesis (as it applies to consciousness), can only be undertaken together. (This of course is just a general fact about investigating cognitive and neural architecture.)

Since it is not clear what neural structures support consciousness and co-consciousness—in any of us, to some extent, and in acallosal subjects, in particular—the finding that acallosal subjects do have a single stream of consciousness would not, in and of itself, weigh against the isomorphism thesis. For while the two hemispheres of an acallosal subject are (largely) divided at the cortical level, they are not divided at a “lower” level.²¹ It is only in the context of a commitment to the cortex's involvement in conscious experience and to the corpus callosum as a mechanism of unifying consciousness that an acallosal subject's brain would be expected to support

²¹ In addition, they are cortically connected via the anterior commissure.

two streams of consciousness in the first place. But how much do we know about the *functional* neuroanatomy of acallosal subjects?

There is at least a debate about how acallosal subjects' great degree of behavioral integration is achieved. Three positions in this debate are most prominent. The first is that some non-cortical structures of an acallosal subject may be more developed and/or may serve additional functions that they do not serve in a "normal" (or a split-brain) subject. The second is that ipsilateral sensory and motor pathways may be more developed in acallosal than in "normal" (or in split-brain) subjects, resulting in each hemisphere having increased sensory and motor powers, and therefore each capable of perceiving and doing *more*, and of acting independently of the other to a greater and more sophisticated degree, than is true in split-brain subjects and in "normal" subjects in particular. A third possibility is that an acallosal subject exploits certain kinds of behavioral strategies to compensate for the lack of a callosum. (Though these strategies would have to be very subtle, since acallosal subjects still exhibit fewer disconnection effects than split-brains even in experimental situations in which care is taken to prevent the use of such unifying behavioral strategies.) (For some time a fourth prominent hypothesis was that acallosal subjects' hemispheres might enjoy a greater degree of bilateral representation of *cognitive* functions than in "normal" subjects do, but it instead appears that acallosal subjects' hemispheres are normally lateralized on an individual basis. See Schmidt, 1994.)

Based on our discussion up until this point, it should be clear that the first hypothesis is of course compatible with the claims that a *split-brain subject*

(particularly a newly operated one) has two streams of consciousness, that an acallosal subject has one stream of consciousness, and with the isomorphism thesis. Conscious unity would have a different neuroanatomical basis in an acallosal than in a “normal” subject—but for both subjects, consciousness would be isomorphic to brain. A similar point may hold for the second possibility, though only *if* ascending and descending motor and sensory pathways constitute a means of unifying consciousness. If they do not, however, this second hypothesis can support a conscious singularity claim for acallosal subjects only upon rejection of the isomorphism thesis—at least in one special case, to be discussed in the next section of this chapter: the case of singularity-through-redundancy.

The possibility that behavioral integration in acallosal subjects is achieved via behavioral means probably seems intuitively less compatible with the claim that they have single streams of consciousness. Hurley, however, disagrees; she in fact believes that *how* acallosal subjects’ behavioral integration is achieved is irrelevant to determining the structure of their consciousness. Even if acallosal subjects depended largely on what she calls “external” or “extended” means of behavioral integration—cross-cuing for example—those means could still serve to unify their consciousness. Given the functional importance of conscious unity, she continues, it is plausible to believe that we all develop in such a manner as to develop a unified consciousness, whatever our neuroanatomy (at least within reason—no doubt she would bar hydrocephaly, etc.). (Gazzaniga and LeDoux make a similar remark at one point, in the context of discussing some of the psychological, neural, and behavioral mechanisms split-brain subjects ultimately develop to integrate their behavior: “It is

as if the brain demands integration, and in the absence of interhemispheric pathways, less efficient ways of achieving mental unity are employed” (1978: 39).) This makes a conscious disunity claim plausible for at least a recently callosotomized subject, since the means and mechanisms her consciousness *had* relied upon to become unified are now altered.

Suppose a recently-operated commissurotomy patient uses a partly external mechanism of integration. For example, if access movements are prevented, a smile appears to signals that ‘yes’ is the right answer to a question. This would naturally be taken as evidence of disunity of consciousness and an attempt to communicate information between two separate centers of consciousness. The fact that the mechanism of integration is partly external here appears to have implications for the structure of consciousness. (Hurley, 2003: 78)

But if consciousness *always* tends to develop in such a manner so as to become unified, a disunity claim for acallosal subjects upon similar grounds has no support. In fact such a claim never has any support, again no matter *what* we discover about the *means* by which acallosal behavioral integration is achieved, even if it is achieved by, say, cross-cuing or other external or extended means:

Suppose the acallosal has always depended in everyday life primarily on extended rather than purely internal mechanisms of integration. This involves subtle forms of cross-cuing and access movements that are not properly described at the personal level as rationally controlled, intentional actions. Rather, they function smoothly and automatically at a subpersonal level to integrate information. It was never the case that the acallosal harbored separate centers of consciousness and agency, somehow formed, structured, and unified independently of the partly external mechanisms. It was never the case that such prior separate units of consciousness hit on the use of external paths as a way of communicating between themselves. This person’s consciousness, including its structure, emerged and developed with partly external mechanisms in play. (Hurley, 2003: 78-79)

And if we accept Hurley's arguments, here, we reject the isomorphism thesis; and if we reject the isomorphism thesis, then perhaps the structure of *split-brain* subjects' neuroanatomy cannot be appealed to in order to resolve how many minds or streams of consciousness *they* have, either.

Of course, the fact that this acallosal subject's consciousness, including its structure, emerged and developed with partly external mechanisms "in play" does not entail that its structure emerged and developed as a single stream of consciousness. This subject might well harbor separate (in the relevant sense) centers of consciousness, perhaps formed and shaped by partly external mechanisms. These mechanisms might even allow the two hemispheres to communicate quickly and with great skill, while not sufficing to create a single subject of experience that can simultaneously introspect the contents any two currently conscious experiences.²²

I don't mean to say that development can be ignored when we try to do psychology, e.g. to determine the structure of a subject's consciousness. But Hurley's emphasis on the ways development shapes consciousness comes at the expense of de-emphasizing the constraints the neural poses on conscious structure. Of course, this de-emphasis is deliberate on her part: Hurley rejects mind-brain supervenience: she believes there are mental phenomena (in human subjects, for instance) that are not neural. If the mind extends beyond the brain then of course "split-brain" neuroanatomy will not necessarily signify anything about split-brain subjects' mental

²² Whether the subject would develop two separate centers of *agency* is, I think, a harder question. Given the links between agency and acting on the one hand and acting and embodiment on the other, it seems plausible that the hemispheres would develop in such a fashion as to *in effect* cooperate with each other a great deal, even if neither hemisphere conceived of itself as cooperating with a second mind.

architectures. If cross-cuing, for example, or even features of the environment, can be a means of mental unification, then the split-brain experiment may well change the mental architecture of split-brain subjects, though their *neural* architecture remains identical in and outside of experimental conditions. In fact Hurley's rejection of the isomorphism thesis ultimately just reduces to a rejection of mind-brain supervenience.

This debate about the extent of the mind is highly relevant to current concerns. In fact a defense of mind-brain supervenience is thematically linked to many of the fundamental claims and positions in this work. I put off a defense of mind-brain supervenience until Chapter Six, however, and for the moment, I will simply assume the truth of some kind of mind-brain supervenience claim.

With that claim in place, acallosal subjects present no reason to think that split-brain subjects have single minds. Indeed, contra Hurley, acallosal subjects could have single minds, and split-brain subjects could have two minds, and the isomorphism thesis could nonetheless *still* be correct: the different mental structures (one mind vs. two minds) such subjects possessed could be isomorphic to the in fact different neural structures each subject possessed. For the neural architecture of split-brain subjects and "natural" split-brain subjects—subjects simply *born* without a corpus callosum—are unlikely to actually be the same. Having developed in the absence of a corpus callosum from the time of an acallosal subject's birth, other neural structures—the anterior commissure, the intertectal commissure—may have developed properties, from the morphological to the functional, that these structures lack in "normal" and in split-brain subjects. So the isomorphism thesis stands, and

both supports and is supported by the fact that the partly neural and the purely psychological criteria for individuating minds will both tend to yield the same answer to the “how many minds?” question.

The best model of split-brain subjects’ mental architecture is, I submit, the mental duality model. That is, a split-brain subject has two minds, one more or less associated with her left hemisphere and one with her right. While the two hemispheres to some extent share a functional foundation, due to their mutual connectedness to non-cortical structures, each hemisphere builds on top of and out of this minimal foundation a rich and unique cognitive and experiential structure—one that meets the criteria for mindhood independently of the rich and unique structure associated with the other hemisphere.

This is a conclusion we could reach using either a purely psychological or a partly neural set of criteria for individuating minds and mental tokens. Given the role the corpus callosum plays in interhemispheric mental integration—a role evidenced by the results of the split-brain studies themselves—the absence of a corpus callosum in a split-brain subject has not just neuroscientific but psychological significance. And the causal, physical independence of left and right hemisphere mental activities has this same sort of significance. Within each hemisphere (and associated non-cortical structures) of a split-brain subject there is a rich web of “direct” interactions, in the sense defined in Chapter Two, and yet between the two hemispheres interaction is predominantly “indirect.” In fact I will argue in Section Six of this chapter that the interhemispheric interaction afforded by non-cortical structures does not serve to *integrate* the mental processing of the two hemispheres but rather to *coordinate* their

processing to a degree, by producing in them some of the same perceptual types and contents, and by receiving and coordinating their motor outputs.

That the two approaches to individuating minds should yield the same result in the split-brain case, however, may seem puzzling, partly because philosophers of mind and neuropsychologists have often tended to reach different answers to the “how many minds?” question for split-brain subjects, and because the former group has been more drawn to a purely psychological and the latter group to a partly neural set of criteria. I do believe that there is one view about the relationship between mental and neural tokens—one that neuropsychologists are less likely to adopt than philosophers—which may partly be responsible for this between-group difference. According to this view, which I call the “singularity-through-redundancy” position, there is at least one exceptional instance in which the isomorphism thesis *isn't* correct. Since this view both draws upon some important functionalist principles, and yet nonetheless should be rejected, it is worth considering it in some detail, as I do in Section Four.

First, however, I want to say something about the role of *behavior* in psychological theorizing and explanation. Those who have rejected the mental duality model for split-brain subjects have done so on the basis of the integrated nature of a split-brain subject’s day-to-day behavior. So what role *does* that subject’s behavior—inside and outside of experimental conditions—play in supporting a model of her mind—or minds? I argue that the role is somewhat more complex than critics of the mental duality model have sometimes suggested.

This section deals with the first major objection to a “two minds” conclusion for split-brain subjects: the argument from unified behavior. Very simply the objection states if split-brain subjects had two minds then they would behave in a disunified manner even outside of experimental conditions. Since they instead behave in a generally integrated fashion, they must have a single mind.

One of the important features of the debate about how many minds split-brain subjects have concerns differing views on the relationship between behavior and psychological explanations of that behavior. In advancing a “one mind” conclusion for split-brain subjects, several philosophers have operated under an overly simplistic conception of this relationship.

Those who advocate a “one mind” position for split-brain subjects rest their case primarily on the behavior of split-brain subjects outside of experimental situations. Outside of experimental situations, “one mind” advocates argue, split-brain subjects act basically like non-split-brain subjects. It is true that *inside* experimental situations, split-brains exhibit some unusual behaviors that non-split-brain subjects do not exhibit even in the same situations; perhaps this behavior should be explained by positing temporary conscious disunity. But since the vast majority of the time they act like us, it is surely *extravagant* to insist that nonetheless they always have two minds or two streams of consciousness or that they are always two persons.

Quite obviously, the claim that split-brain subjects usually act like “normal” subjects only strongly supports a mental singularity claim if normals have one mind. What is the argument for this claim? There is Nagel’s argument that since the concept

of mental unity is based on us, then if it doesn't apply to us, the whole concept ought to be scrapped. But although the concept may have its origins in application to "normal" subjects, why couldn't it turn out that we all have two minds—and why would the concept suddenly lose value? Isn't it valuable to know what "normal" human subjects have two of? What hemispherectomized people have one of?

This is perhaps more easily seen if we think about a stream of consciousness. The very term, "stream of consciousness," reflects its origins in *phenomenology*: consciousness *feels* like a stream in certain respects, and we assume we have just one, because it *feels as if* we only have one. Of course (as both Marks (1981) and Tye (2003) have noticed, by the way), it is arguably impossible for a subject to feel as if she has more than one stream of consciousness, regardless of how many streams of consciousness she in fact has. For by hypothesis any feeling can only occur "in" one stream of consciousness or another and not somehow "outside" of either stream and therefore capable of surveying both. Meanwhile, most philosophers speak interchangeably of having a single stream of consciousness and having a unified consciousness. Now, psychology does presumably owe us an account of the *phenomenology* of consciousness. So suppose a neuropsychologists ultimately discover what accounts for the phenomenology of consciousness, and the phenomenology associated with a stream of consciousness, and suppose they even discover the neural *basis* of the feeling of conscious unity—and then determine that there are *two* such bases in all of us. Do the concepts "conscious unity" or "stream of consciousness" lose all value, simply because they were first rooted in our sense of having a unified consciousness, or a single stream of consciousness—and now a

scientific theory of consciousness says that our consciousness is dual? Of course not. Such a theory would of course need to explain *why* we only feel as if we have a single stream of consciousness even though we have two such streams (and this seems comparatively easy, compared to the extraordinary difficulty of explaining why we have any kind of conscious phenomenology at all), but we might still very much *want* those concepts. For one thing, we would want them because we would want a psychological theory to tell us *why* it feels like we have only *one* stream of consciousness!

Marks and Tye don't seem to be making *conceptual* argument against mental and conscious duality in split-brain subjects, however. Marks and Tye seem to believe that non-split-brain, "normal" subjects have single minds and single streams of consciousness because *they* act in such a *unified* fashion. They then conclude that because split-brain subjects act like these "normal" subjects, this provides very compelling evidence that they, too, have single minds and single streams of consciousness.

One problem with this argument is that split-brain subjects don't usually act exactly like "normal" subjects, even in their day-to-day behavior. Their fine motor behavior is different; they show some memory deficits; they rarely read for pleasure, and so forth. Now, it will probably be claimed that *these* sorts of behavioral differences between split-brain and non-split-brain subjects are irrelevant to the "how many minds?" and conscious unity debates, because, interesting though they may be, they have nothing to do with *how many minds* or streams of consciousness anyone has. The behavior that's relevant to the *unity/disunity* (or singularity/duality) debates

isn't just *any* old behavior but *unified* or *disunified* behavior. The argument from split-brain subjects' *normal* behavior then quickly collapses into the argument from split-brain subject's *integrated* behavior.

The first problem with the argument from integrated behavior—when used to support a “one mind” (or “one stream”) conclusion in split-brain or in non-split-brain subjects—is that it begs the question, since it is unclear how to define “*unified*” behavior except in terms of what (we feel) a single person with a single mind and single stream of consciousness would do. But if unified behavior is *by definition* the kind of behavior exhibited by a single person with one mind and one stream of consciousness, then a subject's “unified behavior” can't simultaneously be used as *evidence* that she is a single person with a single mind and a single stream of consciousness.

The inability to even *define* “unified behavior” apart from some notion of “having a single mind” is clearly a serious problem, and yet we can try to set it aside, as much as possible. Let us say that behavior that *isn't* unified is of the type that two people might engage in when they were in some degree of conflict with each other, and that behavior that isn't like this is unified.²³ Thus split-brain subjects normally exhibit unified behavior insofar as they do not usually slap one hand with the other, or slap their own faces to force themselves to get up after over-sleeping, or violently shake and come to the aid of a loved one simultaneously, or struggle to button and

²³ This does not, of course, offer a definition of “unified behavior” that is independent of the notion of mental unity, because two people aren't in conflict with each other in the relevant sense unless their intentions are in conflict with each other: two people just pushed into each other by third parties in other words aren't conflicting with each other in the relevant (i.e. *psychological*) sense. At this point though I am just trying to describe the type of behavior one mind advocates like Marks and Tye to have in mind when they call such behavior unified or disunified.

unbutton their shirts or pull their pants up and down at the same time. Note that I said “usually”: these are in fact all things that split-brain subjects have been observed to do, *outside* of experimental situations.²⁴

Ultimately, all that “one mind” advocates such as Marks and Tye can really mean when they use the unified behavior of split-brain subjects as evidence for their being single persons is that their daily behavior seems no less the product of a single mind than does the behavior of a “normal” subject, and therefore that their *daily behavior alone* gives us no reason to think that they have more minds than do “normal” subjects.

But given the constraints posed by the hemisphere’s shared *embodiment*, a high degree of behavioral integration is more or less to be expected, regardless of how many minds those hemispheres are associated with. Given the role they make for “unified” and “disunified” behavior in individuating minds, it is not surprising that Marks and Tye are able to find subjects with dissociative identity disorder (DID) more compelling candidates for multiple mindhood than they do split-brain subjects. For the many “personalities” (or “alters”) of a subject with DID generally *take turns* exerting control over the subject’s behavior. Across time, it is easy to behave as if you are at war with yourself: I go on a two thousand dollar shopping spree; I cut up my credit cards that night, and return the clothes the next morning; that afternoon finds me yelling at a Visa representative over the phone that my card has been stolen and I need a new one *immediately*. It is not so easy to act at war with yourself *literally* within a moment, however. What must split-brain subjects *do*, exactly, in order for the “one mind” advocate to find compelling evidence of mental duality? Slit their

²⁴Bogen, 1987; Schiffer 1998; Dimond, 1980; Gazzaniga, 1970; Bogen, 1987; LeDoux 2002.

wrists and call 911 simultaneously?²⁵ The subjects can't, *physically*, take a nap and a quick swim in the ocean simultaneously no matter how *mentally* "disunified" they are. Physical embodiment sets out a huge number of constraints on the degree of behavioral disunity that a *single creature* can exhibit, when "behavioral disunity" is defined as the sort of behavior which *two creatures* in conflict would exhibit.

Nonetheless, let us grant that the *daily behavior alone* of split-brain subjects gives us no reason to think that they have more minds than do "normals." Let us even accept that the unified behavior of split-brains outside of experimental situations provides compelling evidence that they have single minds. Is there any evidence that could nonetheless weigh in favor of the opposite conclusion?

The argument from the unified behavior of split-brains subjects still faces several difficulties. Most obviously there is the fact that split-brain subjects sometimes do not engage in it. "One mind" advocates like Tye of course acknowledge that split-brain subjects act disunified during experimental situations. But, Tye seems to say, split-brain subjects spend more time outside of experimental situations than they do inside, and therefore "act unified" more frequently than they "act disunified," and therefore must *really* have a single mind. So, for example, regarding the claim of conscious disunity or duality in split-brain subjects—a claim we won't examine further until later chapters but that "two mind" advocates and

²⁵ Actually I imagine that there are "normal" subjects (people who have a corpus callosum) who have done these things *nearly* simultaneously—who have swallowed a bottle of pills and a minute later called an ambulance in a panic. This should call into question the extent to which "one mind" advocates rely not only on the "unified behavior" of split-brain subjects but also on the "unified behavior" of *non-split-brain* subjects in order to advance their positions. While swallowing a bottle of pills and then calling an ambulance—or swallowing a gallon of ice-cream and then a bottle of ipecac—are behaviors of the mentally ill, they are also just extreme instances of the kind of disunified behavior that non-mentally ill people exhibit frequently. People break promises to themselves routinely; they begin plotting ways to get out of plans that were their idea in the first place; they wince, internally, even as they say something in anger, etc.

neuropsychologists also generally accept, and that “one mind” advocates generally reject—Tye has written:

The major difficulty faced by. . . . [this] hypothesis, which is. . . . popular. . . . among neuropsychologists, is that, except under the specified experimental controls, there is nothing unusual about the behavior of split-brain subjects. Their behavior is generally just as integrated as yours or mine. What leads to the supposition that split-brain patients have a *disunified* consciousness [and two minds] is their failure to behave in an integrated, coherent way in certain, special experimental situations. But if behavior is the evidence on which the hypothesis of disunity rests, then the fact that split-brain patients behave in an integrated way at other times supports the hypothesis that their consciousness is generally unified
(Tye, 2003: 126; original emphasis)

. . . . and that they have one mind.

But this misrepresents the empirical and explanatory basis of the duality claim. Though it was split-brain subjects’ failure to act in an integrated fashion which first lead to the hypothesis of conscious and mental duality²⁶, that behavior is not the sole evidence supporting it, as it would be for subjects who behaved like split-brain subjects do inside experimental situations but who were in all other ways identical to “normal” subjects. We of course *do* have more evidence than that in favor of the duality claim for actual split-brains. We have neuroanatomical evidence, plus the general realist view that phenomena don’t change the moment we stop staring at them intently. We know that during experiments split-brain subjects behave as if they have two streams of consciousness, and we accept that this is *because of their* abnormal

²⁶ Though even this is in one sense inaccurate: it may not be giving Sperry and Myers all the credit they’re due for the careful design of the original split-brain experiments. Those experiments were designed to test an pre-existing supposition or hypothesis: i.e. that the corpus callosum was a mechanism subserving conscious unity, and that split-brain subjects may always have two streams of consciousness even when their behavior is relatively normal.

neural anatomy. This claim—that the behavior observed during the split-brain experiment is a result of “split-brain” neuroanatomy—is used to understand and to explain the split-brain phenomenon, and is itself given some empirical support by the split-brain studies, studies that seemed to show that dividing the brain divides the mind.

This anatomy, meanwhile, is not a product of experimental design: it is the permanent embodiment of their mental lives. Rey (1975) made this point long ago in the context of arguing that the split-brain cases are “disturbing” because they seemed to show that “it is just these hemispheres that. . . if they are detached from one another, seem to be each sufficiently complex and sufficiently autonomous to be regarded as separate persons. At least they seem so for the duration of the experiments. *But these experiments in no way disturb the underlying structures: they simply lay them bare*” (1975: 7, emphasis mine).

Of course, the claims of mental and conscious duality are supported by neuroanatomical facts about split-brain subjects only given some kind of mind-brain supervenience claim. If we accept externalism about the *vehicles* of cognition and consciousness, then the split-brain experiment may very well alter subjects’ mental architectures. I won’t defend mind-brain supervenience until Chapter Six; if we assume mind-brain supervenience for now, however, it is again unclear how merely sealing a nostril could divide a consciousness, much less create a second mind. Thus while split-brain subjects’ *behavior* of course changes inside and outside experimental situations, surely this is not because their mental *architecture* changes. It must instead be because the *same* architecture can exploit different strategies and

act upon different features of the world in the two types of circumstance. Thus “two minds” advocates point to many sources available to integrate split-brain subjects’ behavior, *besides* “co-mentality” (the relation two mental states bear when they belong to the same mind, or the relationship two mental systems bear when they jointly constitute or constitute in part a single mind). Their primary focus has been on the perceptual redundancy afforded by most environments and by much of the body: since both hemispheres are always in the same place at the same time, and since they have all the same sensory modalities, and have access to representations of many of the same parts of the body and world, they will see, hear, smell, feel, etc., many of the same things.

The reason that mental duality advocates have placed most of their emphasis on the role of perceptual redundancy, when explaining split-brain subjects’ day-to-day behavior, is probably that perceptual redundancy is precisely what the split-brain experiment is designed to reduce. But cross-cuing is another feature often used to explain the integrated nature of split-brain subjects’ behavior outside of experimental situations, and for the same reason. Less usually referred to but (it seems to me) at *least* as important are the behavioral constraints no doubt posed by shared embodiment. Beyond the simple (but vitally important) fact that the two hemispheres cannot *be* two places at the same time, non-cortical structures would play a role in ensuring a degree of behavioral integration, even when the two hemispheres issue very different motor commands simultaneously. (Thus subjects might succeed in engaging in mutually conflicting actions with their two hands at once, but I have never heard of a subject attempting to jump and bend down at the same time—the

two actions would require physically incompatible postural sets, not just intentionally different actions.) Indeed, the constraints posed by shared embodiment go a long way towards explaining why a split-brain subject's behavior is fairly unified even *within* experimental situations. Given these two sources of behavioral integration—perceptual redundancy and, especially, shared embodiment—it will be hard for split-brain subjects to constantly “act as if they are two people.” Meanwhile, shared embodiment is extra-mental, and perceptual redundancy, while mental, doesn't require *interaction* between the hemispheres or integration of their mental processes.²⁷

My point about the role of behavioral evidence in theorizing about the structure of mental architecture isn't, of course, that there is no such role. Behavioral evidence plays just as crucial a role for the advocate of mental duality as it does for the advocate of mental singularity. But it doesn't play this role in nearly so *straightforward* a fashion as Tye implies in the passage quoted above, for several reasons.

First, “one mind” or mental singularity advocates like Marks and Tye suggest that similar or identical behavior should be explained on the basis of (or should be taken as very compelling evidence of) similar or identical mental causes. This seems to assume that we can individuate behaviors on the basis of their physical features alone, and only then begin asking questions about their mental causes. In actual fact

²⁷ In other words, while there are admittedly *mental* causes of integrated behavior in the split-brain subject, not all of these support a “one mind” conclusion. As Gazzaniga (1985) has emphasized, for instance, certain features of human psychology—such as the desire to *see* one's behavior as coherent and rational and voluntarily generated—must also play a role in unifying behavior. Yet these features do not depend on any direct interactions *between* the hemispheres. A similar point applies to cross-cuing behaviors.

we individuate behaviors partly in terms of their mental causes all the time. Telling someone that you have no idea how their car got that huge dent in it is, for the purposes not just of moral but of psychological theorizing, a different behavior depending on whether or not you actually *do* know the origin of the dent. There is, in general, a bi-directional relationship between individuating and classifying behaviors and locating their mental causes. Part of what this means is that there is no entirely theory-neutral way of individuating or classifying behaviors. One cannot, totally independently of a set of assumptions about the mental causes of behavior, individuate all behaviors of X-type (on the basis of a subset of their purely physical features for example), and then use that as *evidence* that they all have the same mental cause. It was the belief or intuition that they all have the same mental cause that was used to identify those behaviors as all being of X-type to begin with. (Or, rather, I suppose that one *could* classify behaviors on the basis of their purely physical features—but such a classificatory schema, which wouldn't distinguish between falling and lying down, would be of no use in psychological theorizing.)

Second, note that the behavior the advocate of the mental singularity model takes as most relevant to the “how many minds?” (or the conscious unity) debate is grossly characterized, everyday behavior. More finely observed and characterized, the behavior of split-brain subjects is less normal.²⁸ Which way should we characterize

²⁸ For instance, they have a moderate memory deficit; they show a slight tendency to confabulate; and, while perfectly friendly (indeed Zaidel, Zaidel, and Bogen (1999) refer to their “inappropriate or exaggerated politeness”), they have an impoverished ability to describe their own emotional experiences (*ibid*). (Might this be because of the voiceless right hemisphere's role in generating emotional experience?) They avoid reading at any length. (Trevorthen (1978), speculates that this may be because reading requires rapid and coordinated switches between right- and left-looking—and split-brain subjects, in contrast to “normal” subjects—apparently look left and right with different latencies.) And they exhibit obvious impairments trying to learn new tasks involving bimanual coordination (Preilowski, 1990).

their behavior, then, grossly or finely? This choice matters. If we classify behavior grossly, a split-brain subject behaves like a “normal” subject, strengthening the case that he has a “normal” subject’s mental life. If we classify behavior finely, or if we direct our gaze upon a much larger (i.e. temporally-extended) pattern of behavior, a split-brain subject behaves less like a “normal” subject, strengthening the case that he does not have a “normal” subject’s mental life.

Note that these early (often implicit) choices about how to classify behaviors also influence our sense of which causes are mental and which merely neural. If we characterize behavior grossly, we may classify and individuate mental causes grossly also, and be content to let the details lie at the level of neural processing. (E.g., split-brain and “normal” subjects have the same mental architecture; any behavioral differences between them result from differences in their neural architecture. We will see a version of this strategy below.) If we characterize behavior very finely, we may only be happy with a very detailed mental explanation, and then the detailed difference between a split-brain and a “normal” subject may suddenly look psychologically relevant after all.

For instance, imagine our split-brain subject, S, in his home, sitting in an armchair, reading. Suddenly he gets up and walks into the kitchen, and proceeds to make a sandwich (using both hands). What is the (psychological) explanation for this behavior? We might intuitively identify “interrupting his reading to make a sandwich” as the behavior to be explained, and explain it by referring to S’s hunger or his belief that it’s lunchtime or that now would be a good time to take a break from reading, his belief that sandwiches are appropriate lunch food and quick to make, etc.

Conveniently, this explanation can be offered without knowing anything about S other than what we've just observed, and as a first shot at a psychological explanation, this one seems fair. So far, no need to mention hemispheres, or even hands.

But if we analyzed S's behavior into smaller components, we might discover that his right handed behavior is controlled only by his left hemisphere and vice versa, or that a single hemisphere was controlling both hands. It might turn out that S's *right* hemisphere was solely responsible for initiating S's standing (out of impatience with reading and a desire to go turn on the television instead) but that his *left* hemisphere was solely responsible for initiating walking to the kitchen, when it interpreted S's standing as being caused not by a desire to watch television (S has recently vowed to cut back on television) but by a desire for lunch²⁹. The left hemisphere belief that getting out of the chair was motivated by a desire for lunch and its beliefs about sandwiches etc. would then be largely but not entirely responsible for the entire behavior, for distinct right hemisphere events—some of which look mental—initiated the whole sequence.

I don't claim that split-brain subjects' behaviors are often caused in this way; I don't think we know these sorts of details yet. My point is more general. Casual observation may not reveal whether a split-brain and a "normal" subject are exhibiting the same behavior in the *relevant* sense, where relevance is (as circular as this may sound) in part a matter of *what, psychologically, caused the behavior*. And

²⁹ Gazzaniga has in fact for a long time (at least since Gazzaniga and LeDoux, 1978) suggested that what he now calls the "left hemisphere interpreter" (see for instance Gazzaniga, 2000) is responsible, in all of us, for in effect continually asking the question, "Why am I doing *this*?" and then constructing answer—one that may or may not describe the *actual* mental causes of behavior.

casual or even careful observation *alone* certainly won't answer the extremely difficult theoretical question of which causes of behavior are mental and which not.

Probably there is no single best way to characterize anybody's behavior, grossly or finely, at all times. Often grossly characterized behavior may be just what we're after. And grossly characterized, split-brain subjects' everyday behavior does in fact seem "unified"; it is certainly adaptive. But grossly characterized behavior often isn't a great help in letting us locate the precise causes of that behavior. We may not always be *interested* in knowing the very detailed causes of behavior, but *when* we are, we may want to individuate characterize the behavioral explanandum more finely.

Gross behavior often isn't a great help in letting us locate precise mental causes especially when everything is working well. The behavior of a psychologically, intellectually, neurologically normal human adult, for example, can even make the notion of a "general purpose learner" or an "indivisible mental substance" look intuitively plausible. It is the blindsighted patients, the subjects with autism and agnosia, the child struggling to learn irregular verbs, or Damasio's (1995) ventromedial prefrontal cortices-damaged patients (who do so well on standardized tests of intelligence and moral reasoning and yet fare so poorly in the game of life) who make it a little easier to discern the architecture of *all* humans' minds. Split-brain *experiments* are specifically designed to identify the *precise* causal, mental basis of split-brain subjects' behaviors; the behaviors that emerge during those experiments, then, may be the best guide to determining the true structure of their mental lives.

Tye seems to suggest, however, that split-brain's disunified behavior at some moments (whether in or outside of experimental situations) isn't even *relevant* to determining how many streams of consciousness they have at other moments, when their behavior is unified. Behavior that occurs at t_2 , in other words, is irrelevant to determining the mental cause of a behavior occurring earlier at t_1 or later at t_3 . This cannot be correct. We do want mental explanations that account for local, immediate behavior, but these explanations should be consistent with more global patterns of observed behavior. If the "how many minds?" question is a question about mental architecture, and minds, rather than being mere *collections* of mental states, are defined architecturally, then we would expect basic *architecture* to remain largely stable over significant periods of time. What is much more fluid is the set of mental states (especially their contents) someone is subject to at any moment; we can easily imagine these changing (they'd *better* change!) in and outside of experimental situations. The "psychological frameworks" Tye (2003) apparently identifies with minds lack the stability we would probably want minds to have, since he defines these frameworks only in terms of coherent sets of mental states. Mental architectures, however, do have this stability.

Finally, the behavior of split-brain subjects also isn't the *only* relevant evidence for determining how many minds or streams of consciousness they have. There are at least some obvious, non-controversial limits to the significance of unified behavior in developing an adequate model of a subject's mental architecture. Imagine two people yoked front-to-back to each other, with some kind of structure encasing their limbs such that two right arms can only move in unison, two left arms can only

move in unison, etc., producing unified behavior. Or imagine that the set-up is even more sophisticated than this; their brains are in fact wired up to the encasing structures, which moves their limbs in unison for them, and whenever conflicting motor plans reach the control unit for the structure, the strongest motor impulses win. (Or perhaps the control unit just gates one person's motor plans for one hour, than the other person's motor plans for the next hour, and so forth.) Or for that matter suppose when you looked inside the skull of a subject whose behavior was perfectly unified you saw to your shock that there were actually two tiny little guys in his brain—with arms and legs and language and everything—controlling his behavior!

Such examples are fantastical, but they show, again, that there are non-controversial limits to the significance of even the most integrated behavior in theorizing about mental structure and in individuating minds. Psychological theories should of course be empirically adequate with respect to behavior. But, as essential as this criterion for theoretical adequacy is, it is still minimal. Psychological theories should also cohere with what we see at the level of neural implementation.

No realist functionlist, of course, would disagree with this claim. But one might still argue that because we have *neural* architecture to refer to in the split-brain case to explain occasional instances of experimentally-induced dissociative or disunified behavior, this obviates the need to refer to any kind of unusual *mental* architecture to explain that same behavior. Thus one might interject, “Why can't we just explain the lack of direct interaction between, say, one of S's conscious (right hemisphere) experiences and another of S's conscious (left hemisphere) experiences, in *purely* neural terms—by referring to S's callosotomy—and leave it at that? The

explanation for occasional instances of clearly disunified behavior in S, is, after all, very easy to explain in broad strokes, simply by referring to S's unusual *neural* architecture. After having offered this explanation, why take the extravagant second step of attributing to S two minds?"

Tye appears to appeal to this sort of explanation at one point in the context of arguing that split-brain subjects aren't two persons. Of a case in which the word "pen" has been projected to the right hemisphere of a split-brain subject, he writes: "The fact that the split-brain patient doesn't say 'pen', when asked what he saw, doesn't show that he doesn't believe that he saw 'pen'. He does believe that. It's just [that], given the commissurotomy, he can't verbally express that belief" (2003: 116). Thus there is a *physical* explanation for why the subject does not say that he just saw a pen. No further psychological explanation is necessary. And because there is such an obvious physical explanation for what might otherwise be puzzling behavior, we need not postulate the existence of multiple minds—merely an unusual neural architecture.

It is of course true as a general rule that apparently inconsistent attributions of mental properties may be easily reconciled when we learn more about the neural architecture underlying them. For example, we can say that a stroke victim with visual agnosia consciously sees a fork, knows what a fork is, and yet cannot verbally identify it as a fork. We can say all this about the stroke victim knowing that the stroke has compromised some high-level visual areas of her brain, depriving her, perhaps, of the visual template for "fork," and therefore making her incapable of *seeing* a fork *as* a fork, though she can verbally identify a fork by touch. Note,

however, that we can say all this about the stroke victim not just because she has an “unusual” neural architecture, but because of what we know about that architecture. Thus in the case of a stroke victim who can verbally identify a fork as a fork after touching it, but not after looking at it, there is of course no need to postulate multiple subjects of experience, for instance—a feeling subject who knows what a fork is and a seeing one who doesn’t.

But commissurotomy is a particular, not a general, case, with at least one strikingly different feature from (most) other unusual neural architectures. While many cases of unusual neural architecture may involve some compromised *access* between mental systems, and while the split-brain case does of course involve this, in the split-brain case each of the mental systems in question—that associated with the left hemisphere and that associated with the right—is capable of supporting a mind “comfortably characterizable as human” (Marks 1981: 47, fn. 18). Both hemispheres can sustain (I am assuming along with Tye) conscious experiences—experiences to which the other hemisphere is not subject. Either hemisphere can guide intelligent behavior, apparently without the help (indeed, sometimes despite the hindrance) of the other. There is even evidence that, in split-brain subjects, the hemispheres separately sustain autobiographical memories and different emotional responses to these memories (Schiffer et al., 1998). Evidence not only from split-brain studies but from those rare cases of left or right hemispherectomy following (basically) normal development lends further support to the conclusion that each hemisphere is at least a

candidate for constituting a unique mind.³⁰ (Tye himself apparently acknowledges this in a slightly different context; see p. 150).

In the split-brain case, in other words, the very neural architecture, reference to which might be claimed to make unnecessary the positing of multiple minds, is in fact one of the strongest pieces of evidence in *favor* of there being multiple minds to begin with. For if the psychological entities we're discussing don't reduce in some sense to these neural systems, then either I am wrong to assume that minds are (constituted by) brains, or else the discussion we're having isn't one of scientific, as opposed to folk, psychology to begin with.

I will continue to take the integrated nature of split-brain behavior as evidence *against* the mental duality model. As I have argued in this section, however, it is defeasible evidence.

4 Perfect Parallelism: Does Redundancy Matter?

I have argued that purely psychological and partly neural criteria for individuating minds will, all other things being equal, tend to yield the same answer to the "how many minds?" question. It might be said, however, that we can conceive of a particular kind of case that shows that the purely psychological and the partly neural criteria will *not* yield the same answer to that question. In fact it might be said that the split-brain case *is* that case. And it might further be argued that in this special kind of case, the purely psychological approach to individuating minds yields the better answer, from the functionalist perspective.

³⁰ See Boatman et al., 1999; Burklund and Smith, 1977; Gott, 1973; Smith, 1966, 1969.

In this section I examine what I call the “singularity-through-redundancy” (STR) position: the claim that multiple neural events of the same psychological type and carrying the same content can be identified with a single mental token, even if these neural events operate causally independently of each other. I will approach this claim largely via the writings of Marks (1980) and Tye (2003), since they have developed and defended this claim most explicitly. But the claim enjoys broader appeal; Dennett, for example, gestures towards a similar position (1991), and even Sperry seemed to feel its pull at times (see his 1975 for example). One goal of this section is to help motivate a “two minds” conclusion for split-brain subjects. A larger goal, however, is to defend a partly neural set of criteria for individuating mental tokens. The underlying concern of this entire chapter is whether the causal relationships that multiple neural events bear, or fail to bear, to each other, constrain the mental events with which we can identify them. I submit that they do.

The STR position has been defended most explicitly as a position on the structure of “split-brain” *consciousness*, rather than “split-brain” cognition or mentality generally. Because I wish to address particular arguments made in existing developments of this position, I will accordingly be talking about consciousness in this part of the chapter, even though a general discussion of the structure of “split-brain” consciousness will need to be postponed until a later chapter. An examination of the *theoretical* reasons why singularity-through-redundancy fails as a characterization of split-brain subjects’ *consciousness*, however, is instructive in the context of determining how many *minds* split-brain subjects have also, and the

conclusions drawn from this examination have some wider applicability to discussions of individuating mental tokens generally.

4.1 STR: Meaning, Motivation, and Empirical Adequacy

Marks and Tye say that “normal” subjects behave in a unified fashion, and that we believe that this is partly due to these subjects each having a single stream of consciousness or a unified consciousness. In their daily lives, meanwhile, the gross behavior of split-brain subjects usually resembles that of these “normal” subjects. Marks asks, “If we account for our integrated behavior, at least in part, by assuming unity of consciousness and can do the same for split-brain patients, why not do so?” (Marks 1981: 22)

Marks and Tye accept that there are occasional times during the split-brain experiment at which split-brain subjects have two streams of consciousness. Notwithstanding such moments they believe that a split-brain subject’s consciousness is normally singular or unified. They therefore face the challenge of showing how the split-brain experimental paradigm could alter the structure of a split-brain subject’s consciousness, especially since it does not alter that of the “normal” subject. (“Normal” subjects do not exhibit the conscious dissociation under conditions of perceptual lateralization that split-brain subjects appear to.) The solution both philosophers take is to say that conscious unity—which they equate to having a single stream of consciousness—can supervene on certain properties of conscious *contents*. When present, the corpus callosum ensures conscious unity even under conditions of lateralized perception, by allowing some kind of interhemispheric *communication* of contents. But even without a corpus callosum, the two hemispheres of a split-brain

subject will normally be subject to the same (or to highly similar) conscious contents, Marks and Tye (and many others) believe, because they will be receiving the same (or highly similar) information from the environment and the body. The singularity-through-redundancy claim asserts that what Marks calls the “independent duplication of information” suffices for a unified consciousness. Multiple neural events of the same mental type (e.g. conscious experience), and bearing the same content, can together constitute a single mental token, even if there is no causal interaction between them.

One distinction that I will draw again in Chapter Four, on the structure of split-brain consciousness, but which is important to make here as well, is that between having a *unified* thing or things, and having a *single* thing, unified or disunified. In our daily language we generally treat “unified consciousness” and “single stream of consciousness” as near-synonyms—as mere different parts of speech. Yet some of those who have argued that split-brain subjects have two minds and two streams of consciousness prefer not to speak of mental or conscious *disunity* so much as *duality*.³¹ They believe that, particularly since the hemispheres may be often or even always associated with similar psychological states, histories, and dispositions, it might be misleading to describe the hemispheres as mentally or consciously disunified, for the term “disunity” connotes conflict and discord.

Ascribing a *dual* consciousness has fewer of these potentially misleading implications, and so, too, I submit, should ascriptions of multiple streams of consciousness. Since I am concerned with individuating mental tokens, and not with providing a qualitative analysis of their character, I prefer to speak of conscious

³¹ Bogen (1990) seemed to prefer this formulation, for example.

singularity and duality, of having one stream of consciousness and having two streams of consciousness, rather than speaking of unity and disunity.

Defenders of the singularity-through-redundancy characterization of split-brain subjects' consciousness are also interested in individuating streams of consciousness (though Tye especially is interested in providing some analysis of their phenomenal character as well). So they do not merely believe that the two "disconnected" hemispheres are largely unified in some sense or other. They claim something stronger. Defenders of the STR position, like Marks and Tye, equate having a single stream of consciousness with having a unified consciousness, and having two streams of consciousness with having a disunified consciousness.³² They therefore claim that to the extent that the two hemispheres are associated with the same conscious contents, they are associated with the same conscious tokens. It is this claim that I reject.

Now one objection to the singularity-through-redundancy characterization of split-brain consciousness is empirical: the two hemispheres are almost certainly *not* subject to identical, or even to highly similar, conscious contents outside of experimental situations. They are no doubt subject to *more similar* contents outside of experimental situations than they are inside of them, and this no doubt offers at least a

³² Tye rejects attributing to split-brain subjects "two separate streams of consciousness that remain two from the time of the commissurotomy" in favor of saying that such subjects "are single persons whose phenomenal consciousness is briefly split into two under certain special experimental conditions, but whose consciousness at other times is unified" (2003: 126, citing Marks 1980 as well). Likewise, as a "rough necessary condition for two simultaneous conscious experiences belonging to the same stream of consciousness", Marks offers, "e1 and e2 belong to the same unified consciousness only if they are known, by introspection, to be simultaneous" (1980: 13, emphasis added), thus tying possession of a single stream of consciousness to possession of a unified consciousness. Both philosophers meanwhile also accept what is of course widely accepted, i.e. that a stream of consciousness is composed of token experiences (or in Tye's case, a single token experience—this is one of the main positions argued for in Tye, 2003).

partial explanation for the fact that split-brain subjects behave differently in the two types of circumstance.³³ Nonetheless, since the hemispheres have different patterns of perceptual access to the world, and since they also have different processing styles and capacities, and appear to experience emotions somewhat differently, and to have access to a somewhat different store of long-term memories—not to mention the fact that one hemisphere can presumably generate a fairly normal stream of inner speech and one hemisphere probably can't—it would be a stretch to imagine that the hemispheres are associated with highly similar, much less identical, conscious contents.

This empirical objection, however, is irrelevant to determining a general method of individuating minds and other mental tokens, and I therefore set it aside in most of what follows. For most of this section I will assume that the two hemispheres of a split-brain patient are subject to *highly similar* conscious contents; eventually I will consider what we should say of a subject who possessed two hemispheres that independently generated neural events bearing truly indistinguishable psychological properties at every moment.

³³ The split-brain experiment, however, is also typically designed with the goal of eliciting a response from a hemisphere that might otherwise not respond; this probably plays a role in the increased degree of behavioral disunity patients exhibit in the labs. Schiffer, Zaidel, Bogen, and Chasan-Taber (1998) elicited different answers to questions about an unhappy childhood experience from the two hemispheres of a split-brain subject, for example, by putting two sets of pegs, five in each set, in front of the patient but obscured from his vision, where the leftmost peg in each set represented “none” and the rightmost represented “extreme,” and then requiring the subject to answer questions like, “How much did this experience upset you at the time?” using both hands at once. The subject P.S. (see Gazzaniga, LeDoux, 1978) similarly gave different responses when asked what he wanted to do professionally as an adult, depending on whether he was asked to respond verbally (using his left hemisphere) or in writing using his left hand (using his right hemisphere). Same stimuli, in other words, but different processing and thus different, “disunified” behavior.

4.2 Constituting a Single Token

The singularity-through-redundancy position holds that multiple neural events may jointly constitute a single mental event regardless of the nature of their causal relation to each other.³⁴ This section examines two analogies Tye (2003) offers in support of this position and argues that both analogies are inappropriate to that purpose. The first seems simply non-analogous, and the second is subtly question-begging against those who would claim that the two hemispheres of a split-brain subject are associated with distinct mental tokens.

4.2.1 The projector analogy

Tye anticipates that some will object to STR by saying that, “a single experience cannot have as its physical basis neural events in the left and right hemispheres that are themselves causally unrelated.” But, he asks, “why not?”:

Consider the following example. Two movie projectors each project an image onto a screen at time *t*. Only a single image is present on the screen at *t*, since identical slides are in the projectors and they are aimed at exactly the same part of the screen. There are two projections but only one image. One projection is redundant. Each projection on its own suffices for the screen image.

(Tye 2003: 127)

There are three difficulties with this analogy. Note, first, that the redundancy in the “contents” of the projections is in fact irrelevant to their generation of a single image; one machine could project an image of a child at play, the second of a hulking

³⁴ At least, so long as they are both located in the same creature or brain; I am sure that neither Marks nor Tye would allow that one of my neural events and one of my Twin-Earth-twin’s neural events could jointly constitute a single mental token. But it is not clear if this is just supposed to be a brute fact about creatures or brains. Part of what I am suggesting is that some of the same considerations that would weigh against identifying my neural event and my twin’s neural event with a single mental token also weigh against identifying a split-brain subject’s right and left hemisphere neural events with a single mental token.

monster, and the result would be a single image of a monster lurking over an unsuspecting child at play. The generation of a single image is a function of the direction in which the projectors are aimed, rather than of the contents they are projecting.

More troublingly, the two projections appear not to *realize* but to *produce* the image. One could, for example, increase or decrease the number of images without manipulating the projections, simply by moving (or removing altogether) the screen. But everyone will accept that a mental event can causally depend upon two independently acting neural events. The relationship between realizer and realized is more intimate than the relationship between cause and effect, and Tye needs to provide an analogy more clearly involving the former.

Even if the projections did jointly *realize* the image, however, the doubly-projected image offers an inadequate analogy to a case of *mental* singularity through *neural* redundancy in particular, because the individuation criteria for images and experiences are simply too different. Clearly an audience would *see* a single image, so long as all projections terminated on the same portion of the screen. And individuating images is arguably just a matter of determining how many images a normal viewer (a human being whose visual system is functioning normally) would see. But as Tye himself admits, the fact that split-brain subjects don't "see double" or *experience* having multiple streams of consciousness is to be expected, *regardless* of how many streams of consciousness any of them actually have. When it comes to counting images on a screen, in other words, phenomenology is everything. But when

it comes to counting experiences and streams of consciousness, particularly those that *do* have redundant contents, phenomenology may not tell us much at all.

If the disanalogy between individuating images and individuating experiences isn't immediately obvious, it is because this first analogy plays off of our naïve or pre-theoretic notion of conscious experience as, to borrow Dennett's (1991) oft-borrowed phrase, a *Cartesian theater*. Once we rid ourselves of the illusion that conscious experience is a thing we *watch*, it is unclear what a conscious experience is supposed to be analogous to in this example. Maybe to the beam of light? But there are *two* of those.

4.2.2 The restaurant analogy

In his second analogy, Tye, while seated at a restaurant, waves both of his arms at once to catch his server's attention. There are two arm-wavings here, he says, and yet one event of signaling the server.

Note that this fails to show that either arm-waving is *redundant* in the sense of something unnecessary to giving the signaling event its *character*. There are, for example, social and psychological differences between trying to catch your server's attention by raising one hand, and trying to catch his attention by waving both hands wildly in the air, just as there are social and psychological differences between trying to catch his attention by waving both hands wildly in the air, and trying to catch his attention by shouting "Hey, boy!"³⁵

³⁵Tye himself introduces the example by asking us to imagine a "case in which I am in a restaurant, and, *being anxious to leave*, I signal the waiter by raising both hands in the air and waving them" (2003: 127; emphasis mine). Tye's feeling it necessary to imply that of course he wouldn't wave both hands in the air to catch his server's attention unless he were particularly anxious constitutes an inadvertent admission that neither arm-waving is truly *redundant*, since both arm-wavings are necessary to give the signaling event its precise social and psychological significance.

Nonetheless, even if the event of signaling with one hand and the event of signaling with two hands have a different character, each appears to be but a single event, and *individuating* events (tokens), rather than providing a qualitative analysis of their nature, is our current concern. Still, it is worth asking why Tye seems right that waving both arms in a restaurant is one way of realizing a single event. How do we know that waving two biological arms in the air isn't just one way of realizing *two* waving-one-arm events—waving two prosthetic arms being a different way of realizing two waving-one-arm events?

Actually, waving two biological arms in the air may be one way of realizing two waving-one-arm events. But Tye *believes* that there are two *waving-one-arm* events. His claim is that there is just one *signaling* event. How many events some spatiotemporal region constitutes or contains is relative to a particular level of description. There is a single *signaling* event in Tye's restaurant case because *signaling* one's server is a *communicative* act, and we individuate such acts partly in terms of *intention*. Thus if two parties seated near each other talked and decided (their situation being somewhat desperate) to both wave their hands at the same time in the hopes of finally catching their server's attention, they arguably both participate in a single signaling event. But if the same two parties both waved their hands at once by mere coincidence, then there were arguably two, simultaneous signaling events.

Part of the reason we may find it hard to conceive of a reason we would want to say, of the original case that Tye provided, that there are two signaling events, is because we are just so used to thinking of a single human organism as having a single mind and a single set of intentions (at a time) that are (in *some* sense) integrated prior

to behavior. In Tye's original "restaurant" analogy, the two arm-waving events aren't initiated causally independently of each other in the way we deem relevant to individuating communicative actions, because we attribute to *Tye* a *single* intention to *signal by waving both arms*.

But of course precisely what is under consideration is to what extent this same sort of attribution is warranted for split-brain subjects. Identifying Tye's two arm-waving events with a single signaling event rests upon the assumption that there was a single mental event—a single intention to signal by waving both arms—*causally responsible* for *both* realizing (arm-waving) events. Proponents of the conscious and mental duality models claim that the two hemispheres of a split-brain subject are associated with *distinct* mental events that drive the subject's (unified-seeming) behavior, causally independently of each other, to a significant degree. Tye obviously disagrees, but his restaurant analogy offers only question-begging support for his position.

Note again that it isn't just the *number* of token intentions to signal that determine how many signaling events we locate. In the case in which two parties waved at once, each possessed a unique token intention to wave; whether we see one or two signaling events depends upon whether there was an *interactive process* leading to the formation of the two token intentions. We ask the parties, "Did you talk to each other, and come to a mutual decision to jointly signal your server? Or did you each come to the decision to wave independently—was it mere coincidence that you both waved at once?" We try to determine whether there was a certain kind of causal relationship between the admittedly distinct token intentions that produced the

wavings. So it is, too, when the events we're counting aren't communicative actions but cognitive acts.

4.3 Three Claims on Irrelevance

This subsection turns to three distinct claims supporting the singularity-through-redundancy position, particularly as applied to the structure of split-brain subjects' consciousness. These claims all concern what, if any, psychological significance attaches to the causal relationship that holds or that fails to hold between the neural events in whose mental identity we are interested.

4.3.1 Multiple realizability and the type-token distinction

The first claim supporting the singularity-through-redundancy position asserts that facts about the realizers of mental phenomena are mere implementation facts, of no significance to their psychological identity.

Fodor (1975) has argued that the kinds postulated by psychology will not reduce to the kinds postulated by any purely physical (non-functional) science, because the kinds of psychology—the types of entities to which it will refer in its laws—are *multiply realizable*. What makes a given mental event a pain event, or an event of believing something, as opposed to some other kind of mental event or no mental event at all, isn't a matter of its intrinsic physical properties, but the role it plays within a functionalist story, a role connecting it to sensory inputs and to motor outputs and to other mental events. Thus creatures whose physical design was radically different from ours could nonetheless possess some or all of the same types of mental states that we do. All that would be required is for their physical design to

somehow implement a functional design similar to that implemented by our own brains (whatever design that turns out to be).

The principle of multiple realizability is foundational to the functionalist program, and has been appealed to in order to support the singularity-through-redundancy claim. Marks answers the question, “Why should neural processes unrelated by direct causal routes not be the physical basis for a single mental state?” (Marks 1980: 23) in part by citing Fodor (1975). Tye says that when a split-brain subject’s right and left hemisphere, causally independently of each other, each generate a neural event of the type *experience* and carrying the same content, “There is. . . . a single experience. . . with a neurological realization or constitution that is partly redundant and that differs from the neurological realization or constitution” that that single experience would have “in normal subjects” (Tye 2003: 127). Same mental events, different realizers, in other words.

By casting the neuro-functional differences between split-brain and “normal” subjects as mere differences in *how* experience is realized, rather than as differences in *how many* experiences are realized, those sympathetic to the singularity-through-redundancy position seek to ward off a simple objection. This objection states that while the hemispheres of a split-brain subject, S, may bear the same *types* of experiences (where contents are included in type), they surely bear distinct token experiences, for S’s hemispheres realize the experiences they realize causally independently of each other, and do not have a means of transmitting to each other the contents of at least the bulk of their experiences (absent behavior).³⁶ This

³⁶ Due to the fact that the hemispheres remain connected to each other via mutual connectedness to non-cortical structures (and often via the anterior commissure), there do appear to be *certain* types of

objection rests on the intuition that facts about the causal properties of a neural event constrain the mental events with which we can identify it.

Drawing on the language of multiple realizability, Marks and Tye respond that such facts are facts about how psychological phenomena are *implemented*, but are not psychological facts properly speaking. We can, and should, refer to those facts at some times—such as when we need to explain why a split-brain subject’s consciousness, and not that of a “normal” subject, becomes dual or disunified³⁷ during the split-brain experiment. Marks writes that this disunity:

is itself explained by the fact that the experimental controls defeat the mechanisms which are, as a result of commissurotomy, responsible for unity of consciousness in these patients. Similarly there is a natural explanation for the behavioral differences between split-brain patients and normal controls. The mechanisms which subserve unity of consciousness in the normal controls differ in ways that make them immune to failure in the experimental situation.

(Marks 1980: 22-3)

But facts about functional neuroanatomy are facts about the mechanisms subserving various mental phenomena, not facts about *which* mental phenomena are actually being subserved. Marks believes that the moral of the split-brain cases is that:

bilateral neural representation is a physical basis for unity of consciousness; but *it is irrelevant how such representation is*

mental states that the hemispheres generate or experience in an interactive way. Emotional states may represent an exception to the general rule that the two “disconnected” hemispheres generate experiences independently of each other; certain other feeling-like states, “affective mental auras” (as Sperry, Zaidel, and Zaidel, (1979), once put it) or impressions seem the product of interhemispheric interaction as well. My fundamental concern in this section isn’t with the structure of split-brain subject’s consciousness, however, but with the singularity-through-redundancy claim. That claim doesn’t hinge upon the possibility of *any* kind of interhemispheric *interaction* in the generation of experience; it rather asserts that even *without* such interaction, two hemispheres can jointly generate a single stream of consciousness. I therefore ignore here as much as possible many complications and subtleties regarding interhemispheric interaction in split-brain subjects. Many of these will be returned to in Section Six, however.

³⁷ Recall that for Marks and Tye “dual” and “disunified” are the same thing.

achieved, whether through the commissures or through mechanisms for independent duplication.
(Marks 1980: 22, emphasis added)

Facts about how conscious representation is achieved in the “normal” and in the “split” brain should not constrain the individuation of psychological entities—such as streams of consciousness—which are, after all, multiply realizable.

If the principle of multiple realizability is correct, then creatures with a wide variety of physical constitutions could possess streams of consciousness. This does not mean that physical facts are irrelevant to individuating *particular* streams of consciousness, however. A mental token *is* a physical thing—a realizer of a particular mental *type*. The thesis of multiple realizability, and the fact that, as Marks says, the “general account of mind advocated by philosophers as diverse as Fodor and Grice. . . does not require corresponding types of neurological states for each *type* of psychological state” (Marks 1980: 23, emphasis added), provides no support for the *further* claim that causally unrelated neural events can form “the basis for a *single mental state*” (Marks, 1980: 23; emphasis added). The latter is a claim about mental *tokens*. But the principle of multiple realizability doesn’t say anything about how to individuate experiences or other tokens. It is silent on this matter. Since the principle of multiple realizability provides no guidance where counting mental tokens is concerned, we must turn elsewhere for this guidance.

4.3.2 Is causal independence psychologically relevant?

The defender of STR must claim that multiple neural events (again, at least within a single creature or brain) can constitute a single mental token regardless of their causal relationship to each other. There is a broader and a narrower way of making this

claim. The claim could be interpreted simply as stating that mutually causally independent neural events can nonetheless together constitute a single mental token. Alternatively it could be interpreted as stating more narrowly that there is a particular *class* of cases in which neural events can constitute a single mental token causally independently of each other: that class of cases in which the neural events in question have “redundant” psychological properties. I will consider the broad interpretation first and the narrower interpretation next.

Marks says that “the crucial causal principle” underlying the intuitive objection to the singularity-through-redundancy claim, i.e. the principle that the neural events which constitute a single token of a mental type must be themselves be causally related in some way, “is not strongly motivated. Why should neural processes unrelated by direct causal routes not be the physical basis for a single mental state?” Psychofunctionalism certainly doesn’t “require any direct causal links between the neural events which are the physical basis for a single psychological state. It would be sufficient if causally unrelated neural events jointly, though separately, produced effects which were, from the standpoint of the psychology, the basis of a single mental state” (Marks 1980: 23).

Anyone familiar with the *binding problem* (or problems) may second Marks here: no direct causal links (on an intuitive notion of “direct”) are required between neural events in order for them to realize a single mental state. The binding problems (roughly) concern how it is that we perceive objects possessing multiple perceptual properties, and entire scenes composed of such objects, given the fact that perceptual information, both across and within sensory modalities, is processed in a distributed

manner in the brain, such that color and shape, for example, are represented in different parts of the brain. If we state *a priori*, as it were, that two neural events can only constitute a single mental event when they are directly physically linked (in some intuitive sense of “directly”), then we may never find the neural correlates of the mental tokens that we surely possess. Marks would presumably say that a psychofunctionalist theory need not await a neurophysiological solution to the binding problem in order to speak confidently of the mental token that is my *visual experience of my laptop* (one representing it as having a certain shape *and* color *and* location).

But even assuming that all this is correct, the kinds of causal independence the binding problem concerns appear irrelevant to the kind of causal independence we’re talking about in the split-brain case. To begin with, behavioral evidence alone allows us to conclude that *whatever* neural mechanisms are responsible for phenomenal and functional binding in split-brain subjects, they operate intra-hemispherically (at least largely).³⁸ Under conditions of perceptual lateralization, perceptual information is *not* bound, either phenomenally or functionally, across the hemispheres.

Kingstone and Gazzaniga (1995) showed the right and left hemispheres of split-brain subjects different halves of compound words or word pairs, and yet the percept of each half was not joined phenomenally with the percept of the other half: subjects would report (left hemisphere) having seen the word “dog” rather than the (compound) word “hotdog”, for example. The two percepts also did not appear to be

³⁸ Again, I am ignoring the potentially major exception of binding via subcortical structures, largely because the singularity-through-redundancy position does not require any wholly brain-mediated interaction between the hemispheres. The shape these structures give to split-brain consciousness will be considered in the chapter dealing with split-brain consciousness in particular.

bound together functionally somehow: when asked to *draw* the referent of the word they'd seen, subjects sometimes drew a dog, and sometimes a sun or a stove or a thermometer, but never a frankfurter.

Supporters of the singularity-through-redundancy position might point out that on those trials for which they were allowed visual feedback while drawing, subjects did sometimes draw a dog sweltering under a hot sun. They did sometimes draw a picture integrating the referents of both halves of the compound word pairs, that is, if not a picture of the (single) referent of the compound word. Doesn't this show that *looking* and *drawing* behaviors can be used to bind right and left hemisphere percepts functionally, if not phenomenally? But as the authors pointed out, the production of this sort of drawing, "does not reflect an *internal*. . . transfer of. . . information." "Indeed," the authors continue, the only time that right and left hemisphere "word information. . . [is] integrated is *on the sheet of paper in the drawing itself*" (Kingstone & Gazzaniga, 1995: 324, emphasis added). Though the *behavior* might have been "unified" in some sense or other, interhemispheric percepts and processing were *not* bound in the sense relevant to the binding problem, nor in the sense relevant to individuating mental tokens. After all, if you were shown the word "hot" and I was shown the word "dog" and we both took a turn with a pencil, we might produce the same drawing.

Again, then, *whatever* neural mechanisms are responsible for binding, they operate (largely) intra-hemispherically in split-brain subjects. But there is a more general principle lurking beneath this empirical point. The various proposed solutions to the binding problem (see for instance Golledge, Hilgetag, and Tovée, 1996;

O'Reilly, Busby, and Soto, 2003; Robertson, 2003; Treisman, 1996; Treisman and Gelade, 1980) all seem to require that multiple neural events constituting a single mental token at least have causal connections to further neural events, some of which also have a mental description. But interhemispheric interaction in split-brain subjects by and large does not supervene on wholly neural and mental events. It supervenes partly on behavioral and environmental events—as in the study cited above, in which a split-brain subject, S, can produce an “integrated” picture only if S’s right hemisphere can *watch* S’s left hemisphere *drawing* a dog, after which point, S’s right hemisphere can add its own pictorial contribution.³⁹ Interhemispheric neural event pairs in split-brain subjects do not *jointly* produce mentation. They do jointly produce *behavior*: there are no doubt tons of behaviors in which split-brain subjects engage that involve the joint participation of both right and left hemisphere. But then there are many behaviors in which my sister and I engage which involve the joint participation of events located in her brain and events located in mine.

I suspect Marks is right that two neural events need not interact with each other directly, physically, in order to constitute a single mental event.⁴⁰ While the various and diverse proposed solutions to the binding problems all seem at least compatible with this assertion, they also all recognize a limit to how causally distant from each other these neural events can be, before we cease being able to recognize

³⁹ Of course, subjects will draw a frankfurter outside of the split-brain experimental paradigm, when they are allowed to read the compound “hot dog” naturally, scanning from left to right such that both parts of the compound are scanned in both visual fields, so that each hemisphere sees the compound. But, first, the production of such a drawing relies upon behavioral and environmental events, just as above, and second, it is quite likely that the drawing will be the product of a single hemisphere (probably the left, given other results that obtained in this study) rather than of both, thus still not reflecting any interhemispheric binding.

⁴⁰ In fact they might not need to *interact*, technically, with each other at all, so long as they both interacted with a *third* neural event in an appropriately direct way; then, perhaps, the three neural events together could jointly constitute a single mental token.

them as constituting a single realizer, rather than as distinct realizers, of mental phenomena. Right and left hemisphere neural event pairs in split-brain subjects are for the most part located at a point past this distance, for their interaction does not occur prior to the behavior that they cause. When duality theorists say that in a split-brain subject the right and left hemisphere generate experience causally independently of each other, *this* is the kind or degree of causal independence they have in mind.

4.3.3 Is neural redundancy psychologically relevant?

There is a narrower claim that the STR position might be interpreted as making. This is the claim that even if, as a general rule, the mutual causal independence of two neural events requires identifying them with distinct mental tokens, there is an important exception to this rule: those instances in which the two neural events in question have indistinguishable mental properties. In such instances, either set of events seems *redundant* from a psychological standpoint. Why, then, identify each set with a unique set of mental tokens?

After all, Marks calls it “a commonplace that the neural structure of human brain [sic] is highly redundant”; why should it, “be surprising if the redundancy is sometimes irrelevant for the purposes of psychology” (Marks 1980: 23-24)? Tye, too, says that:

There seems no obvious reason why nature should not have made us so that, in certain circumstances, there is redundancy at the neural level in the generation of perceptual experience. After all, it is well known that the human brain has a neurological structure that is highly redundant anyway. Why not also here?
(Tye 2003: 128)

The purported redundancy is from the psychological perspective but at the neural level: there are two *realizations*, but a single mental token, of a mental type. The realizing events are thus psychologically redundant, and the fact that there are two of them is psychologically irrelevant.

This analogy to the “well known” redundancy of the human brain may be subtly question-begging, however. Two neural events need do more than bear the same content (and instantiate the same mental type) in order to be redundant. In most contexts, for instance, *your* neural event representing a rapidly approaching truck will hardly be redundant simply because *I* already have neural event with the same content. Psychological redundancy seems to occur within a mental system, where what constitutes a single mental system seems itself to be a matter of causal organization.

It makes sense to speak of the structure of the brain (and perhaps even of some of the contents of the brain) as to some extent redundant, assuming that the brain is a functional system characterized by a certain functional organization and so forth. Thus if one area of the brain is damaged, another area can functionally compensate for its loss.⁴¹

But those who advocate a conscious and mental duality model for split-brain subjects believe that the two “disconnected” hemispheres are distinct mental systems, and that right hemisphere neural events are therefore not psychologically redundant.

41. Though even here there may not be exact redundancy in many instances; sometimes one area of the brain functionally compensates for the damaged area, at least to some extent, without performing the exact same operations. We should be careful not to individuate mental functions and capacities too grossly. (See the criticisms of the STR position in Section 5.3.)

Indeed, it seems fairly simple to show that right and left hemisphere neural events aren't really redundant in the split-brain subject. If, for example, something catastrophic suddenly happened to S's left hemisphere visual system, his right hemisphere visual system would not be able to compensate for it, nor would his right hemisphere visual representations be able to serve as "back-up" copies of the representations he's just lost, at least for many cognitive purposes. For instance, S wouldn't be able to *describe* (via his left hemisphere) his right hemisphere experiences.

This seems to suggest that attempting to formulate the STR position narrowly, as stating that the causal independence of two neural events is psychologically irrelevant so long as the events in question are redundant from a psychological perspective, doesn't really change the main issue. For what *qualifies* as genuine redundancy *itself* seems to depend upon facts about causal interaction and independence.

But the defender of STR might press that at most left and right hemisphere neural events aren't redundant in the split-brain subject because the two hemispheres aren't functionally identical, for instance, and thus right and left hemisphere neural events won't have truly identical functional roles. But if a subject *did* possess two genuinely psychologically redundant neural events, couldn't *they* jointly realize a single mental token—no matter the manner of their causal connection to each other?

4.3.4 Functional type and other psychological properties

It is worth emphasizing again that the two hemispheres of a split-brain subject surely do not have “redundant” psychological properties. The hemispheres have somewhat different patterns of perceptual access to the body and world; they have differing access to the visual field, for example, and to tactile information from the hands. But even if the hemispheres had identical perceptual access to the world, the right and left hemispheres are not functionally identical; they appear to process auditory (e.g. verbal) and visuospatial information differently, for example, and therefore surely generate mental states of different types (where content is included in type). Moreover, minds, certainly minds like ours, are so complex that, even if the hemispheres had identical patterns of perceptual access to the world, and even if they were functionally identical, it is highly unlikely that they would really be subject to all the same types of mental state (where content is included in type) at every moment (absent some kind of interhemispheric mental interaction to ensure this, that is).⁴²

Furthermore, while a description of an event’s mental type and content tells us a lot about its psychological properties, it doesn’t tell us everything. Even in a case in which the two hemispheres of a split-brain subject each generated neural event of the type “belief that X”, the belief might well be put to different uses in each hemisphere. Of course, the two uses would still have to have a lot in common in order for both states to remain *beliefs* (with the same content); they would have to share a single *core* or *central* causal role. Still, there are many things—probably indefinitely many

42. As Gazzaniga once put it, “it is unlikely that the two independent mental systems (each with its own sensory input, processing and storage mechanisms, and motor output) would maintain equivalent attentional and motivational states over an extended period” (1978: 117). The same goes for other types of states as well.

things—that can be done with a given mental state, even if we keep its functional type and content constant.

The redundancy condition—the condition that two neural events must be truly redundant with respect to their psychological properties if they are to constitute a single mental token via causally independent processes—will be very difficult to meet. Certainly it is not met by the two hemispheres of split-brain subjects. Thus there will frequently be a way to distinguish between even two type- and content-identical mental tokens, even within a single creature, *by attending to their psychological properties alone*, i.e. by looking at the other (instantiated) types of mental states they interact with.

The STR position may still be founded on an important principle, however. This principle says that, whatever criteria we use to individuate neural events, when it comes time to individuate *mental* events or tokens, we should use *purely psychological properties* to do so. Thus if two neural events truly do have indistinguishable psychological properties, a psychological theory need not distinguish them, but may rather identify them with a single mental token.

The next subsection of this chapter considers whether this is correct, or whether certain physical properties, in particular properties concerning the causal relationship a neural event bears to other neural events, is always an essential part of a neural event's psychological identity. It does this by conceiving of that creature for whom the case of singularity-through-redundancy could be made most strongly. I argue that even in such a case, we should reject that characterization of that creature, and thus accept a method of individuating mental tokens which is *always* sensitive to

the causal properties of those tokens. It follows that in the rare case in which a purely psychological and a partly neural approach to individuating minds might yield different counts, the answer yielded by the partly neural approach is the better one.

4.4 Mental Tokens

I begin by describing the creature for whom the case of singularity-through-redundancy could be made most strongly. I nonetheless argue that the STR characterization would be inappropriate even in this case, for a method of individuation which is sensitive to the causal properties of realizers is required by those of us who are realists about functionalist explanation.

4.4.1 Perfect parallelism: the best case for singularity-through-redundancy

We have seen that the principle of multiple realizability provides no support for the claim that causally unrelated neural events may constitute a single mental token. We have noted that the neural redundancy, sensory decomposition and binding that exist in the “normal” brain are not analogous to the kinds of redundancy and causal independence the STR model of split-brain subjects’ consciousness concerns, and therefore not relevant to evaluating that model. Can anything else be said in favor of STR—or can anything be said more decisively against it?

In fact I think that something more can be said in favor of the position, though I also think that ultimately, it still fails. To evaluate the position at its strongest and most plausible, consider a hypothetical case of what I will call “perfect parallelism.” Our “perfectly parallel” subject, PP, unlike our split-brain subject S, has two *functionally identical* hemispheres that operate *wholly* independently of each other (they interact neither at the cortical level nor indeed via any wholly neural route).

Unlike those of our split-brain subject, PP's two hemispheres operate in *perfectly* "redundant" or parallel fashion: every time PP's left hemisphere generates a neural event realizing the belief that X, PP's right hemisphere does also, and at the exact same time, but via a wholly independent causal processes. Every time PP's right hemisphere generates neural event realizing the desire for Y, PP's left hemisphere does too, at the exact same time, wholly independently. And every time one of PP's hemispheres generates a motor plan, and initiates a motor impulse to, say, grab Z, PP's other hemisphere does also, at the exact same time and again wholly independently, such that the set of neural events in either hemisphere alone would have sufficed for that exact same action having been performed. How many beliefs that X does PP have? How many streams of consciousness? How many minds?

Perfect parallelism is clearly a fiction, but one that may suggest that those sympathetic to the STR claim still have an important *theoretical* point against proponents of the duality model, who have tended to see neural facts and properties as *in and of themselves* highly relevant to the individuating of mental tokens. The defender of STR would presumably press, here, that while PP's neuroanatomy may be striking and fascinating from some neuroscientific standpoint, the fact that PP's beliefs, desires, experiences, intentions, memories, motor commands, and so forth, are realized "disjunctively" by *two* causally independent sets of neural events, is now without question psychologically *irrelevant*. For surely we *lose nothing*, from a psychological perspective, by attributing to our genuinely perfectly parallel subject, PP, only a single belief that X, a single desire that Y, and so forth, at every moment. And surely we *gain* nothing, from a psychological perspective, by attributing to PP

two sets of mental tokens with indistinguishable psychological properties at every moment. Occam's razor alone would suggest that we attribute the less extravagant number of mental tokens to PP. In which case the singularity-through-redundancy position, even if not tenable with respect to counting minds or streams of consciousness in split-brain subjects for purely empirical reasons, is premised upon a sound method of individuating mental tokens: in virtue of their purely psychological properties, and *not* their neural ones. In which case many members of the neuropsychological community have made a significant error.

4.4.2 Functionalism, physicalism, and the realist commitment

Defenders of STR might at this point propose that we accept the following principle: that at any point in time, a given creature cannot have two or more mental tokens with indistinguishable psychological properties. Should functionalists accept this principle?

In the abstract, functionalism doesn't say much about the nature of mental tokens; functionalism proper is neutral with respect to what sorts of things can occupy the roles tagged by mental state terms, and is therefore in principle compatible even with ontological dualism. Of course the vast majority of functionalists are physicalists, who believe that all existing phenomena are actually physical (or have physical properties). Some functionalists—such as Rey (1994) and Tye and, so far as I can tell, Marks—are also *realists*. To commit to realism about functionalist explanation is to believe that there *really are things* occupying the roles defined in functionalist theories, that mental tokens of functionally defined types actually exist. The functionalist who is also a realist thus has an additional reason for being a physicalist. Recall from Chapter Two the point that causation just seems in fact to

require physical stuff (Wiggins, 1976; also Rey, 1977); identifying mental tokens, the realizers of mental types, with physical things, offers a way that functionalist explanations can be genuinely *causal*.

As I mentioned in Chapter Two, this work is obviously written from the standpoint of a functionalism with realist commitments: it assumes that emotions and experiences and so forth are the types of things that are *actually* realized by physical (I am assuming neural) events, and is concerned with the identity relations between mental types and neural tokens. For the functionalist who commits to realism, neuroscience has the potential to vindicate a psychofunctionalist theory, to show that the theory is not just empirically adequate or useful with respect to explaining and predicting behavior, but that the story the theory tells is moreover causal and correct. Of course the realist runs greater risks as well; neuroscientific discoveries can falsify the realist's best functionalist theory, in a way that they can't falsify the best theory of the non-realist. The functionalist who is a realist expects and in fact requires that at multiple points in the development of a psychological theory, neuroscientists will hunt for neural things playing one of the psychological roles the functionalist has described. If the functionalist is lucky, neuroscientists will find such things. If she is less lucky, such things won't be found, and she will either have to repeatedly revise her theory or (perhaps, ultimately, following enough failures of the right sort) drop the theory altogether. And sometimes neuroscientists may not only find *something* playing the role her theory described—they may find more of those things than she expected.

From the perspective of the functionalist who is also a realist, causal relationships to (instantiated) *types* of mental state are essential to a neural event's mental type and arguably token identity. But are causal relationships to instantiated *types* of mental event all that matter to a neural event's (mental) *token* identity? Or are causal relationships to *particular mental tokens* also essential to a neural event's (mental) *token* identity?

It seems quite clear that relationships to *particular* tokens matter. Clearly, that one of your neural events and one of my neural events are identical with respect to the *types* of mental state to which they are causally related provides no reason for identifying these two neural events with a unique mental token. For the two neural events are causally related to distinct actual occupiers of these mental types. The point can be put in the following way. A mental token's *type* identity is a matter of its bearing certain relationships to *any* mental tokens of *particular types*. But a mental token's *token* identity is matter of it bearing certain relationships to *particular* mental tokens. The relationships that matter are of course causal and counterfactual ones: a realist functionalism, again, is distinguished by a commitment to the reality, to the causal efficacy, of the entities and activities described by (some) functionalist psychological theory.

Although functionalism is a theory about the types of things mental states are, not just mental types but mental tokens are of course familiar and important in functionalist psychology. Think about what it is to have an episodic memory of having done X, for example; there must be an actual causal connection of a certain sort between the doing of X and your memory of having done it. Thus my twin on

twin earth doesn't remember my mother, but not because she's suffering from amnesia. She may have lots and lots of memories of someone exactly *like* my mother. But she has absolutely no memories of *my* mother, and therefore her memories are not mine.

Now, how any of these mental tokens are individuated will be in part a physical matter. A mental token is an *occupier* of a given functionally-defined role; counting tokens is a matter of counting *how many occupiers of that role* there are. Occupiers are physical things; physical properties are relevant to determining how many such things exist. Admittedly, the physical, causal properties that an occupier of a psychological role possesses may not *all* be relevant to its individuation as mental token; indeed, a great number of them (e.g., the number of dendritic spines synapsed upon) are no doubt irrelevant. For that matter, not all of a neural event's *psychologically defined, actual* causes and effects may matter to its identity qua mental token. A token percept that gives rise to a certain belief, for instance, might have been that *same* percept even if it *hadn't* given rise to that belief. So not *all* of the causal relationships that a mental token *does* and *does not* bear to other mental tokens need matter to its token identity. But certain *counterfactual* relationships, for instance, the causal relationships a mental token *can* and *can not* bear to other role-occupiers, is clearly the sort of physical, functional property that matters to its *token* identity.⁴³

At least, it matters from the perspective of a functionalism committed to the view that mental tokens are real causal actors.

43. See also Rey (1977) for an argument that the identity of a mental token hinges upon its causal relationship to other particular mental tokens (and not just any mental tokens of a particular type).

I submit that the realist must concede that the perfectly parallel subject PP has two sets of mental tokens with indistinguishable psychological properties at every moment. Granted, by offering a complete characterization of the mental architecture of one of PP's hemispheres—a characterization given in terms of interactions between the various *types* of mental states that hemisphere contains tokens of—you will have offered a complete characterization of the mental architecture of PP's other hemisphere as well. Granted therefore that to offer a complete characterization of the structure of PP's mental life as a whole, you need not refer to even a single further (kind of) psychological *property*, once you have described all the psychological properties associated with one of PP's hemispheres. Nevertheless, a characterization of the structure of PP's cognitive life would not be complete without the acknowledgment that there are in fact *two* of every psychological token and interaction heretofore mentioned in this characterization: two of every type of mental state, two of every type of psychological process, two streams of consciousness, and two minds.

Granted, it would by stipulation be impossible to see any evidence of these *two* sets of tokens with indistinguishable psychological properties in PP's behavior. (Except counterfactually, through manipulations, e.g. by functionally incapacitating one of PP's hemisphere and noting that PP's behavior doesn't change.) Granted a psychological theory for PP would therefore never have to mention the existence of *two* sets of mental states with indistinguishable mental properties in order to be empirically adequate with respect to even a single one of PP's perfectly unified behaviors. Instrumentalists, therefore, need not recognize two such sets. But as

Ramsey, Stich and Garon (1990) point out, the realist commitment means that fit with behavioral data isn't all that is relevant to evaluating a functionalist story. Whether that story *correctly* identifies the events that actually caused the behavior is also relevant.

Defenders of the “singularity-through-redundancy” view might worry, however, that by attributing distinct mental tokens to PP's two hemispheres, we actually *lose* explanatory adequacy with respect to behavior. For now having a single stream of consciousness can't be an explanation for unified behavior, and having multiple streams of consciousness can't be an explanation for disunified behavior. But this misses the point that perfect parallelism would be extraordinary. To the degree that an overlap in conscious contents *ever* plays a role in producing unified behavior, then we should *expect* two streams of consciousness with identical contents to produce unified behavior. It is just that it is highly unlikely that two such independently-generated streams will ever actually exist.

Even if split-brain subjects had two hemispheres associated with highly similar or in fact identical conscious contents, this would not mean that such subjects had single streams of consciousness. Mental tokens are causal actors. Causally distinct mental actors—i.e. neural events that have psychological properties and that are causally independent in the strong sense described above in 3.3.2 and 3.3.3, or in any stronger sense—are distinct mental tokens. Well motivated functionalist principles give no reason to think otherwise. In fact, while functionalism says little about individuating tokens, a functionalism with realist commitments should use some physical properties to individuate mental tokens. Doing so respects the role that

mental state types and contents play in functionalist explanations, *and* the status of mental tokens as causal actors.

5 *On Integrating Theory and Practice*

A major theme of this chapter so far has been the tension between neural and behavioral data regarding split-brain subjects—or, indeed, whether there *is* such a tension. In Section Two, Hurley (2003, 1998, 1994) argued that neuroanatomical facts about split-brain subjects don't provide any objective evidence for mental duality, because (this is ultimately what her argument seemed to reduce to) we don't have any principled reason for privileging *neural* over other physical (bodily, environmental) phenomena as constitutive of the mental. In Section Three I showed that the unified or integrated character of split-brain subjects' daily behavior cannot play as simple or decisive a role in supporting a mental singularity model as Marks (1981) and Tye (2003) seem to suggest. And in Section Four I argued that neuroanatomical and neurofunctional facts about split-brain subjects play *more* of a role in supporting the mental duality model than Marks and Tye are willing to acknowledge.

In this section I pause to examine *why* this theme, that of a tension between neural and behavioral evidence, has emerged several times in the discussion of this chapter. More specifically, I wish to explore whether there might be some underlying motivations for the “one mind” model of split-brain subjects that I reject, since I don't think that the criteria for individuating minds that Marks and Tye, at least, offer to support such a model, are in and of themselves convincing or even *prima facie*

appealing. Might there be something *else*, some deeper motivation, driving them towards “one mind” conclusion?

Recall that in my view, while the integrated nature of split-brain subjects’ behavior does provide some compelling evidence for the “one mind” model of “split-brain” cognitive architecture, philosophers such as Marks (1981) and Tye (2003) have still drawn *too* strong an evidential link between split-brain subjects’ day-to-day behavior and the mental singularity model. I see two likely and distinct motivations for their having done so. It is possible on the one hand to read Marks (1981) and Tye (2003)—less so Nagel (1979)—as having fallen prey to verificationism about the mental. Alternatively, these philosophers might believe that the “two minds” or mental duality model for split-brain subjects is incompatible with the way that we do or would understand and relate to these subjects in practical, social, legal, moral, and personal contexts. Nagel (1979), for one, expresses quite explicitly his concern about our ability to integrate theory and practice with respect to such subjects. In fact he seems to despair of our ultimately achieving any way to integrate our scientific and practical understandings of *any* human subjects. Perhaps Marks and Tye are motivated to defend the proposition that split-brain subjects have one mind apiece because they think doing so is the only way to avoid Nagel’s despair.⁴⁴

5.1 Entities and Evidence

Verificationism is a theory of meaning, according to which the meaning of a proposition, including or in particular a scientific proposition, is equivalent to the

⁴⁴ Hurley’s (2003, 1998, 1994) arguments, which I also discuss in this chapter and which similarly concern the evidential links between mental structure, brains, and behavior, are I think differently motivated, and that motivation requires separate discussion, which I undertake in Chapter Six.

evidence for that proposition. This theory had close ties to logical positivism, of which *behaviorism* can be seen as an extension. Behaviorism was a theoretical approach in psychology that said, first of all, that the goal of psychology was to explain behavior, rather than to explain the mind, and that also said that psychological *explanation* should consist, in some sense, of statements about behavior (and behavioral dispositions and environmental stimuli). So for the behaviorist, behavior constitutes the proper explanandum, and the proper explanans, in psychology.

The sense in which psychological explanation should consist of behavioral statements varies according to the form of behaviorism in question. *Radical* behaviorism proposed an entirely non-mentalistic explanation of behavior and was thus eliminativist with respect to mental states; psychological explanations should consist entirely in statements about behavior and environment. (Radical behaviorists such as Skinner believed that the *cause* of all behavior was directly or indirectly environmental—mental states and processes were not necessary to explain any behavior.) *Analytic behaviorism* was not eliminativist about mental states, but claimed that a mental state is just a behavioral disposition.

Methodological behaviorism, meanwhile, was the view that scientific psychological procedures and explanations should be concerned with behavior, and not, for instance, with neurophysiological data. Rey (1997) has suggested that methodological behaviorism lives on in the form of what he calls *superficialism*, the view that “for every mental process there is some or other *piece of outward behavior* that would, in a particular context, be criterial (or *decisive*) of its presence” (1997: 197; emphasis added). As Rey notes, some have defended superficialism explicitly;

Dennett for instance admits: “I unhesitatingly endorse the claim that, necessarily, if two organisms are behaviorally exactly alike, they are psychologically exactly alike” (Dennett 1993: 922).

Marks (1981) and Tye (2003) are not behaviorists or superficialists—at least, not explicitly or consistently. While some forms of functionalism incorporate elements of behaviorism, Marks and Tye both appear to be psychofunctionalists, who believe that the best (functional) analysis of mental states will come from a developed psychological theory (or theories). This developed psychological theory, it seems clear at this point, won’t be a behavioristic one.

But nonetheless some of their arguments for mental singularity in the split-brain subject do recall superficialism. Marks for instance claims that if a developed psychological theory explains the integrated behavior of “normal” subjects in part in terms of mental and conscious singularity, there will be great pressure on such a theory to explain split-brain subjects’ behavior *in the same way*, though note that he does not go so far as Dennett as to say that similar or identical behavior absolutely *requires* identical explanation:

Since, apart from the experiments, the split-brain patients exhibit the same degree of behavioral integration we do, one would like to explain it in the same way. . . . This line, as opposed to treating each brain-half as a psychological subject, has the advantage of making the integrated action of both split-brain patients and us depend upon the same thing. . . The appeal to disunity. . . will be limited to those spots *where it does some genuine work*, namely, to where it explains a lack of behavioral integration. So there will be considerable pressure on the psychology to give a one-mind account of split-brain patients. (Marks 1981: 41-42; emphasis added)

Or recall Tye saying:

What leads to the supposition that split-brain patients have a disunified consciousness is their failure to behave in an integrated, coherent way in certain, special experimental situations. But *if behavior is the evidence* on which the hypothesis of disunity rests, then the fact that split-brain patients behave in an integrated way at other times supports the hypothesis that their consciousness is generally unified.

(Tye 2003: 126; emphasis added)

One interpretation of these passages is that they reflect a confusion between the *nature* of a thing and the *evidence* we have of its nature, or between the metaphysical and the epistemic. There is no *equivalence* between disunified behavior and mental duality on the one hand, and unified behavior and mental singularity on the other, such that every time a subject behaves in a disunified fashion she must have two minds at that moment, and every time she behaves in a unified fashion she must have one mind at that moment. The behavior of a split-brain subject *during* the split-brain experiment provides evidence concerning the deep structure of that subject's mental life—a deep structure that is explained (in part) by the structure of the subject's brain, something that doesn't *change* even when her *behavior* changes.

Behavior that looks “single-minded” may provide evidence, perhaps compelling evidence, perhaps *very* compelling evidence of *actual* single-mindedness. But it does not *decisively* support that conclusion. Neither would behavior that looked disunified provide decisive evidence of mental duality. The degree of “disunified” behavior split-brain subjects exhibit during the split-brain experiment only provides compelling evidence of conscious or mental duality *in conjunction with hypotheses* about *how* and *why* that behavior was elicited under *those* circumstances. If, for instance, a certain class of subjects had seemed to exhibit the behavior characterizing

the split-brain phenomenon, but had no neuroanatomical abnormalities, and, moreover, if they exhibited this behavior not *specifically* during the split-brain experiments but rather when participating in *any* kind of psychological experiment whatsoever, the explanation for their behavior might not have been mental duality at all. Perhaps they would have been dubbed mentally ill, or simply uncooperative. The important point is, again, that mental duality (or “disunity”) doesn’t *reduce* to disunified behavior, nor vice versa, and mental singularity (or “unity”) doesn’t reduce to unified behavior, nor vice versa.

For whatever reason—or probably for numerous reasons—the pull of verificationism often seems harder to resist in psychology than in other sciences. This is why Putnam’s (1975) refutations of verificationism using non-psychological examples, such as that of *disease*, remain useful. Medical professionals first become acquainted with a disease, X, through encountering the cluster of symptoms, W, that X normally produces. Indeed it may take quite some time before a particular cluster of symptoms, W, is hypothesized to be *a cluster* of symptoms, signifying the presence of a *particular* disease. And then it might still turn out that W *isn’t* really a cluster, or that it is a cluster signifying the presence of one of a number of diseases, or signifying either disease X or allergy Y, and so forth. Meanwhile, other people may acquire X without exhibiting W at all.

The concept of a *disease*, in other words, is that of an underlying *cause* of the symptoms that first provided evidence that the disease existed. But the disease is not *equivalent* to the symptoms it (normally) causes. And similarly, Putnam continued, the concept of *pain* is not that of “the presence of a cluster of [stereotypical pain]

responses, but rather the presence of an event or condition that normally causes those responses” (1975: 330). Pain is what (normally) produces pain responses. Pain is not the cause of pain (pain is not *touching* a hot stove, for instance, even though touching a hot stove does normally cause pain), and pain is not what pain causes. Pain is that which is normally (but not necessarily) caused by e.g. touching a hot stove, and that which normally (but not necessarily) causes e.g. wincing. And we tell similar stories for other mental entities.

Including minds. The perfectly parallel subjects we considered in the previous section were in some respects not too far removed from Putnam’s “super-super-Spartans. (Who, for ideological reasons, give *no* behavioral sign—not via report, not by wincing, or by screaming, or by clenching their fists, nothing—of experiencing the pain that they do nonetheless feel just as vividly as we feel ours.) Lacking the normal behavioral evidence for mental and conscious duality, we nonetheless concluded that perfectly parallel subjects had two minds and two streams of consciousness because we had other evidence—neural evidence—that convinced us of their mental duality. Except that we are, if anything, in an even better position to conclude that perfectly parallel subjects have two minds and two streams of consciousness, than we would be to conclude that super-super-Spartans feel pain. For one thing, we had built into the thought experiment that we *already had* a good psychological theory for perfectly parallel subjects, one that gave us confidence that we knew the mental character of all their neural goings-on. And second, we are in a better position, I think, to describe the sorts of behavior that pain *normally* causes than we are to describe the sorts of behavior that having a single stream of consciousness, for instance, normally causes.

For (this is at least part of the explanation for this difference) we do (often) know when we're in pain, and therefore can really observe the sorts of behavior that we and other people engage in when we're in it. But, I would submit, we don't *know* that we—by which I mean we human animals, and not necessarily we subjects of experience—have a single stream of consciousness. It's possible that "normal" human subjects have two streams of consciousness as well, and therefore that much of the behavior we've thought resulted from having a single stream of consciousness really didn't. And even if we do have single minds and single streams of consciousness, and even if we know that we do, I still don't think that we know as much about the typical behavioral effects of having a single mind and a single stream of consciousness, as we know about the typical behavioral effects of being in great pain.

Here I suppose Nagel (and perhaps Marks and Tye also) might object that my borrowed analogy linking mental duality to pain, or to polio, gets me in trouble. If polio is the thing responsible for the cluster of symptoms associated with polio, why can't mental singularity be *whatever* normally causes unified behavior? We could just *stipulate*, that is, that whatever the mental causes of integrated behavior, it shall constitute a single mind. But this is what we *don't* do with diseases. While the medical community might *hypothesize* that there is *a single disease* responsible for symptom-cluster W, this is a *hypothesis*, one that they are willing to revise in light of further discoveries about the cause, or *causes*, of W. It could have turned out that there were *two, distinct* diseases (or one disease and one allergic reaction) *both* responsible for the cluster of symptoms associated with polio, for instance, which the

medical community would not claim are now both a single disease by prior stipulation. And similarly it could turn out that there are *two* minds responsible for my integrated behavior.

This is all of course familiar, but the verificationist turn that Putnam (1975) sought to bury may still haunt some corners in psychology (and philosophy of psychology). So perhaps verificationism is the best explanation for Marks' (1981) and Tye's (2003) (less so Nagel's (1979)) repeated appeal to split-brain subjects' integrated behavior to support a mental and conscious singularity model. The appeal makes sense up to a point, for integrated behavior no doubt does provide *some* compelling evidence for some kind of mental integration or at least coordination. (The right kind or degree of which *would* produce mental singularity, in my account.) But the appeal can only be taken so far. Unified or integrated behavior is, again, *defeasible* evidence for mental and conscious singularity, just as a perfectly calm demeanor is compelling but defeasible evidence for the absence of great pain.

This dissertation tries to defend the positions that split-brain subjects have two minds and two streams of consciousness. But the defenses of those particular positions are of course tentative. We don't yet have an adequate, developed psychological theory or set of theories, and so we can't yet say with certainty whether the kinds of causal interactions between right and left hemispheres in split-brain subjects isn't of the intra-mind sort, in part because we don't know the full extent of interhemispheric interaction in split-brain subjects, but in large part also because we don't yet know with certainty what the "right" kind of causal interaction—interaction that suffices to integrate what would otherwise be two minds into one—really is. A

lot of what I can do at this point is just provide *some* evidence for supposing that there *is* a distinction between interhemispheric and intra-hemispheric interaction in split-brain subjects, and that this distinction is *psychologically relevant*, relevant to identifying those hemispheres with minds.

The “how many minds?” question, applied to split-brain subjects (or for that matter to “normal” subjects), is like the question about whether a particular subject or group of subjects who perhaps aren’t showing any *symptoms* nonetheless all suffer from a common *disease*. What we cannot yet answer with certainty, but what we want to *know*, is whether there is some kind of unifying causal process, or some kind of unifying or integrating system of causal interactions, that’s absent between the hemispheres, though present *within* them, in split-brain subjects. And, if the answer to *that* question is yes, what we want to know is whether this causal process is somehow explanatorily *deep*, and thus so important that its presence *trumps* the absence of superficial behavioral “symptoms” of duality. The answer to one or both of these questions might be *no*, and the “one mind” model might be correct. But *verificationism* cannot lead us to the right answer.

5.2 Minds and Persons

But there may be something else going on in Marks (1981) and in Tye (2003) and in Nagel (1979) that is responsible for the (in my view) exaggerated emphasis they place upon split-brain subjects’ grossly characterized, day-to-day behavior, in theorizing about their mental structure. A second interpretation of these philosophers is that they are simply concerned about how on earth to reconcile a scientific model of split-brain subjects’ minds as dual, with our pre-theoretic, social, legal, moral, personal, and

practical ways of understanding and relating to those subjects. For we would naturally respond to such subjects, we would naturally relate to them, as if they were single persons, with single minds, and with a single unified consciousness, and so forth. We might call this the problem of integrating *theory* and *practice*.

Tye, for instance, emphasizes that “Split-brain patients *typically* act as if they are single persons with a unified conception of the world and what is going on around them” (2003: 115; original emphasis)—in contrast, he says, to subjects with dissociative identity disorder (DID), or what used to be called “multiple personality disorder.” A split-brain subject—with a discernible and seemingly normal personality, temperament, and range of behavior types—*unlike* a DID subject—is not the sweetest, most unassuming, if a bit shy, person you’ve ever met on one day, and a loud-mouthed, arrogant, obnoxiously aggressive person the next. And Marks, too, believes that a subject with DID is a better candidate for possessing multiple minds than a split-brain subject is:

The successive sets of propositional attitudes we use to explain her [a DID subject’s] State 1 behavior have a familiar kind of internal coherence, and they follow each other, with various changes and overlaps, in the way we expect; and those for State 2 do so as well. Each, separately, looks like part of a typical psychological history of a single person. But if we attempt to combine them, that is, take the successive sets of propositional attitudes in their true temporal order, we can maintain the internal coherence of the individual sets on each side of a transition between State 1 and State 2 only by making it unintelligible how they could be parts of the psychological history of a single person. (Marks 1981: 27)

Marks makes this worry about integrating theory and practice particularly *personal* when he points out that it is conceivable that technologically superior alien

life forms performed outwardly undetectable, immediate-recovery laser-callosotomies on *all* “normal” human subjects last year, and that we simply haven’t discovered it yet, because the vast majority of us have never been subjects in a split-brain experiment. “For my part,” he says, “I find it scarcely conceivable that my individuality depends upon the falsity of this surgical fantasy, or upon any facts about how I might behave in . . . [a split-brain] experiment” (Marks 1981: 11). Note the last clause in particular: Marks is barely able to concede that his *own* behavior under conditions of perceptual lateralization could provide compelling scientific evidence that *he* has two minds (or at least, that he isn’t an individual); how can he then in good faith be any less reluctant to ascribe mental duality (or at least, dual identity) to split-brain subjects?

The worry about integrating theory and practice is most explicit in Nagel:

It seems strange to suggest that we are not in a position to ascribe all those experiences to the same person, just because of some peculiarities about how the integration is achieved. The people who *know* these patients find it natural to relate to them as single individuals.

(Nagel 1979: 159; original emphasis)

Even more troublingly, for Nagel as for Marks, what we say about split-brain subjects has implications for what we say about “normal” subjects as well. Split-brain subjects are just like us in this important respect, Nagel says: they have two hemispheres, each of which *could* support a mind in the other’s absence; as it is, though, those hemispheres don’t function entirely independently of each other. They interact with each other in various ways, and if they don’t interact with each other in

all the ways and to the extent that ours do, this just makes our *degree* of “mental unity” greater than theirs. . . . But not absolute. “Even if we analyze the idea of unity in terms of functional integration, therefore”—as I have suggested we do—“the unity of our own consciousness may be less clear than we had supposed” (Nagel 1979: 164).

And while:

The concept of a person might possibly survive an application to cases which require us to speak of two or more persons in one body. . . . it seems strongly committed to some form of whole number countability. Since even this seems open to doubt, it is possible that the ordinary, simple idea of a single person will come to seem quaint some day, when the complexities of the human control system become clearer and we become less certain that there is anything very important that we are *one* of. But it is also possible that we shall be unable to abandon the idea no matter what we discover.
(Nagel 1979: 164)

Nagel, in other words, ultimately suggests that no integration of theory and practice is possible. Either one way of understanding split-brain subjects (and “normal” subjects as well) will have to bow to another—or neither will yield, nor admit of integration with the other. The scientific worldview, in other words, and the non-scientific worldview, will both survive. But they will describe different worlds.

Of course the general problem of integrating scientific with other modes of understanding is a deep one, and one that many people, not just philosophers and not just scientists, may be increasingly familiar with. And that the problem exists in this particular instance will probably be clear to anyone familiar with the split-brain phenomenon. One of the things that makes the phenomenon so puzzling and so fascinating is, of course, the disconnect between their ordinary, every day, and their experimentally-evoked behavior, and Nagel is (as usual I think) insightful in casting

this disconnect as one between, again, scientific and other (legal, social, moral, etc.) modes of understanding human beings. Fortunately, I think that the theory-practice problem in *this* instance is less serious than in other, more familiar cases embodying this problem, such as the free will vs. determinism case.⁴⁵

This is because the mental duality model, or at least some plausible versions of it (including the one to which I would ascribe), understands mental duality in such a way that we can attribute two minds to a split-brain subject and still rationally and in some sense correctly—not just pragmatically, in other words—view that subject as an individual, probably in most of the ways in which our pre-theoretic understanding of human beings suggests that we should. Thus there is, I believe, less of a conflict between the scientific or theoretic and the non-scientific or pre-theoretic image of human beings in this case, than there is in the free will case, for instance.

The conflict between theory and practice is not so great in this instance for three reasons. First, the mental duality model of split-brain subjects needn't attribute mutually incoherent psychological states to the two hemispheres—so we are free to view those subjects' mental lives as coherent and thus unified in *some* sense, just as Marks and Tye and Nagel do. Second, there are plausible accounts of personal identity according to which a single person can have two minds, at least as some plausible versions of the mental duality model understand mental duality. And third,

⁴⁵ The problem of integrating scientific and non-scientific understanding of human beings with respect to free will is that science (or much of science, anyway) presents us with a deterministic universe. Yet some sort of vision of other people and of ourselves as freely acting agents is important not just in social contexts in general, and not just in spontaneous, personal interactions with other people, but also for making (and justifying) moral judgments. It's important in legal contexts—to some extent the concept is even *codified* in law!

Perhaps another example of this is that between the scientific, value-free universe, and the pre-theoretic, moral universe: morality, religion, society, law, teach that some deaths are unjust, killings wrong; but scientists discover only the deaths, only the killings, only the *causes* of death and of killing, only the *beliefs* about right and wrong; science doesn't discover wrongness or injustice itself.

while the two hemispheres of a split-brain subject don't stand in the sort of functional relations to each other that they would need to in order to constitute a single mind, on my view, the relations that they stand in to the subject's body and to the subject's world make it so that even if the subject were somehow constituted by two persons, these would be persons who should justly and rationally be treated as one (at least in most instances). These three claims are related, but let me discuss each of them separately.

Marks and Tye believe that we have more reason for ascribing multiple minds to a DID than to a split-brain subject, because only the former subject is one to whom it is tempting or potentially necessary to ascribe incompatible personalities, incompatible character traits and values and goals, and so forth. But in most instances we aren't tempted to ascribe incompatible personalities or characters and so forth to a split-brain subject at all. So whence the motivation for attributing multiple minds to a split-brain subject?

But the mental duality model is obviously not offered in terms of incoherence or incompatibility between mental states or mental systems. The two minds of a split-brain subject may be very similar. If my arguments concerning the perfectly parallel subject are correct, two minds can be not just perfectly coherent but qualitatively identical and *still* be two minds. So the fact that split-brain subjects seem, most of the time, like "normal" individuals with "coherent" mental lives is not in any conflict with the mental duality model. The motivation for the mental duality model lies

elsewhere. What makes it *rational* to relate to a split-brain subject as a “single individual” may, perhaps, depend upon the overall coherence of her mental states.⁴⁶

Second, the passages quoted above in this subsection also exaggerate the seriousness of the problem of integrating theory and practice with respect to split-brain subjects by conflating seeing or treating a subject as (constituted by) multiple *persons*, and attributing multiple *minds* to that subject. But individuating minds and personal identity are, I think, distinct issues. This is one of the reasons why I have tried to steer clear of personal identity questions (and why I have felt that I *could* do so), but let me just briefly explain why, when we understand the meaning of the (scientific) attribution of mental duality to these subjects, we see that it is not necessarily incompatible with regarding such subjects as single persons.

While both personal and mental singularity may be a matter of integration, as Nagel says, being a single person may require a different kind of integration, and/or integration of different things, than being a single mind requires. For instance, in the

⁴⁶ To be clear, I am not actually convinced that coherence and unity play as significant a role in governing our understanding of and interactions with other people that Marks and Tye and Nagel seem to believe. There’s no doubt that certain kinds of unity play a significant role in such interactions; I expect that if I tell someone something important, he’ll remember it the next day, for instance. (Although I trust I’m not the only one who’s been disappointed in the past.) But it seems to me that we beholders tolerate, if barely, a fair degree of what our eyes take to be inconsistency, incoherence, etc., in other people. And even in extreme cases in which we *don’t* tolerate it, cases in which a subject’s apparent incoherence, irrationality, inconsistency, or disunity, baffles or disturbs us greatly, I have further doubts about whether, in these cases, we *actually* resort, in a meaningful way, to treating such a subject as multiple individuals. (Do jurors accept the insanity plea of a plaintiff diagnosed with DID because they *really* think his body is occupied by multiple persons, only some of whom can be justly sent to prison? Or do they just think the guy’s *nuts*?) Of course we would *treat* very differently (and understand differently) the angry, aggressive, potentially violent “alter” of a subject whose other “alter” is fearful, retiring, gentle. But responding to an angry person very differently from how one would respond to a fearful person, whether they’re the same person (at different times) or not, is just basic social competence. Still, since Marks and Tye and Nagel, and perhaps many other philosophers, are all convinced of the importance of seeing other people as unified and coherent agents, I don’t push this point. Here I just claim that coherence and unity may enjoy a significance in non-scientific ways of understanding *people* that they don’t enjoy in scientific ways of understanding *minds*; there thus seems to be no conflict between the two modes of understanding, however; in contrast to the free will case where there is at least a *prima facie* conflict between free action and determinism (even if *compatibilism* shows the conflict to be merely apparent).

next section, I am going to argue that the generally integrated character of split-brain subjects' behavior results less from interhemispheric *integration* of mental states and processes and more from interhemispheric *coordination* of output (and input). To take one example, imagine that, if right and left hemisphere simultaneously issue competing and conflicting motor plans, the *stronger* issuance wins, and *that* action and not the other is initiated. (Though within some very short time period, presumably, the other hemisphere could generate an even stronger "vetoing" response, checking or reversing the other hemisphere's motor plan.) This could be modeled as (friendly) competition between two still distinct mental systems (or even as a form of cooperation between the two distinct mental systems). And, admittedly, it *could* perhaps be modeled as a form of competition (or perhaps cooperation) between persons as well. But this could also be easily modeled as intra-personal competition between or coordination of desires, as well. If "part of" Carter, say, wants to leave the server's tip right where it is, but "part of" him wants to just go ahead and grab it, and if his desire to take cash is *stronger* than his desire to leave it where it is, and if his *stronger* desire is the one that actually "prevails" over his behavior, yielding an efficacious motor plan that results in him stealing someone else's tip. . . . Many people, probably most people, would probably say that we *can* hold Carter responsible for stealing someone's wage, and that we can view him as a single person. And, interestingly, it seems as if there are least some instances in which we could hold Carter to be a single person, morally culpable for this theft, even if Carter is a split-brain subject and the competing desires were located in distinct hemispheres. (Admittedly the details could matter here.) Of course Carter is an

ambivalent thief, and there may be better and worse angels of his nature. . . . But this is mundane, really. Perhaps the right thing in the end would be to say that Carter, the whole organism, wanted to be good (RH) and he wanted to be bad (LH) and in the end he especially wanted to be bad (LH).

This isn't to say that we *should definitely* see Carter as a single person, responsible for stealing his server's tip, in the example I just sketched; for one thing, the sketch is too bare, and anyway, much deeper thought about personal identity, not to mention moral responsibility, would be called for. All I hope to have suggested is that accepting the mental duality model, properly understood, for split-brain subjects, leaves room for negotiation and maneuver where constructing an account of personal identity, even a psychological account of personal identity, is concerned. The mental duality model could very well be compatible with a wide range of positions on matters that are of great personal and social and moral and legal and practical importance to us.

I said above that even though being a single person probably is in part a matter of integration, just as being a single mind is also a matter of integration, the kind of integration that matters may differ between the two sorts of cases. But it is also arguable that personal unity or individuality depends less upon mental integration than does mental singularity or one mindedness. Personal individuality might depend more upon other factors—perhaps factors like unity and coherence, for instance, as Marks and Tye believe. The psychological similarity between a split-brain subject's right and left hemisphere could have more relevance to that subject's personal identity than to how many minds he has. And I suspect that history matters

more to personhood than to mindhood as well; the fact that the two hemispheres of a split-brain subject have a similar history, that is, may be of greater importance to the subjects' *personal* identity than it is to individuating their minds. For the two hemispheres *do* share a history, including a social history, and a history of acting together. (Not necessarily acting *as one*, mentally speaking, but they've jointly contributed to many of the same actions. And personhood is I think more closely tied to notions of agency than is mindhood, which would make sense, given the different domains to which the concepts of persons and minds are belong.) And they share not just a past, of course, but also a future, because they remain inside the same body. Indeed regardless of how little integration there is between right and left hemisphere mental processes, there is still just one *life* to be led, one set of things to experience, if not one shared set of experiences, and one past to be remembered, if not one shared set of memories.

The shared history of the two hemispheres provides another reason why it is not just understandable but rational for the loved ones of a split-brain subject to continue to relate to the subject as a single person after his callosotomy. It isn't as if, after their only son's callosotomy, S's parents are suddenly confronted with two strangers, Rightie (RH) and Leftie (LH). They have known Rightie and Leftie for all of S's life: S has *always* been constituted in part by both Rightie and Leftie. Now granted Rightie and Leftie may now bear a different relationship to each other, and may now be associated with distinct minds. (For that matter, maybe S's right and left hemispheres were *always* associated with distinct minds.) But all the same elements are there, and no doubt S's parents related to him previously as a single person—their

one son. (After all, while his brain and his two hemispheres are an important part of him, to say the least, his parents might well balk if you stated that they were the *entirety* of him.)

Finally, even if S *were* associated with two distinct persons, Rightie and Leftie, it is not clear how one could *treat* S as two persons, in any way that would significantly diverge from treating S as one person. S isn't like a pair of conjoined siblings, joined at the hip, each of whom can at least do *some* different things, can have *some* different relationship with a unique marital partner. S isn't like a DID subject, whose body is occupied by different persons (if that is what you think about DID) at different *times*. There is no need to say things twice, to S, and no way to have a relationship with Rightie that you don't have with Leftie; there is little way, on a day-to-day, practical basis, to *do* something to Rightie that you don't do to Leftie.

If S's two hemispheres had very different characters, and each initiated behaviors that differed, morally, from that of the other, and if neither hemisphere could act quickly to *stop* the other hemisphere's immoral behavior, then there might be some difficulties about, say, imprisoning S for a crime he'd committed. But I don't think this is a genuine worry, first, and second even if we did somehow have good evidence that Rightie really wanted to stop Leftie from stealing the car but was just too functionally depressed or something, we wouldn't, I don't think, resort to treating S like two persons. We would again at most have to treat him as someone who wasn't fully culpable for his actions, because he (S, that is, not Rightie or Leftie) wasn't capable of sifting through his various motivations in the normal way, didn't have his

full range of beliefs to be used in the same reasoning process, or couldn't use some of his desires to guide his action.

I don't mean to say that these few quick and rough things that I've said here on the subject of personal identity are uncontroversial, or that they can definitely be used to develop the best account of persons. I do hope to have made plausible my belief that accepting a model of a split-brain subject's cognitive architecture according to which she has two minds still leaves *room* for ascribing those minds to a single person. In some cases multiple minds may bear relationships to each other that allow them to jointly constitute (or constitute in part) a single person—without bearing the relationships to each other that would make them into a single mind, from the perspective of a scientific psychology. For while scientific findings will probably contribute to our understanding of persons and personhood in general, in all of us, the important interests we have in persons are not equivalent to or exhausted by the scientific interests we have in minds.

6 Duality versus Partial Co-mentality

In this section I turn to a final objection to the “two minds” conclusion for split-brain subjects: the objection posed by the existence of subcortical structures connecting the two *cortically* separated cerebral hemispheres in split-brain subjects. That the two hemispheres of a split-brain subject remain connected via structures which figure in cognition and experience might in fact seem to call into question the very meaning and import of the “how many minds?” question. For their existence, especially in

combination with the isomorphism thesis, may show that a split-brain subject has an indeterminate number of minds.

6.1 Partial Co-Mentality (Or, an Indeterminate Number of Minds?)

In the previous chapter I concluded that a question can escape being “merely verbal,” if it can be shown to be an empirical question, or an interesting conceptual question. (Some such conceptual questions have a role to play in scientific theories, though as I suggested in Chapter Two, there is no reason to think that *all* interesting conceptual questions do.) A merely verbal question, however, can only be answered by arbitrary *decision*.

I think that the concept of a *mind* does have a role to play in a psychological theory, or at least in a functionalist psychological theory—a role which is both important and very subtle. This role relates to the relationship between type-identifying mental states and individuating token minds. Within the functionalist framework, the types of interactions a mental state can and can’t engage in—the nature of its causal relationships to other tokens of particular mental types—is in fact a large part of what makes that mental state the type of mental state that it is. In saying that a belief is *unconscious*, you say something about the systematic relations it bears to other mental states of various types, something about its causal role in a larger psychological economy. The relationship between mental state type and mental interactions within the computational-representational theory of mind means that one of the theory’s most basic forms of explanation is what we might call explanation-by-functional-type.

One of the things a psychological theory for a particular species tells us is which types of mental state that members of the species have and the different ways those different types of mental state can interact with each other. A psychological theory might tell us that beliefs and desires can interact with each other without the mediation of perceptual states, for instance, or that the activation of stored episodic memories, in contrast to the activation of stored semantic memories, consists in part of the generation of mental images. But these general rules—rules that in part define what it *is* to be a belief, an episodic memory—are only meant to apply *within* a mind; the inability of a belief and a desire to interact absent a perceptual process doesn't even constitute an *exception* to the general rule that beliefs and desires can, *ceteris paribus*, interact absent a perceptual process, if the belief and the desire are each located in a distinct mind; the rules are defined over a mind, but not across them. This is, I think, part of the work that minds do within a psychological theory, though they do this work subtly and implicitly.

The boundaries of a mind are functional but have theoretical and explanatory significance. To mark something as a mind is to mark it as a system with the kind of structure that requires psychological explanation. But also, the functional relationships between mental tokens *within* a mind require a different sort of psychological explanation than do functional relationships between mental tokens *across* minds. The “how many minds?” question is then a question in part about where and when we apply our psychological theories, a question about carving up the domain of psychological explanation, and perhaps a question about the range of psychological laws as well.

A purely psychological and a partly neural approach to individuating minds should tend to yield the same answer to the “how many minds?” question for any particular creature, given, among other things, the necessity of providing the neural with a psychological taxonomy before using neural facts to individuate minds in the first place. I have argued, however, that for the rare creature for whom the purely psychological and the partly neural criteria for individuating minds may yield a different answer, the correct answer is that yielded by the partly neural method of individuating minds. These claims both support and are supported by the isomorphism thesis, according to which cognitive architecture and the neural architecture underlying it are isomorphic.

But while the isomorphism thesis provides some reason to reject the “one mind” claim for split-brain subjects, it does not necessarily support a “two minds” conclusion for such subjects. In fact it might be claimed that the isomorphism thesis supports the claim that split-brain subjects represent a case intermediate between having one mind and having two. Let the term “co-mentality” refer to the relationship mental tokens bear to each other when they belong to the same mind. Following Lockwood (1989), we might reason as follows. (Lockwood argued that the neural architecture of split-brain subjects affords them a *partially unified consciousness*. His model of split-brain subjects’ consciousness will be examined in Chapter Five.) First, physical connectedness comes in degrees. (Imagine an island composed of two small mountains and a valley in between, and now imagine the ocean levels rising continually, such that this island *gradually* becomes two islands.) Second, if mental structure is isomorphic to neural structure, then co-mentality may come in degrees

also. What would it mean for co-mentality to come in degrees? It is difficult to make this idea precise without an account of what functional relationships specific types of mental states should bear to each other when they *do* belong to a single mind, but let us continue to try to use co-mentality as a placeholder for *whatever* kind or kinds of relationships two mental states bear to each other when they belong to the same mind, and say that, for a given set of mental states, their degree of co-mentality is absolute when every mental state in that set is co-mental with every other state in that set. Their degree of co-mentality is zero when no mental state in that set is co-mental with any other state in that set, and intermediate degrees of co-mentality fall in between these two extremes.

Perhaps split-brain subjects have neither a single mind within which all mental states are co-mental with each other, nor two wholly discrete minds between which no two mental states are co-mental with each other. For while the two hemispheres of a split-brain subject are divided at the cortical level, they remain connected to each other via their mutual connectedness to subcortical and other non-cortical structures. These non-cortical structures meanwhile make complex and essential contributions to mental operations. So perhaps, in a split-brain subject, there are right and left hemisphere mental states that are not co-mental with each other, but subcortically-borne mental states which are co-mental with both right and left hemisphere states. Therefore a split-brain subject has neither one perfectly “unified” mind nor two wholly “discrete” minds.

But accepting that co-mentality can come in degrees might call into question the meaning and import of the “how many minds?” debate itself. If there can be cases

intermediate between having one mind and having two minds, then the “how many minds?” question may not be answerable—not in whole number terms—absent making some kind of arbitrary decision, at least for split-brain subjects. So why agonize over the “how many minds?” question?

This worry, that the possibility of partial co-mentality challenges the entire import of the “how many minds?” debate for split-brain subjects, is itself easily dealt with. *Prima facie*, at least, cases of partial co-mentality seem possible. There seems no reason to insist that minds *must* be wholly discrete from each other, in the relevant (intra-mind) sense. Conjoined twins who partially shared a cerebral lobe might not have wholly discrete minds, for instance. Isomorphism between mind and brain may give reason to think that multiple minds need not be wholly discrete. But this doesn’t make it a waste of time to wonder, in any particular instance, whether co-mentality holds or does not hold at all between multiple mental systems. And as, in general, two things need not be wholly discrete or distinct from each other in order to be two things, there could still be reason to wonder whether a split-brain subject is more helpfully and accurately conceived as having two minds or one mind.

So *do* the two “disconnected” hemispheres of a split-brain subject constitute a single mind, between which only partial co-mentality holds? It is certainly true that, due in very large part to their mutual connectedness to non-cortical (neural) structures, the two hemispheres of a split-brain subjects are psychologically unified in some special ways. In the rest of this section however I explain why I nonetheless believe that the “two minds” or mental duality model is still the best model of split-brain subjects’ cognitive architectures: sub- and non-cortical structures are essentially

involved in cognitive operations, but in a way that is more consistent with mental lateralization or bilateralization than with mental unification; non-cortically-mediated interactions between the two hemispheres serve more as a means of coordinating the activities of the two hemispheres than they do as a means of integrating their mental processing.

At least so I believe at the present time. Ultimately, whether sub- and non-cortical structures do compromise a mental duality model or better support a partial co-mentality model in part depends, of course, on what those structures are doing, and on their place in the overall mental architecture associated with a split-brain subject's brain. But we don't *know* these things yet. So everything I say in this section will of necessity be very tentative. I speak in broad strokes in part for that reason, and in part because I am interested in the philosophical question of whether, assuming that non-cortical structures are essentially involved in cognitive processes, and assuming they mediate some form of causal interaction between the two hemispheres, the two hemispheres *cannot* be associated with distinct minds. I argue that they still can.

6.2 Subcortical Structures: Plausible Sources of (Interhemispheric) Mental Integration?

This subsection describes introduces some important terms and players (neuroanatomical structures) in the debate about “subcortically mediated” interhemispheric interaction in split-brain subjects, to give some sense of what this debate is about. In this subsection I deal with some empirical and terminological issues and challenges to the debate; subsequent subsections turn to the more philosophically relevant conceptual and theoretical issues.

There are a few different ways of classifying cortical and non-cortical neuroanatomy, but I will use “cerebral hemispheres” to refer to the cerebral cortex, the basal ganglia, and the cortical white matter (including the corpus callosum). I will use “cerebral cortex” to refer to the frontal, parietal, occipital, and temporal lobes, and also to the olfactory cortex and hippocampus. I will also classify as *cortical* the corpus callosum, the anterior commissure, and also the hippocampal commissure. By “subcortical structures” I mean to refer to the thalamus and hypothalamus, and to the striatum, and to the amygdala. (There are, again, other ways of classifying some of these structures, but I think this way makes best sense of the split-brain literature on cortical and subcortical structures.) See Figure 3.1.

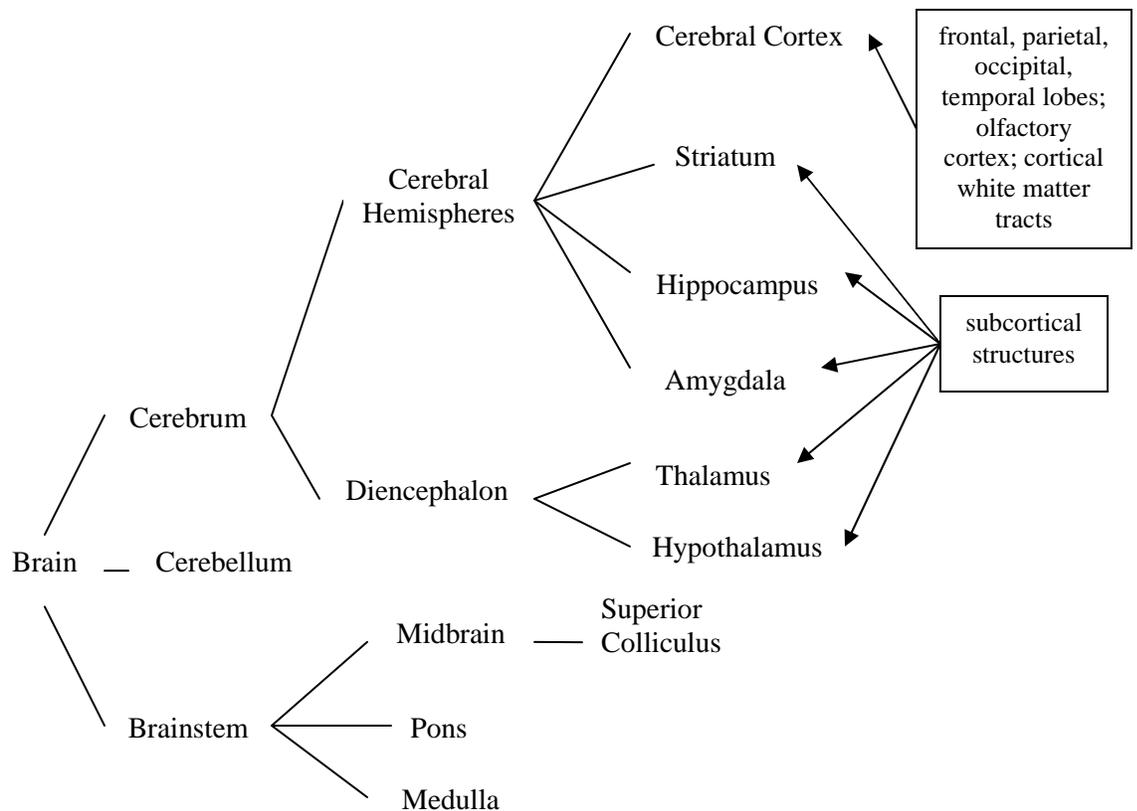


Figure 3.1: Gross Anatomical Divisions of the Brain

When neuropsychologists (and philosophers) refer to *subcortical structures* as a possible means of interhemispheric mental unity in split-brain subjects, to which structures can they be referring? To begin to answer this question we need to know which structures communicate across the midline of the brain, either via commissures or by virtue of being midline structures (rather than bilateral, paired structures) to begin with.

In addition to the corpus callosum (and the optic chiasm, at which information from the nasal visual field from both eyes crosses to be carried to the hemisphere contralateral to the eye from which it originates), the brain contains a number of other crossing pathways: the anterior commissure and the posterior commissure, mentioned just above, and the massa intermedia, the hippocampal commissure and fornix, and the intertectal commissure.

Note in addition that at every level of the *spinal cord* there are both gray and white matter commissures carrying sensory and motor information. On the motor side—the white matter commissures of the spinal cord are contributed to by the corticospinal tract—these play a role in applying motor commands potentially initiated by only a single hemisphere to muscles on both sides of the body. Despite these commissures, the two hemispheres of a split-brain subject often don't have strictly speaking identical motor control over the two sides of the body.

In the brainstem, most pathways ascending from the spinal cord run contralateral to the side from which they originate, and most pathways descending to the spinal cord run contralateral to the side on which they terminate. The situation

with cranial nerves and their pathways is much more complex and difficult to sum up. Pathways from cortex to cranial nerves and vice versa tend to be bilateral (Mosenthal, 1995), but with a mixture of crossed and uncrossed (contralateral and ipsilateral) pathways, and not every nerve shows this pattern. Significantly ascending cranial nerve V, however, which carries sensory information concerning the face, projects bilaterally (Gazzaniga, 1970) which is why a split-brain subject can cross-localize for points on the face and neck.

Finally the brain contains some midline structures, most notably, for our purposes, the cerebellum. The lateral hemispheres of the cerebellum (the cerebrotocerebellum) send and receive “crossed” information: the right side of the cerebellum generally receives information concerned with the right side of the body, that is, and sends its output to the left hemisphere (which is also primarily concerned with the right side of the body). Thus right cerebellar lesions can produce right side hemineglect (perhaps at least in part due to the right lateral cerebellum’s connectivity with the left hemisphere).⁴⁷ Also note, though, that the *vermis* or central portion of the cerebellum, between the two cerebellar hemispheres, is concerned with movements of the trunk. Thus right and left side body maps may be coordinated via their mutual connectedness to central cerebellar body maps.

Many bilateral, paired, non-cortical structures are not *directly* connected to each other; there is no commissure of the amygdala, for instance, and no striatal commissure. Of course, even absent any commissural connections between, say, right and left amygdala, these structures could still interact. In fact so long as both are

⁴⁷ Silveri, M., Misciagna, S., Terrezza, G. 2001. Right side neglect in right cerebellar lesion. *Journal of Neurology Neurosurgery and Psychiatry* 71: 114-117.

functioning they are bound to *interact* at *some* point or other; everything in the brain will interact with everything else in the brain via *some*, possibly very long, very convoluted, causal chain. But then, again, I'm sure every structure in my brain has interacted at *some* point with every structure in my sister's brain; we're concerned here with questions of causal intimacy.⁴⁸ We have prima facie reason to *suspect* (though of course not to be certain) that two brain structures connected directly by some neural pathway are in some sort of important functional relationship with each other—a more intimate functional relationship than two structures that are only very indirectly connected, in part because direct pathways of course allow for faster communication, and speed matters in the brain. So let us consider the most promising pathways—the most direct routes between one hemisphere and the other—in split-brain subjects.

Of these, several don't look at this point to play any significant role in the interhemispheric integration of mental states and processes. The massa intermedia connects right and left thalamus, a structure that is arguably so inextricably linked with cortical areas in its functioning that it could be considered a cortical structure, *functionally*. But so far as I know there is very little communication between right and left thalamus, and whether the massa intermedia really has any functional significance is unclear, especially since it is apparently absent in a very significant portion (about 20%) of the human population. (It is noteworthy that the massa intermedia isn't necessarily a commissure per se but rather just a fused group of

⁴⁸ One strong way of expressing this point would be to say that we're concerned with which causal chains constitute natural (functional) kinds in psychology. And in particular with which causal chains constitute the natural kind of an *intra-mind interaction*. But this assumes that there is such a kind, which of course there might turn out not to be.

midline thalamic nuclei, and thus we might not expect it to serve a “unifying” role.) The functional significance of posterior commissural fibers is also not clear, though they may be involved somehow with pupillary light reflexes. (The functional significance of the habenular commissure, which I did not mention earlier, is also not clear, and I have never heard the habenular nucleus specifically implicated in the debate about subcortical communication in split-brain subjects.) There is also a habenular commissure connecting right and left habenulae—limbic structures. It is possible that this commissure serves some sort of role integrating or coordinating the states of the two hemispheres with respect to the sorts of activities the habenular nuclei are involved in, including regulation of sleep-wake cycles, pain processing, and reproductive and nutritive behavior. (The habenular nuclei output to midbrain tegmental areas and to the pineal gland.) I have not read anything in the split-brain literature implicating it in any function (much less interhemispheric function), however.

This leaves, as commissures likely to serve as significant routes of interhemispheric interaction, in addition to the corpus callosum, the anterior commissure, the hippocampal commissure, and the intertectal commissure. Sectioning the posterior callosum, however, as is done in a full callosotomy, necessitates sectioning the hippocampal commissure as well (Gazzaniga 2000). In a complete *commissurotomy*, meanwhile, the anterior commissure also sectioned. Thus, depending on the extent of the surgery, there are fewer or greater opportunities for cortical commissural communication between the hemispheres.

The anterior commissure “derives its fibers from the temporal lobe and from subcortical ‘limbic’ structures, in particular, the amygdala, and projects to the same regions in the other hemisphere” (Gazzaniga and LeDoux, 1978: 153). Gazzaniga and LeDoux accordingly speculate that the anterior commissure may play some role in the interhemispheric transfer of affect in callosotomized (though not commissurotomized) subjects. Others raise doubts, however; Doty writes that “the amygdala, as the principal organizer of emotional reactions, is decidedly unilateral in its anatomical relations and electrophysiological manifestations” (1989: 70)). He drew this conclusion in part after performing unilateral lesion studies on non-human primates; at least one study using rats found that anxiety could be increased by stimulating the right amygdala, but not by stimulating the left (Adamec and Morgan, 1994), suggesting an absence of at least left-to-right amygdala transfer. Studies like these are more difficult to run using human subjects for obvious reasons, but Damasio, Adolphs, and Damasio (2003) theorize that one should expect even greater degrees of lateralization of emotion in human subjects. (Actually there is a good deal of evidence for the lateralization of emotions in human and in some other mammalian subjects, but since cortical areas are significantly involved in emotional circuits as well, it is difficult to determine how lateralized subcortical emotional circuits are from the mere fact that emotion is lateralized in the hemispheres.) Since I don’t know how to rule out that the anterior commissure plays some role in the interhemispheric transfer of affect in some split-brain subjects I will assume that it might. If the anterior commissure *does* serve as a means of interhemispheric communication, then subjects who have undergone full callosotomy and those who have undergone

complete commissurotomy (and certainly those have undergone only partial callosotomy) may constitute importantly different cases, however, where comparing the mental duality and the partial co-mentality models is concerned. Still I will continue to speak generally of “split-brain subjects” and only distinguish between (full) callosotomy and commissurotomy cases if it seems particularly relevant.

The split-brain literature on interhemispheric subcortical interaction is unfortunately slightly more difficult to sort through than it probably has to be. One difficulty is terminological. The term “subcortical” is used sometimes to refer only to non-cortical components of the cerebral hemispheres: the thalamus and hypothalamus, the striatum (or basal ganglia), and the amygdala. In some of the literature on subcortically-mediated interaction in the split-brain subject, however, the term “subcortical” is used to refer more broadly to any *non-cortical* brain structure. (E.g., Handy, Gazzaniga, and Ivry (2003) write: “Evidence that a putative internal clock may be linked to *subcortical mechanisms* has come from a variety of sources. This work has focused on the *cerebellum* and basal ganglia, two structures that have extensive reciprocal connections with the cerebral cortex” (1461; emphasis added). Or, Corballis and Forster (2003) write of “a subcortical visual route, presumably via the *superior colliculus*” (337).) I prefer to use the term “subcortical” more strictly, to refer to non-cortical components of the cerebral hemispheres, and to use the broader term “non-cortical” to refer to non-cortical neural structures: to midbrain structures, to the cerebellum, and in some instances even to the cord. But this terminological issue makes it difficult to tell even the *range* of structures that neuropsychologists

(and philosophers) intend to implicate when they speak of possible “subcortical” interactions in split-brain subjects.

For much of the literature on split-brain subjects uses the term “subcortical” imprecisely in a further way also. Works referring to “subcortical” interhemispheric interaction in split-brain subjects often don’t specify, or even speculate as to, *which* sub- or non-cortical structures are mediating the interaction in question. Granted, the reason for this is understandable: many times, authors no doubt simply don’t know which non-cortical structures are sustaining the interaction in question. But this is still a significant problem, for two reasons.

First and more obviously, “subcortical” and “non-cortical” are neuroanatomical and not functional terms, and subcortical structures, no less than cortical and non-cortical structures, are not helpfully spoken of as a unitary ensemble. (In fact even *one* particular subcortical structure, the thalamus, contains at least four *functionally* distinct nuclei group, one concerned with arousal and awakesness and attention, and one limbic, one motor, and one sensory.) It is informative to be told that subcortical structures serve a function no more frequently than it is informative to be told that the cerebral cortex serves a function. Meanwhile many *subcortical* structures are not, *prima facie* at least, even plausible candidates for mediating interhemispheric interaction (except indirectly), since again many of these paired structures are not *linked* across the midline functionally or even physically, and others are often sectioned in the course of a “split-brain” surgery.

The intertectal commissure, between right and left superior colliculus, is an exception; it is functionally significant in human subjects, and indeed my own sense

is that a stronger case can be made for interhemispheric mental (in particular motor-perceptual and attentional) integration via the superior colliculus than for interhemispheric mental integration via any other non-cortical structure. See Savazzi et al. (2007) for discussion. (Note that the superior colliculus is a midbrain structure, and not properly speaking a subcortical structure. Note also that the superior colliculus contributes some fibers to the posterior commissure.)

I am sure that the authors who refer to subcortically-mediated interhemispheric interactions in split-brain subjects, without implicating *particular* sub- (or non-) cortical structures, would agree that, ideally, we would all know more about which *particular* sub- or non-cortical structures mediate such interactions. Still, they would say, referring instead simply to “subcortical structures” as a group is at worst non-ideally vague. But, these authors would presumably continue, in their own defense, *some* form of integration is achieved by split-brain subjects in a number of behavioral tasks, suggesting that some form of interhemispheric *mental* integration has been achieved. And since cortically-mediated routes to such integration have (essentially) been blocked in the split-brain subject, the integration exhibited *must* be proceeding via sub- or non-cortical routes. This justifies referring broadly to subcortical (or non-cortical) sources of mental integration, even though ideally, again, one would be able to narrow down which structures in particular are responsible for the integration.

But there is a second, more significant, problem with sweeping references to “subcortical” sources of (interhemispheric) mental integration (especially when “subcortical” is used broadly to include non-cortical structures like the superior

colliculus or the cerebellum). Some such references may be at worst not just vague, and therefore not as helpful or informative as they would ideally be, but actually misleading: they may refer to sources of interhemispheric mental integration that don't actually *exist*.

For not all integrated-seeming or unified behavior, even behavior that is the product of both hemispheres' *participation*, need be the product of the hemispheres *interacting with each other*. And not all communication between the hemispheres amounts to interhemispheric *integration* of mental states and processes. So if you observe integrated behavior in a split-brain subject, or even if you conclude that some form of interhemispheric communication has occurred, you *can't* infer that some form of non-cortically mediated, interhemispheric mental integration has occurred. Instead, unified behavior may be the result of some sub- or non-cortically mediated coordination of the hemispheres' inputs and outputs. (For instance, Savazzi et al. speculate that "it might well be that after callosotomy, input to the extrastriate area in one hemisphere could provide descending input to the pontine nuclei bilaterally through brainstem commissures and then to the contralateral cerebellar cortex which finally, in connection with the motor cortex on the same side, will carry out the response" (2007: 2424). Depending on the form that this "connection" between the ipsilateral cerebellar and motor cortex takes, no interaction between the two hemispheres per se need be involved.) Meanwhile, this coordination of interhemispheric inputs and outputs may or may not itself be a form of *non-cortical* mental integration, depending upon the form *it* takes. And, again, communication between the hemispheres may occur via a route which is not characterizable as a form

of intra-mind interaction. (Note that interhemispheric communication between the two hemispheres of a split-brain subject will indeed always or almost always be sub- or non-cortically-mediated. But this would be true even if the two hemispheres never interacted except via a route that involved behavior. This is because of the brain's "vertical" functional organization: both perception and action involve non-cortical structures and pathways.)

Knowing which non-cortical structures were mediating interhemispheric interaction (assuming we also knew some things about the functional properties of the non-cortical structures in question) would *help* determine the psychological nature of that interaction. It would help determine whether the function those structures were performing in fact serves as a form of mental integration: as a form of intra-mind interaction. There may be some cases in which one can read off the behavioral data and *know* that some form of interhemispheric interaction has occurred; if a split-brain subject, S, can describe (LH) the form of a visual stimulus presented only in her LVF (RH), for instance, this suggests interhemispheric interaction. (Though even in this case, we would need some way of knowing that RH processing was involved, for it seems possible that instead visual information was sent to the right superior colliculus and then directly transferred via the intertectal commissure to the left superior colliculus and then ultimately to the left hemisphere, without any right hemisphere involvement. Holtzman (1983) raises this issue.) But in many instances, behavior that appears to involve both hemispheres' participation really might not. Or it might involve both hemispheres' participation, but not their interacting (directly or indirectly) with each other; rather, their participation might consist in their having

joint effects on some downstream, non-cortical processing. Or it might involve both hemispheres' interacting, but still without entailing any intra-mind integration. (To take an *obvious* example, suppose that that you present a photograph of a church to S's LVF (RH) and she says (LH) "It's a church." Clearly interhemispheric interaction has occurred. But suppose that it occurred via this route: S *drew* a picture of a church with her left hand (RH), and brought her sketch of the church into her RVF (LH).)

This is all just prefatory, to urge caution and clarity (as much as is possible) in discussions about subcortically- (or non-cortically-) mediated interhemispheric interactions in split-brain subjects. In the rest of this section, I suggest that on the whole sub- and non-cortical structures do not serve as a means of integrating what would otherwise be two minds into one, and that the "two minds" or mental duality model for split-brain subjects is the most useful and accurate.

6.3 What Kind of Interhemispheric Relations Do Non-cortical Structures Afford?

Non-cortical, including subcortical, structures obviously play a complex and important though admittedly not yet well understood role in coordinating the activities of the two hemispheres: in perception and emotion, in attention and action. That they play such a role, however, does not mean that they integrate what would otherwise perhaps be two minds into one. While it is not possible to draw any decisive conclusions at this point about whether and to what extent non-cortical structures serve as mechanisms for inter-hemispheric co-mentality, this section presents some reason to think that non-cortical operations are consistent with a mental duality model for split-brain subjects.

6.3.1 Non-cortical representations

Non-cortical structures utilize representations: representations of temporal duration, for instance, and of visually perceived stimuli, and of motor effectors. Also, and perhaps most famously, the split-brain cases illustrated that non-cortical structures either play some crucial role in generating, or permit the interhemispheric transfer of, *emotional* or *affective* information (or both). (Non-cortical structures may also generate and/or transfer what Sperry, Zaidel, and Zaidel (1978) referred to as general mental *aura*, though the anterior commissure may also be implicated here.)

Yet even in these cases, it is difficult to say whether the two “disconnected” hemispheres *share* the same set of mental tokens, in virtue their mutual connectedness to intact non-cortical structures in which these tokens are located, or whether they instead both have a non-cortically-mediated ability to access or to generate distinct tokens of some of the same of mental state. For while, on the one hand, some kind of affective information appears able to travel, via *some* route, into one hemisphere from the other, the hemispheres *do* distinct things with these states.

In one famous example, a frightening film (about fire safety) was played only in V.P.’s *left* visual field (right hemisphere). After viewing, V.P. said (LH) that she didn’t know what she saw—“I think just a white flash”, she said, and, when prompted further, “Maybe just some trees, red trees like in the fall.” When asked by her examiner (Michael Gazzaniga) whether she felt anything watching the film, she replied (LH), “I don’t really know why but I’m kind of scared. I feel jumpy. I think maybe I don’t like this room, or maybe it’s you. You’re getting me nervous.” Turning to an assistant, she said, “I know I like Dr. Gazzaniga, but right now I’m scared of him for some reason.” (Gazzaniga 1985: 76-77).

In another case, Schiffer et al. (1998) elicited different answers to questions about an unhappy childhood experience from the two hemispheres of two split-brain subjects. The experimenters put two sets of pegs, five in each set, in front of each patient but obscured from his vision, where the leftmost peg in each set represented “none” and the rightmost represented “extreme,” and then required the subjects to answer questions like, “How much did this experience upset you at the time?” using both hands at once. (Via his right hemisphere one subject expressed a higher degree of disturbance by these memories than he did via his left. The second subject’s right hemisphere expressed both a more positive self-image, and more negative emotions, e.g. greater loneliness.) While the results of this second study could be interpreted as showing that the hemispheres *don’t* experience even the same affective states, despite subcortical intactness, it is equally possible that the hemispheres *did* experience the same affective states, and yet, again, the hemispheres *responded* differently to these states. (E.g, the first subject’s left hemisphere might have “played it cool” about being bullied in childhood; the second subject’s left hemisphere might have deemed it socially inappropriate to express an overly strong self-image or admit to overly negative emotions.)

Meanwhile, the route by which the affective information travels between the hemispheres is also of course highly relevant, and at this point I think still unclear, at least in many cases. Consider three ways that affective information could hypothetically have carried from V.P.’s right to her left hemisphere, in the example cited above. (These are obviously not the only three possible routes, and indeed I do not even mean to suggest that the routes I’ll describe are consistent with existing

evidence concerning the generation of affect and emotion; the examples I give are just to illustrate that the particular neural route taken may have consequences for how we describe the psychological interaction between right and left hemisphere.) One possibility is that her right hemisphere visual system in conjunction with her right amygdala generated an affective representation, a representation of a negative affective state. (There is disagreement about the content of affective and emotional states and even whether and to what extent such states are representational; the nature of such states is of course relevant to current concerns, but is also beyond our scope, so I am deliberately vague, here.) Then V.P.'s left hemisphere gained access to *this* token representation via her anterior commissure. (I believe that the anterior commissure was not sectioned in V.P., but in any event for this example assume that it wasn't.) So V.P. said that she felt a little jumpy.

A second possibility is that V.P.'s right hemisphere visual system in conjunction with her right amygdala generated an affective representation, a representation of a negative affective state. Then, via V.P.'s anterior commissure, the content of this representation was communicated to V.P.'s left amygdala, which formed a representation with this same content. Her left hemisphere received this left amygdala affective representation (perhaps via her left thalamus) and so V.P. said that she felt a little jumpy.

A third possibility is that V.P.'s right hemisphere visual system in conjunction with her right amygdala and hypothalamus resulted in a particular physiological state of arousal: slightly elevated heart and breathing rate, sudden muscle tension, and so forth. Representations of this physiological state were made accessible to V.P.'s left

hemisphere via ascending spinal cord tracts, and perhaps via left-side subcortical limbic structures, and so V.P. said that she felt a little jumpy.

Obviously, the former possibility looks more like a genuine case of subcortically-mediated interhemispheric mental unity than the latter. And this is so even though subcortical structures (the amygdala, the hypothalamus), and loops of activity from RH to subcortical structures to LH, characterize *all three* possible routes of affective transfer. What differs between the three possible routes is *where* the transfer *from right to left* occurs, whether the hemispheres share any mental tokens, and (perhaps most crucially) whether the transfer involves some *non-representational* step.

Picking apart whether and when the hemispheres share mental tokens, and what sorts of mental tokens they share, is admittedly very difficult. Careful study may be required in order to determine the exact *form* that non-cortically-mediated, interhemispheric coordination takes. For example, it may appear that the hemispheres are communicating via shared representations even when they are not. Consider an interesting study by Naikar (1996) concerning the observation that split-brain subjects can often perceive apparent motion across the midline. When two dots are presented very near to each other and in quick succession, subjects, split-brain and “normal,” tend to perceive a single dot *moving* from the location of the first dot to the location of the second. Interestingly, split-brain subjects still experience this phenomenon even when the first dot is presented to one hemisphere and the second dot to the other hemisphere. In order to correctly indicate whether the motion was from left to right or from right to left, then, it would appear as if at least one hemisphere would have to

have access to *both* representations—to the representation of the first dot and to the representation of the second dot. Moreover, there was a plausible candidate mediating this apparently interhemispheric motion perception ability—the superior colliculus. For it was already known that striate cortex lesions, at least in macaques, does not eliminate motion sensitivity in higher visual regions (MT—often called the equivalent of human V5), while lesioning the superior colliculus *on top* of striate cortex does eliminate motion sensitive responses in MT. Thus, the superior colliculus is apparently capable of sustaining direction sensitivity in MT in the absence of input to that same area from the striate cortex.

At the same time, however, neurons in the superior colliculus are not *themselves* direction-sensitive (and any degree of direction sensitivity in the pulvinar is eliminated by lesioning of the striate cortex). Thus, Naikar could conclude that, “visual input reaching MT by the tectopulvinar route must be nonselective, and MT must have the capacity to *produce* directionally selective responses from nonselective input” (Naikar 1996: 1042, emphasis mine). This is an example in which subcortical structures allow for the cortical generation of a representation they themselves do not seem to bear.

Naikar ultimately explained the observed phenomenon not in terms of shared access to the same set of representations, but rather to attentional shift, hypothesizing that what allows split-brain subjects to distinguish the direction of (apparent) motion across the midline were non-cortically-mediated *shifts of attention* from the location of the *first* dot to the location of the *second* dot. The shift of attention from, say, left visual field to right visual field, was something that either hemisphere could (in a

sense) perceive or experience, and from this could *infer* that the motion was from left to right. This hypothesis was supported, in the following way: the two dots in this experiment were each of a different color—color being a property that the superior colliculus does not appear to represent at all—and subjects were asked to indicate the color of “the” dot. On those trials for which motion was reported, subjects gave the color of the *second* dot—which is what one would expect if the hemisphere receiving the *second* dot (the hemisphere to which attention had been shifted) were indicating the color. On those trials for which motion was *not* reported—as if attention had never shifted from the first dot, and from the hemisphere that had received that dot—the color of the *first* dot was given. Naikar concludes that “there are two. . . processes that can mediate motion perception; a subcortical, involuntary process that is insensitive to object features, and a cortical, voluntary process that tracks visible stimulus features” (1996: 1048).

Now, admittedly, *some* kind of representation must be used *by the superior colliculus* in shifting attention from one hemisphere’s visual field to the other. Indeed it has been known since at least the 1970’s that the superior colliculus allows split-brain subjects to verbally describe (LH) left visual field (RH—and right superior colliculus) visual stimuli. (See Trevarthen and Sperry, 1973.)⁴⁹ This midbrain-

⁴⁹ This is actually a fascinating paper in part because Trevarthen and Sperry detail some of the verbal and other behavioral responses of their subjects to the suite of perceptual tasks (presentation of LVF (RH) stimuli, or bilateral presentation of visual stimuli, where the stimuli either differed or did not differ in form or in motion between the two fields. One gets a sense not just of the important individual differences between subjects but of the fragility of bilateral cooperation at least when verbal response is required. N.G., who , “became aphasic and distressed in attempting to respond to double questions such as ‘what kind of movement? and which way did it move?’ She made contradictions or irrelevant replies and refused to continue, saying she could not tell or even see the stimulus. Then again, with simple choices between ‘quick’ and ‘slow,’ she responded calmly and correctly.” (Trevarthen and Sperry, 1973: 560). “In general, L.B. was not able to describe more than one or two distinctive features of stimuli confined to the left peripheral field. . . In attempting verbal responses he was frequently

mediated exchange of visual information is usually referred to as “crude” or “primitive”—split-brain subjects can’t describe (LH) LVF (RH, right superior colliculus) visual stimuli in great detail, or sometimes at all—but still, they are capable of making some discrimination of shape and movement. Indeed, as I said before, the intertectal commissure, connecting right and left superior colliculus, seems to me probably the best candidate for mediating a genuinely intra-mind form of interaction between right and left hemisphere in the split-brain subject.

But even here I am uncertain, for the following reason. The superior colliculus, like many other non-cortical structures, is clearly involved in some clearly mental (especially motor-perceptual and attentional) processes. But it is, I think, through its interactions with cortical areas, that it is involved in these processes. It is less certain that colliculo-collicular interactions are themselves clearly mental. So while the intertectal commissure may allow a way for the two hemispheres to interact with each other, and to interact with each other via their connectedness to a midbrain structure that does figure importantly in some mental processes, it is still not clear that this interaction serves to really integrate the *mental* processing of the two hemispheres. Intertectal transfer may not take the right form for this to occur.

confused and he mumbled, or his speech was temporarily arrested. On many occasions he made elaborate and totally inaccurate confabulations describing details of visual appearance which were not even remotely related to the stimuli. Confabulation could be obtained by forcing him to give elaborate responses, or it could be greatly reduced by cautioning him to give simpler responses, more carefully and slowly. Occasionally, however, he could give a clear spontaneous description of one or two elements of form” (ibid, 560-561). For instance, when two parallel black bars were presented to his LVF (RH) he said, “straight, like square blocks. No, not square; long and thin, like caskets, two of them.” The paper is also fascinating from the perspective of trying to distinguish conscious from non-conscious perception, and raises the question (see Dennett, 1997) of just how informationally rich a representation has to be before we take it as conscious—and just how informationally impoverished a representation must be before we take it as non-conscious.

So, again, while subcortical operations make no doubt crucial contributions to genuinely cognitive processes, it is not clear that they themselves constitute such processes. It is via their interactions with cortical structures that non-cortical structures most clearly participate in mental processes. And these cortico-subcortical (or cortico-pontine, or cortico-tectal, or cortico-cerebellar, etc.) connections tend to be lateralized. The right superior colliculus is in communication with the right visual cortex. The left thalamus with the left cerebral hemisphere. The right cerebellar cortex with the right motor cortex. There are bilateral sensory and motor projections between some of these higher and lower structures; in these cases, we *might* think it relevant to the “how many minds?” question to know whether these bilateral pathways engage in some kind of processing or merely transmit (without transforming) representational contents, though perhaps this makes no difference. This is the sort of subtle and difficult issue I’m not sure we’re in a position to answer yet.

In the next subsection I consider more directly the nature of cortical-non-cortical interactions, rather than the nature of non-cortical representations. Insofar as non-cortical structures and functions serve to coordinate the functioning of the two hemispheres, they do so, I believe, largely in a manner similar to the way in which shared body and environment coordinate their activities: these structures operate to ensure that similar contents are produced *by the two hemispheres*, and give both hemispheres a similar ability to control many parts of the body. This sort of coordination between the hemispheres is consistent with their still constituting two

minds, albeit two minds that causally interact with each other richly and at every moment. But this interaction, in other words, may still be of an inter-mind sort.

6.3.2 Coordination versus integration

The criteria for individuating minds that I sketched in Chapter Two were founded on the claim that a single mind is characterized by *some* kind or kinds of cognitive integration, even if we must wait for a developed psychological theory to give a characterization of that integration. Accordingly, when two mental systems significantly integrate their activities, they become a single mental system.

Whatever the degree of sophistication of non-cortical operations, and whatever the psychological status of non-cortical representations, there remains the separable question of to what extent non-cortical structures serve to *integrate* rather than to simply coordinate the activities of the two hemispheres. Many neuropsychologists believe that it is this sort of integration, this intra-mind interaction, that the corpus callosum affords. But the inputs and outputs of two systems may be highly *coordinated*—and yet there can remain two systems.

A similar sort of distinction, though using slightly different terminology, can be found in Seymour et al. (1994), in which is urged “the need to distinguish between the *transfer of information* between the hemispheres via some extra-callosal route versus the integration of *outputs* from the two hemispheres by non-cortical structures. . . with the latter form of integration neither hemisphere has *access* to the *processing* operations of the other, but rather some other process *coordinates the outputs of both*” (113; emphasis added). The authors say this after remarking that in many cases, behavior that looks to be the result of interhemispheric *cooperation* in the split-brain subject may in fact be the result of the hemispheres working separately, particularly if

there is some mechanism in place that allows the hemisphere more likely to be *right* in a particular instance to dominate performance.

Some non-cortically mediated interaction between the hemispheres in split-brain subjects actually bears, upon careful examination, a striking resemblance to more obvious forms of cross-cuing. In an experiment allowing for extended viewing of lateralized stimuli, Myers and Sperry (1985) found for instance that patient L.B. was pretty good at naming letters, digits, and pictures presented to his LVF or RH. But he could only do this if he had previously been shown the pool of *choices* each individual presentation would be taken from. Corballis notes that “L.B. seemed to use an active strategy of rehearsing the likely alternatives, presumably with his left hemisphere, until a matching one ‘sticks out’” (1994: 164)—until, presumably, his right hemisphere, the only hemisphere that knows the answer, can respond somehow (perhaps by producing an affective state). Then, “reading” the response, the left hemisphere can take over again.

In another study (Gazzaniga, 1987, designed in part to replicate results from Sergent, 1983), pairs of letters were presented bilaterally, one in each visual field, to split-brain subjects, who were then asked to press one button if *either* letter was a vowel and a different button if *neither* letter was a vowel. Split-brain subjects were above chance at this task, but their “no” responses took significantly longer than their “yes” responses. From the observed time differences between affirmative and negative responses Gazzaniga (1987) inferred that although the hemispheres *cooperated* in performing this task, they were forced to do so without actually exchanging information. Corballis (1994) writes of this same study, in striking terms,

“each hemisphere can *observe* whether or not a response is forthcoming from the other hemisphere, and if no response occurs within a certain time period it can *infer* that no vowel was presented” (Corballis, 1994: 168; emphasis mine). He calls this a kind of “poker” strategy, and it is one that appears to play a prominent role in the degree of integration or at least cooperation the hemispheres exhibit.⁵⁰ That is, both hemispheres perceive and process independently of the other, and then each hemisphere (or perhaps just a single hemisphere) waits to see what the other hemisphere’s “move” will be: in some instances the move is not an actual physical movement, but rather some other kind of output (such as an *intention* to respond).

In other cases the interaction between the hemispheres appears to take the form of one hemisphere actively *interrogating* the other hemisphere for a response to some kind of *binary* question. The hemispheres appear capable of transferring information about many things, as long as the transfer can take the form of a question with only two possible answers: yes/no, even/odd, etc. For instance, imagine that a split-brain subject is bilaterally presented with pairs of objects—a coat and a hat, a coat and a coke, a cherry and a coke—and that the subject’s task is to say whether the two objects belong to the same or to a different category. Receiving the image of a hat, the LH can in effect “ask” the RH: “Is it clothing?” The exchange of binary information can be used in some sophisticated ways: if a split-brain subject is presented with bilateral number pairs and asked to say whether the numbers add up to greater or less than 10, and her LH receives the number 4, the LH can in effect “ask”

⁵⁰ At least during split-brain experiments; whether the two hemispheres would cooperate to this same degree outside of the split-brain experiment is difficult to say. Outside of the split-brain experiment there may be less pressure on either hemisphere to cooperate with the other, since a single hemisphere can simply grab the information it knows it wants from the environment directly, and since there is no experimenter to ask questions that “stump” a hemisphere.

the RH: is *your* number *big* or *small*? Receiving this information alone will allow the subject (or his LH) to give the answer on most trials. (It is possible that the transfer of binary information plays a very significant role in the set of studies in which split-brain subjects exhibit interhemispheric interaction; Myers and Sperry (1985) themselves noted that interhemispheric transfer of information in split-brain subjects does not appear passive or automatic but rather a matter of “active search”.)

It is instructive to compare the kinds of interhemispheric interactions that exist in split-brain subjects to those afforded by the corpus callosum in “normal” subjects (and perhaps by other structures in acallosal subjects). Weissman and Banich (2000) for instance found that, with the corpus callosum intact, the hemispheres couple or uncouple their activities depending on the complexity of the task subjects are engaged in, practicing a divide-and-conquer strategy only for more complex tasks. Hochman and Eviatar (2004) found evidence that the division of labor between the hemispheres for complex tasks often takes the form of one hemisphere (perhaps the one dominant for that task) engaging in the initial processing and responding, while the other hemisphere concurrently monitors the first hemisphere’s performance, engaging in error correction, post-response, when necessary. These studies used “normal” subjects, and their results may explain both why split-brain subjects often suffer greater impairments in performance as task complexity increases (e.g. Kreuter, Kinsbourne, and Trevarthen, 1972), and why they fail to engage in spontaneous self-correction (Kaplan and Zaidel, 2002), the latter of which is something that “normal” subjects apparently almost always do (according to Kopp and Rist, 1999). Of course, even in the examples above—particularly the latter—the interaction between the

hemispheres interestingly seems to resemble close communication and coordination *between two minds*. It is nonetheless clear that in comparison to such instances, the interaction between the two hemispheres of a split-brain subject is of a *less* intra-mind sort.

Clearly the two hemispheres of a split-brain subject do not function entirely independently of each other: in fact, they will *causally* interact with each other essentially at every moment. The question relevant to individuating minds in split-brain subjects is whether there is some psychologically important distinction between intra- and interhemispheric interactions. It is of course too soon to be certain that there is such a psychologically important distinction—but it *seems*, at this point, that there is. As Reuter-Lorenz and Miller summed up the literature on subcortically-mediated (and non-cortically mediated) interhemispheric interaction:

In fact, the separated hemispheres are virtually isolated from one another, so that one hemisphere is rarely conscious of stimulation presented to the other. Visual and tactile information presented to one hemisphere cannot be compared with information presented to the other hemisphere. The separated hemispheres are also unable to share abstract codes or semantic information, despite previous claims to the contrary. Coarsely coded information about spatial location may be shared, but this capacity may be limited to central portions of the visual field. Moreover, the separated hemispheres can implement different visuomotor programs concurrently and with minimal interference. . . Thus, the autonomy of the hemispheres produced by callosal section extends beyond perceptual processing into the stages of decision making and response preparation.

(Reuter-Lorenz and Miller, 1998: 16)

This last point is particularly significant. Sometimes it is argued that while the hemispheres function independently at the level of some “high-level,” especially

conscious perceptual, processing, the mental architecture of a split-brain subject remains “unified” at some deeper level.⁵¹ In response to such arguments it is important to recall that the functional *split* between the hemispheres itself goes very deep. Processing becomes “unified” at some point: for instance Reuter-Lorenz and Miller speculate that the reason split-brain subjects so rarely exhibit interhemispheric *conflict* in their behavior is that “it has been shown that each hemisphere is able to inhibit, or gate, the *responses* of the other” 1998: 16; emphasis added). So some form of coordination is achieved eventually. Arguably, however, this coordination occurs at *so* deep a level, at so functionally late a stage, that it constitutes coordination *between minds*.

Let me end by quickly clearing up one likely misunderstanding of my view of non-cortical mental processing. I do believe that non-cortical structures are essential to the operation of the mind (or minds). And I even accept that at least some non-cortical processes are not just *causally* necessary to mental operation but are themselves *components* of, partly constitutive of, mental operations. So my view is not that only cortical operations are mental. My view is that, first, non-cortical operations are mental in large part in virtue of being incorporated into cortical operations. Interactions between right and left superior colliculus in and of themselves, for instance, may not be mental. But interactions between striate and extra striate cortex and the superior colliculus are mental.

⁵¹ Sperry said this at one point about even the conscious structure of split-brain subjects: because “mental-emotional ambience or semantic surround generated in one hemisphere promptly spreads also to the second hemisphere”, he suggests that the “structure of the conscious system in the divided brain. . . [is] Y-shaped, i.e., divided in its upper, more structured levels but undivided below” (1986: 15)

This may not be the case in all animals. Berridge for instance notes that “Humans can be devastated, rendered into vegetative states, by large neocortical lesions, whereas a rat can lose its entire neocortex and continue on remarkably normal” (2003: 41). Berridge is articulating something like the view I’m defending here; indeed this view is not novel; Berridge himself cites Gallistel (1980) and Hughlings Jackson (1958), a 19th century British neurologist. Berridge calls the evolutionary (and probably developmental) process by which subcortical processes become incorporated into cortical ones, “encephalization” (the term “encephalization” has other uses and meanings as well.) “Encephalization has modified the autonomy of subcortical structures, so that neocortical inputs have become incorporated into subcortical function to a greater degree than in nonhuman animals, so much so that human subcortical function cannot be maintained normally when cortical inputs are suddenly removed” (Berridge, 2003: 43). Merker, even in the context of arguing that *consciousness* is *not* a cortical function, admits that “In adults massive bilateral cortical damage will typically issue in a so-called persistent vegetative state” (2007: 65). Admittedly, “This by itself does not, however, allow us to make an equation between cortical function and consciousness, because such damage inevitably disrupts numerous brainstem mechanisms normally in receipt of cortical input” (ibid). But I do not need, nor intend, to claim that the hemispheres function to sustain consciousness or anything else *independently* of non-cortical operations. And again, I do not even need to claim that non-cortical structures contribute only non-mental inputs to (and coordinate the outputs of) mental, cortical processing. *Noncortical-cortical interactions may be mental. Noncortical-*

noncortical interactions may not. There are two mental systems in the split-brain subject, a right hemisphere-noncortical system and a left hemisphere-noncortical system. Taken by themselves, right and left hemisphere superior colliculus, for instance, or right and left hemisphere thalamus, are not mental systems. Non-cortical structures owe their mental status to the cortical elaboration of their activities.

Finally, the degree to which non-cortical processes are incorporated into cortical ones has another consequence as well: that of making the overall *mental* architecture of the human brain largely “vertical” as opposed to “horizontal.” This idea is difficult to cash out precisely; to convey it at this point is more to convey a general impression resulting from sifting through a bunch of data, than it is to convey even a *sketch* of a model that would systematize this data in some way. But let me convey what I can. What I mean by saying that perception, cognition, and motor intentions are functionally instantiated in the human brain in “vertical” loops of activity, is that these loops pass through spinal cord and cerebral hemisphere and thalamus and cerebellum and midbrain superior colliculus, and back up to cerebral cortex and back down to spinal cord and via the midbrain reticular formation, etc. . . . And there are comparably few loops passing (“horizontally”) from one side of the brain to the other. There is some “direct” communication between the hemispheres, via the intertectal commissure, for instance, but by and large, right and left superior colliculus seem to interact more with right and left hemisphere, respectively, than with each other. Moreover, when there is communication between right and left side structures, when there is a circuit running from one side of the brain to the other and back again, it is largely just means of transfer, just a means of making contents

available for processing by both sides of the brain, rather than consisting in a kind of (physically) central processing itself. The actual processing, the computation, the transformation of representational contents, seems still to occur either *within* a right or left side structure, or between structures, but *vertically*. The processing isn't only cortical; it's in part sub- and noncortical. But it is still largely lateralized.

Until the corpus callosum, that is. The corpus callosum suddenly adds a substantial “horizontal,” left-right element to the otherwise mostly vertical, up-down, functional organization of the brain. With that pathway in place, there is a case for the two hemispheres constituting a single *mental* circuit. They interact not simply because they're “hooked up to” the same body and environment—via a loop that passes outside the boundaries of any mind—but because they are engaging directly with each other. When the corpus callosum is sectioned, however, the functional organization of the brain is again on the whole vertical.

7 *Conclusion*

This chapter argued that split-brain subjects have two minds, whether we individuate mind on the basis of purely psychological criteria or partly neural criteria. The two methods of individuating minds yield the same result in part because there is, at a very abstract level, an isomorphism between mental and neural structure.

I rejected verificationism in the “how many minds?” debate; while behavioral evidence is of great importance in determining how many minds a creature has, integrated behavior does not provide decisive evidence of mental singularity, since there are many sources of behavioral integration that are mental but do not reflect

intra-mind interactions, and other sources of behavioral integration that are not mental at all.

The “how many minds?” question is properly a question about mental *architecture* and mental *tokens*, and thus type-identity (identical *contents*) cannot make a single mind.

While noncortically-mediated interhemispheric interactions could of course in theory integrate the mental processing of the two hemispheres so as to yield a single mind, and may in fact turn out to do so, I gave some reasons to think that the two hemispheres of a split-brain subject are in fact associated with distinct minds.

Finally, while acknowledging that accepting the mental duality model poses something of a challenge to our various practical ways of interacting with and relating to split-brain subjects, I also suggested that integrating our theoretical and our practical understanding of such subjects is not quite so difficult as some philosophers have feared.

Chapter 4: On Co-consciousness and Conscious Unity

1 Introduction

In this chapter I turn to the structure of split-brain subjects' consciousness. In the first part of the chapter I lay the groundwork for comparing competing models of the structure of split-brain subjects' consciousness. In the second half of the chapter I defend the model I advocate, the conscious duality model, from the most obvious challenge it faces: that of explaining the integrated nature of split-brain subjects' behavior.

The unity of consciousness, as a scientific subject, is a jungle: rich, complicated, and vast—a terrain in which are woven together the strands of various other areas of study, from perception and attention to action and agency to the binding problem. One loses one's bearings here very easily. There is so much more that *could* be said about this topic than what I can discuss here, that I accordingly spend a good deal of this chapter simply stating which issues I will and will not be dealing with. I also draw a few important distinctions and explain how I will be using certain contested terms in this terrain. Then I offer a simple account of *co-consciousness*—the relation composing conscious representations into streams of consciousness—based in the Global Workspace theory of consciousness. This account of co-consciousness is then used to defend the conscious duality model for split-brain subjects.

By way of introduction both to the debate about the structure of consciousness in split-brain subjects, and to a more general debate concerning what phenomenology and conscious contents can tell us about conscious structure, I will in the next section introduce one particular voice in this debate: that of Puccetti (1989, 1981), who argued that not just a split-brain but also a “normal” subject has two minds, two streams of consciousness—and is (associated with) two persons. While I don’t draw many of these conclusions, I, too, defend a conscious duality model for split-brain subjects.

2 *Puccetti*

Probably the most infamous philosopher’s stance on split-brain consciousness, cognition, and personal identity, is Puccetti’s (1989, 1981). Puccetti believed, first, that split-brain subjects have two streams of consciousness, a conclusion that many neuropsychologists also accepted, but one that philosophers were (and continue to be) more reluctant to accept.

Next Puccetti argued that we *all* have two streams of consciousness. This second claim was considered more extreme; Bogen, it is true, has also advanced a (qualified) conscious and mental duality claim for “normal” subjects (Bogen, 1981, 1990), but even most neuropsychologists believed that “normal” subjects have single minds and single streams of consciousness. (Sperry and Gazzaniga certainly believe this; see for instance Sperry 1977, 1968, and LeDoux and Gazzaniga, 1981.)

Finally Puccetti argued that each “normal” human subject is also two persons, or rather, that there are two persons associated with a “normal” subject’s brain,

though only one person (the left hemisphere person), speaks. (So, *I* am writing this paper, but *I* am only one of the people associated with my brain and my body. The other person associated with my body, the right hemisphere person, doesn't fight to make its—his or her?—presence known, having learned submission to the dominant hemisphere's person from an early age.) Even Bogen (1981, 1990) hadn't gone this far, insisting that there remained enough *integration* and commonality between a “normal” subject's two streams and two minds to generate only a single person.

Puccetti's argument for this startling conclusion takes the following form⁵²:

- P (1): There are neural processes in the left hemisphere of right-hemispherectomy (or split-brain) subjects that suffice for (or constitute) conscious experience; there are neural processes in the right hemisphere of left-hemispherectomy (or split-brain) subjects that suffice for (or constitute) conscious experience.
- P (2)/Conclusion 1: Therefore there are neural processes in the left and right hemisphere of a “normal” subject that also suffice for or constitute conscious experience.
- P (3): The corpus callosum does not serve as a “fusing” mechanism that somehow fuses these two sets of experience into a single stream of consciousness. (Or, at least, the burden of proof lies on those who believe that the callosum *does* serve as such a mechanism, to show that and explain how it performs this function.)
- P (4)/Conclusion 2: Therefore each “normal” subject (like each split-brain subject) possesses two streams of consciousness, one associated with each hemisphere. In a “normal” subject, the corpus callosum acts to ensure that the two streams have duplicate contents. Thus right now for example, both of my hemispheres are generating a conscious visual experience of my laptop.
- P (5): But we do not (usually) experience “double vision.” I only see one laptop, that is.
- P (6)/Conclusion 3: Therefore each stream of consciousness, each hemisphere's set of experiences, is associated with a unique subject of experience.
- P (7): There is a unique person for each subject of experience.
- P (8)/Conclusion 4: A “normal” human subject is associated with two persons: a speaking LH person and a silent RH person. So,

⁵² Some of what follows is only implicit in Puccetti.

the “I” who sees only a single laptop is really just the person associated with *one* of my hemispheres, my left hemisphere. (Since “I” am typing this. There is another, mute person, associated with my right hemisphere, which is the subject of the second experience of this laptop.)

Several of Puccetti’s commentators within the philosophical community took issue with his first premise: with the claim that the isolated right hemisphere can generate conscious experience. This claim was much controversial in the neuropsychological community, although there were some here who questioned this claim as well (e.g. Davidson, 1981).⁵³

For many other commentators, especially those in the neuropsychological community, premise (4) was particularly problematic. It is true that of course the burden of proof in general lies on those who would insist that some particular thing *does* exist or that some existing thing *does* have some property where there is no (widely acknowledged) evidence for this positive claim. But similarly the burden of proof lies on those who would insist that some particular thing *doesn’t* exist or *doesn’t* have some property when there is (widely acknowledged) evidence against this negative claim. In this case, there is one widely acknowledged piece of evidence for the positive claim, i.e. for the claim that the corpus callosum does act as a fusing mechanism: the split-brain phenomenon itself!

⁵³ Another non-philosopher responded:

As a clinical neurologist I have been trained not to deal with the concept of consciousness. I can deal with responsiveness, for that I can test with a stimulus, grading the response or noting the nonresponse. What is going on in that patient’s brain between the stimulus and the response is his own province. Presumably, what is going on is consciousness or awareness of the stimulus and response—but it is still the private domain of that individual. Therefore, I will deal with responsiveness and hope that ‘consciousness’ fades into a well-deserved obscurity.

(Joynt, 1981: 109)

The inference from premise (5) to premise (6) was not subject to particular criticism, and yet I believe it is the most interesting move in his argument—though not necessarily the most convincing. This is where Puccetti reasons that *if* my brain is currently generating two experiences of my (one) laptop, and *if* I nonetheless do not see *two* laptops, then each experience must belong to a different stream of consciousness, and a different subject of experience. Note that this is in essence an argument for conscious duality grounded in phenomenology. These are different grounds from those in which I ground my arguments for the conscious duality model of split-brain subjects; I don't believe that phenomenology tells us very much about the *structure* of consciousness. This is mainly because there is nothing it is like to have (or no phenomenal experience of having) two streams of consciousness. We cannot *feel* that we have two streams of consciousness, because conscious feelings are located *within* one stream of consciousness or another; there is no phenomenal perspective external to a particular stream of consciousness from which to experience multiple streams.

We should furthermore pause to consider whether having two co-conscious experiences, E1 and E2, both with content C, would *feel* any different from having just E1 or just E2 (with content C.) Imagine, for the moment, that Puccetti is right that each hemisphere of my brain is generating a set of experiences, and that each hemisphere is generating its set of experiences with some degree of causal independence from the other. (We have already seen the difficulty of distinguishing direct from indirect causal interactions in the generation of mental phenomena, but let us simply say here that the conscious representations or experiences within each

hemisphere interact with each other in certain subtle and complex ways, ways in which they *don't* interact with the conscious representations of the other hemisphere.) Now imagine that each set of experiences contains one or more experiences of my laptop right *there* before me: call these E1 and E2. According to Puccetti, I should see two laptops—unless these two experiences of the laptop belong to distinct streams of consciousness and distinct subjects of experience. . . and *I* am really only one of these two subjects of experience.

But it is not immediately obvious why I should be expected to see *two* laptops in front of me, no matter how many streams of consciousness or subjects of experience are associated with my brain. What I *see* is a function of what my experiences represent, and what they are representing is *a single* laptop at a *single* location. Why, then, would I see *two* laptops unless I had either one (or more) experiences representing the presence of *two* laptops before me, or two experiences, each representing a laptop at a different location? (Note that this is what “double vision” involves: seeing two objects that do not occupy the exact same space, though they might overlap in space.) Hurley (1998) seems to make a similar point when she writes that “There is nothing it is like for there to be one experience with content *r* rather than two duplicate experiences at the same time with content *r*; there is no subjective contrast here. Therefore, we cannot imagine what it is like for there to be one experience rather than duplicates” (Hurley 1998: 165).

While the *content* of conscious experience tells us *some* things about the structure of consciousness, there is much that it does not tell. The structure of consciousness cannot be determined from its contents in many instances but can only

be determined by considering the functional properties of and the functional relations between the various representations carrying conscious contents.

3 Some Claims and Distinctions

In this section of the paper I lay out a number of distinctions and positions concerning consciousness and conscious unity in general, and the structure of consciousness in split-brain subjects in particular.

3.1 Right Hemisphere Consciousness

I will be assuming that the right hemisphere of a split-brain subject is associated with conscious phenomena. There are two sorts of data that I believe provide especially compelling evidence of conscious processing in the “disconnected” right hemisphere. First, there is data concerning the right hemisphere’s emotional responsiveness: V.P.’s distress after her right hemisphere was shown a film in which people are thrown into flames (Gazzaniga, 1985); the “wide, sheepish, and (to all appearances) . . . self-conscious grin” that accompanied the “thumbs-down” L.B. gave in response to a (LVF/RH) photograph of himself (Sperry, Zaidel, and Zaidel, 1979); the right hemisphere’s positive feelings about justice, abortion rights, and sex, and dislike of taxes (Schiffer et al. 1998).

Split-brain subjects also engage in some right hemisphere-controlled behaviors that strongly suggest conscious experience and processing. They can draw objects whose names only their right hemispheres have been presented with (Kingstone and Gazzaniga, 1995). During a tactile learning test, one split-brain

subject “broke out laughing” upon first encountering (with his left hand/RH) a stimulus consisting of a square wooden block with a tack nailed into its center. “Every time he felt it, he would pick it up and twirl the block about the axis and would chuckle heartily when doing so. When asked what was funny he would say, ‘I don’t know, something in my left hand I guess’” (Gazzaniga, 1970). The right hemisphere also performs comparably to the left on the Raven Progressive Matrices Test (Zaidel, Zaidel, and Sperry, 1981), a fairly involved test of non-verbal reasoning. Shallice explains:

To perform correctly on this test, the relation between two items has to be abstracted and then extrapolated so as to infer the third item in a progression; finally, the results must be matched to one of a set of possible answers. The processes involved are far more demanding than, say, those used in picture-word matching in the number of components, the level of abstraction, and the involvement of more than the operation of a routine schema. If this level of performance could be obtained unconsciously, then it would be really difficult to argue that consciousness is not an epiphenomenon.

(Shallice 1997: 264)

This isn’t to say that the conscious experience and cognition of the two hemispheres are the same. They surely aren’t. The hemispheres have non-identical patterns of perceptual access to the body and to the world, and moreover there are hemispheric asymmetries with respect to perceptual *processing*, that seem likely to result in differences in conscious contents. And the left hemisphere’s dominance for productive and syntactic language suggests that only the left hemisphere will be

associated with a (normal) stream of inner speech. This in and of itself would amount to a difference in conscious contents, of course, but the difference between the two hemispheres with respect to their linguistic capacities may have further effects on conscious cognition as well.

For instance, Zaidel, Zaidel, and Sperry (1981), in the study cited above, actually gave split-brain (and hemispherectomized) subjects two versions of both the Raven Standard Progressive Matrices and the Raven Color Progressive Matrices tests, a book version and a board version, the latter of which is designed to encourage a trial-and-error method of solving the problems. (In the book version of the test, subjects select the correct answer from a series of choices by pointing. In the board form, each problem appears on a board with the “solution piece” missing; the answer choices are a number of moveable pieces each of which fits neatly into the missing space, and subjects are encouraged to try out different solutions until they’re satisfied with their answer.) While the right hemisphere tends to slightly outperform the left on the book form of at least the somewhat simpler Raven Color Progressive Matrices test, the left outperforms the right on the board form, suggesting that the left hemisphere benefits more from the ability to take a trial-and-error approach to the task. The authors note that the finding that the book form of the Raven Progressive Color Matrices test favors the right hemisphere, and the board form the left, “is in accord with a study of thirty-two 7-8 year-old children in which it was found that verbalization of strategy during and after solution of the problem improved scores in the board form of the test but actually decreased scores in the standard book form” (Zaidel, Zaidel, and Sperry, 1981: 175, citing study by Carlson et al., 1974).

So, again, I don't assume that the hemispheres are subject to the same conscious contents, nor even that their phenomenal experience is qualitatively the same—visuospatial processing differences between the hemispheres could result in different sorts of phenomenal experiences—nor that they are equally capable of all the same kinds of conscious cognition. The left hemisphere may be capable of more sophisticated processing in some cases due to its superior linguistic abilities, and its capacity to sustain a (normal) stream of (inner) speech, for instance. And it may be difficult (for some) to imagine engaging in conscious cognition that doesn't rely on inner speech, or to imagine possessing a stream of consciousness from which a stream of inner speech has been stripped. Despite this, the evidence does make it reasonably safe to conclude that the “disconnected” right hemisphere has some sort of conscious mental life.

3.2 Access versus Phenomenal Consciousness

Much of the current philosophical literature on consciousness draws a distinction between access consciousness and phenomenal consciousness (Block, 1995). If there is something it is like to possess a mental state, that mental state is said to be *phenomenally conscious*. A mental state that is available for some kind of central cognition is said to be *access conscious*.

The major model of access consciousness is the *global workspace theory* of consciousness (Baars, 1988, 1997; also Dehaene and Naccache, 2001). While there is no general or complete or accepted theory of consciousness, the global workspace theory has a number of defenders both among philosophers of psychology and among neuropsychologists, and I accept the rough adequacy of the theory in this work.

According to this account, non-conscious perceptual representations may play an important role in cognition and behavior, but they do so in virtue of their accessibility to only one or a few local processing systems. For instance, Milner and Goodale (1996; also Goodale and Milner, 1992) showed that some motor control systems use non-conscious visual representations that are not available consciously and therefore cannot be reported by the subject or used in the subject's practical reasoning. (For instance, in the Müller-Lyer illusion, subjects typically report that the line ending in outward pointing arrows is shorter than the line ending in inward pointing arrows. But when you ask them to pretend that they are reaching for the arrows to pick them up, their fingers are the same distance apart for the two lines.) Conscious representations may be used in practical reasoning to select actions to perform, but the actual performance of visually-guided actions uses representations for online control and guidance that are not conscious, that are made locally available only to motor systems.

Some perceptual representations, however, achieve a state of *global* cognitive importance; these representations—access conscious representations—are broadcast to a large number of consumer systems at one time. (I will call these global consumer systems “global consumers.”) Their joint accessibility to multiple global consumers at one time allows them to be used in a kind of “central” or “domain-general” cognition. Representations that *are* broadcast to global consumers are said to be located within a “global workspace.” (This workspace is functionally defined, though there may be workspace areas of the brain as well.) Some plausible global consumer systems are: long-term and working memory; an emotional system or systems; a mind-reading

system; a conceptual system; a verbal report system; and a practical reasoning system.

The nature of the distinction between access and phenomenal consciousness—conceptual or metaphysical—is open to dispute. Some believe that a theory of access consciousness also offers a theory of *consciousness*, period (e.g. Dennett, 2001; Dehaene et al. 2006). Others deny that a theory of access consciousness offers a theory of phenomenal consciousness, while accepting that all phenomenally conscious representations are nonetheless also access conscious (Carruthers, 2005a). Others deny even this, and assert that there can be phenomenally conscious states that are not access conscious (e.g. Block, 2007; Lamme 2004, 2003). (Note also that some who assert that all phenomenally conscious states are also access conscious do not assert the converse; that is, they believe that there may be access conscious representations that are not phenomenally conscious (e.g. Prinz, 2007a).) Finally, some question even the *conceptual* distinction between access and phenomenal consciousness. (E.g. Church 1997; Dennett 1997.)

Phenomenal consciousness is often taken to be a “harder problem” for psychology than access consciousness. The latter phenomenon easily lends itself to functional characterization. But some of those who deny that a theory of access consciousness can explain phenomenal consciousness seem to do so because they believe that what is most mysterious about phenomenal conscious are *qualia* or qualitative properties: the redness (as opposed to the greenness) of the experience of red; the pleasantness (as opposed to the painfulness) of pleasant experiences. And they do not see how mere global *accessibility* can confer or explain phenomenal

properties. In fact some doubt whether *any* functional property can explain these qualia, and would refer instead to intrinsic neural properties.

In addition to accepting the global workspace theory of access consciousness, in this work, I also accept that all phenomenally conscious representations are also access conscious, or (according to the global workspace theory) *globally broadcast*. But I do not, in this work, take a stand on whether or not global workspace theory also offers an *account* of phenomenal consciousness. The claim that all phenomenally conscious representations are also access conscious is less controversial than the claim that there is just one kind of consciousness, or that the global workspace theory offers a theory of phenomenal consciousness. In fact I may not even need the claim that all phenomenally conscious representations are also access conscious. But I will claim that phenomenal *co-consciousness* requires access consciousness.

3.3 Co-consciousness

Right now as I type this I am conscious of the feel of the laptop keys under my fingers. So I have a tactile experience (or perhaps multiple tactile experiences) of myself typing this. But I can also (consciously) *hear* myself typing this (the keys click as I type). So I have an auditory experience (or perhaps multiple auditory experiences) of myself typing this. Moreover, my tactile and my auditory experiences bear a certain relation to each other. When this relation holds between two conscious or consciousness-composing representations, it is *because* it holds that there is *something it is like* to have both of those representations, together, rather than there simply being one thing it is like for me to possess one representation and a distinct thing it is like for me to possess the other representation, and nothing it is like for me

to have them both at once. I will call this relationship “co-consciousness” (as is fairly standard, although Bayne and Chalmers (2003) prefer to speak of subsumption, and thought Tye (2003) rejects the concept altogether).⁵⁴

I have characterized co-consciousness thus far in terms of *what-it’s-like-ness*; this might easily be interpreted as meaning that I am only interested in *phenomenal* co-consciousness in split-brain subjects. (For two states to be *access co-conscious* is presumably just for them to be *jointly* available for some kind of “central” or “domain-general” cognition.) I am actually interested in both phenomenal and access co-consciousness, though the structure of split-brain subjects’ *phenomenal* consciousness may be more controversial than the structure of their access consciousness (assuming the structures of the two kinds of consciousness can diverge in some way). Bayne and Chalmers (2003) for instance suggest that the split-brain phenomenon more clearly involves disunity of *access* than of phenomenal consciousness. But while I may continue to speak more explicitly of phenomenal than of access co-consciousness, implicitly I will really be talking about both, for I believe that access consciousness and phenomenal consciousness in split-brain subjects do not diverge in structure, for several reasons. First, I share some of Dennett’s (2001, 1997) suspicions about suggestions that there are conscious experiences that cannot be reported (verbally or otherwise) in any way. I also echo Tye’s sentiments that

⁵⁴ Tye rejects the use of the term because he rightly believes that its standard use implies that two co-conscious representations are each themselves conscious experiences. My use will not assume this, though; two co-conscious representations must belong to or compose a stream of consciousness, but my use of the term will not assume that these co-conscious representations are themselves conscious phenomena. The term “co-consciousness” is simply very convenient given the distinction I will draw between conscious disunity and conscious duality, for I can speak of the presence or lack of co-consciousness between representations without referring to their being unified or disunified. I will continue to speak of “conscious representations,” and not just “representations composing a stream of consciousness,” but nothing I say hinges upon this choice of phrasing.

surely the simplest explanation for why a split-brain subject who has just seen the word “pen” in his LVF (RH) and “knife” in his RVF (LH):

has no access consciousness of ‘pen’ next to ‘knife’ is that he has no visual experience of the one design to the left of the other in his visual field. He has an experience of ‘pen’ and, on this basis, he undergoes a cognitive state representing ‘pen’. He also has an experience of ‘knife’ and, on this basis, he undergoes a cognitive state representing ‘knife’. But he does not experience the two words together. That’s why access consciousness of both [together] is missing.

(Tye 2003: 125)

Finally, the account of phenomenal co-consciousness that I offer will simply be incompatible with the suggestion that split-brain subjects have two streams of access conscious representations but one stream of phenomenally conscious experience. So while there may be a conceptual distinction between phenomenal co-consciousness and access co-consciousness, just as there may be a conceptual distinction between phenomenal consciousness and access consciousness, I submit that we needn’t draw this distinction in determining the co-consciousness relations between conscious representations.

A distinction that is more than just conceptual, however, is that between synchronic co-consciousness—co-consciousness at a time—and diachronic co-consciousness—co-consciousness across time. The concept of a *stream* of consciousness actually probably connotes diachronic as much as synchronic co-consciousness: one experience flows into another, a moment later, and it is difficult to say when one ends and the next begins. Intuitively most of us feel, that is, that our conscious experience doesn’t stop and start and stop and start at every instant; there is

a fluidity or continuity to our experience. (Though Strawson (2003) thinks that “the basic form of our consciousness is that of a gappy series of eruptions of consciousness from a substrate of apparent non-consciousness” (359), so intuitions differ.)

Diachronic co-consciousness is perhaps even more difficult describe than synchronic co-consciousness in phenomenal terms. Roughly, though, the idea is that, if I have experience E1 of a cat crouching and then a moment later an experience E2 of the same cat beginning to pounce, there is something it is like to experience E2 immediately following E1, rather than there just being one thing it is like to experience E1 and then another thing it is like to experience E2 and nothing it is like to experience the two in quick succession. (Compare the case I just described to a case in which the last thing I see before I drift off to sleep is a cat crouching, and the first thing I see when I wake up is a cat pouncing.)

Note that intuitively synchronic co-consciousness seems *transitive* in a way that diachronic co-consciousness doesn't. Intuitively if E1 is co-conscious with E2, E2 with E3, E3 with E4, and so forth up to E100, *all at a moment*, then there is something it is like to experience E1 and E100 together as well. (This intuition will be challenged, however, once below, and once in Chapter Five.) In contrast, if E1 is *diachronically* co-conscious with E2, E2 *diachronically* co-conscious with E3, and so forth, up to E100, there does not seem to be something it is like to experience E1 and E100 *together*. (Well, after all, they *aren't* experienced together.) So the two kinds of co-consciousness are different in this sense.

The two forms of co-consciousness are of course related nonetheless, and ultimately an account of both is needed. Moreover, it is possible that both synchronic and diachronic co-consciousness require the same kind of account. Nonetheless in this work I deal only with synchronic co-consciousness, offering an account of what it is for multiple simultaneously conscious experiences or representations to be co-conscious with each other. This focus on synchronic co-consciousness is appropriate since at least the major puzzle of split-brain consciousness concerns the relationship between left and right hemisphere conscious experiences at any given moment. And I am able to maintain this narrow focus on synchronic co-consciousness in part because there is no *special* problem of diachronic co-consciousness facing the particular model of split-brain subjects' consciousness that I adopt. Since I believe that split-brain subjects have two streams of consciousness at all times, that is, and not just inside of experimental conditions, I don't face the problem of explaining how a single stream of consciousness could divide in two and then merge back into one again (and on the basis of simply putting on and removing a blindfold, respectively).

3.4 A Unified Consciousness versus a Stream of Consciousness

I define streams of consciousness in terms of the co-consciousness relation. If two (simultaneously) conscious representations belong to the same stream of consciousness, then they are co-conscious, and if two (simultaneously) conscious representations are not co-conscious, then they do not belong to the same stream of consciousness. At any moment, a stream of consciousness is composed of the entire set of experiences that are all co-conscious with each other. (At least, as a first pass. Chapter Five considers a view according to which synchronic co-consciousness need

not be transitive; if that is correct, then the concept of a stream of consciousness needs either revision or rejection.)

Thus, for two conscious representations to be co-conscious is for there to be something it is like to experience them (or their contents) together; for a subject to have a single stream of consciousness, composed of a multitude of representations, all of which are co-conscious with each other, is for there to be something it is like for that subject to possess all those representations at once. (In contrast, for a creature, C, to have two streams of consciousness is for there to be one thing it is like for C to possess some subset of C's representations together, and another thing it is like for C to possess the remainder of C's representations together, and *nothing* it is like for C to possess all of C's representations together.)

Throughout this dissertation I have at numerous points contrasted singularity with duality, rather than unity with disunity. The distinction between singularity and duality on the one hand and unity and disunity on the other is perhaps of particular importance when we examine the structure of *consciousness*. So in this and subsequent chapters I will distinguish between having a *single stream of consciousness* and having a *unified consciousness*. For I submit that these are in fact two different things: one could conceivably have a single stream of consciousness that could not properly be called unified; I even believe that there is at least one sense in which one could have a fairly unified consciousness while still possessing multiple streams of consciousness. Talk of conscious singularity versus conscious duality, or of having one versus two streams of consciousness, encourages a focus on individuating mental token, rather than on providing a qualitative analysis of the

phenomenal character of consciousness. And individuating mental tokens is the subject of this dissertation.

I have said that to have a single stream of consciousness is just to have all of one's currently conscious representations be co-conscious with each other; it is to have there be something it is like to bear *all* of those representations at once. To have a unified consciousness, however, is to have both something more and something less than this. The concept of a unified consciousness is richer than that of co-consciousness, because when we speak of human consciousness as unified, we don't, or don't always, just mean to claim that there is *something* it is like to possess our conscious representations together. We often are attempting to *describe* what this something is like. In particular, we often mean to say that this thing, our phenomenal experience, is not *disorienting*—that the entire phenomenal state-of-being has a kind of stable, centered, coherent feel to it. Sight and sound, touch and smell, are coordinated, and seem to occur from a particular physically localizable perspective. The body and world presented to us *via* our conscious experience of them seem to make sense: objects don't flicker in and out of existence; many motions follow predictable arcs; things stand in spatial relations to each other, and these relations are transitive—and all of this is perpetually confirmed via action, such that the information delivered separately from a particular sense generally coincides with the information delivered separately from the other senses, and such that perception can guide constructive action while action has predictable effects on perception. What we experience now usually makes sense given what we were experiencing a moment ago, and given the intentions we've formed and the actions initiated since that time.

It is at least conceivable, however, that there should be something it is like for a subject to possess all her current experiences together. . . . and that this phenomenal state-of-being should be in fact incredibly *disorientating*. Sight and sound, smell and taste, touch and balance, have all become radically unhinged from each other; it looks as if you're sinking down and feels as if you're floating up; you can't tell if you're hearing a loud noise from a distance or a quiet noise up close. Objects *do* appear to flicker in and out of existence. Walking in the direction of an exit only makes it look as if it's receding further from you. Thoughts and intentions are formed but experienced as coming from *outside* your head, while *outside things*—the wind in the trees, the stories in the news, other people's thoughts and voices and actions—are experienced as being under one's control, perhaps even private phenomena. Indeed, phenomenal states-of-being with some of these features may actually exist—perhaps in subjects with severe schizophrenia, for instance (or subjects who have not yet adapted to wearing lenses that invert directionality). Yet although co-consciousness would seem to be a precondition for even experiencing this sort of radical disorientation, it would be perverse to call such a phenomenal state-of-being *unified*.

In another sense, though, calling experience unified may mean less than calling experiences co-conscious. For, *prima facie* at least, it would seem that a creature *could* have two streams of consciousness, and yet enjoy a significant degree of conscious unity. First of all, each stream might itself enjoy a high degree of conscious unity (coherence of contents, for instance). Second and more interestingly, there might be a great deal of coherence *between* or across the two streams. The streams might have similar contents, for instance, or at least contents that are not in

some sense incompatible with each other. Conscious unity may be grounded in many other things besides similarity (or compatibility) of *contents*, also. For instance, a set of experiences may be more or less unified depending on the role they play in behavior, or depending upon the kinds of spatial relations that the objects they represent are experienced as standing in, relative to each other, or depending upon whether or not they all bear the same kind of accessibility relations to the same set of long-term memories. Experiences might be more unified to the extent that they are caused by the same things in the world.

Someone might press, however, that having a single stream of consciousness and having a unified consciousness are at most conceptually distinct, and that some kind of coherence in contents, for instance, is the basis not just of conscious unity, but also of co-consciousness. When Baars, for instance, writes that “It has been known for at least a century that two simultaneous, incompatible events do not become conscious at the same time” (1988: 126), he seems to mean that two conscious representations do not become co-conscious if their contents are incoherent. He provides as an example of two simultaneous, incompatible events, a speech sound in the left ear and a falling glass in the right visual field, and claims that only one of these events can be globally broadcast at any moment; those who doubt that these two events are truly incompatible can think instead of something like the phenomenon of binocular rivalry.

I agree that there appear to be significant and fascinating causal relationships between co-consciousness and coherence of contents, although I think it is difficult to know whether to say that coherence of contents is a precondition for co-

consciousness, or that co-consciousness compels that contents be made to cohere.⁵⁵ But in either case, this is fine; co-consciousness is clearly an aspect of conscious unity, perhaps even the most important aspect of it, and perhaps there are limits to how disunified a single stream of consciousness can become as well. Again, there probably are deep and important relations between co-consciousness and conscious unity. But there nonetheless *are* cases in which one's conscious experiences are disorienting and confusing in part because they're co-conscious, and it's confusing to experience them together: cases in which something *sounds* as if it's coming from behind you but *looks* as if it's coming from right in front of you; cases in which vision and sensation yield simultaneously incompatible information; cases of damage to early striate cortex in which, in the early recovery period:

the patient will see pure motion (usually rotary) without any form or colour. Then brightness perception returns as a pure Ganzfeld—a uniform brightness covering the whole visual field. When colours develop they do so in the form of 'space' or film colours not attached to objects. The latter develop as fragments which join together and eventually the colours enter their objects to complete the construction of the phenomenal object. (Smythies, 1994: 313)

Such cases need not necessarily be cases of possession of multiple streams of consciousness. In fact, again, the opposite would appear to be true: one can hardly

⁵⁵ I say this because it looks as if co-consciousness may sometimes be a causal precondition for conscious coherence, rather than the other way around: co-consciousness seems to *force* some degree, or at least some kinds, of coherence of conscious contents. Consider the so-called McGurk effect, for instance (McGurk, MacDonald, 1976). People who *watch* someone mouthing the sound "Ga ga" while simultaneously *listening* to someone utter the sound "Ba ba" actually *hear* someone say "Da da." But presumably if the visual representation of the mouth and the auditory representation of the vocalization were not co-conscious, this effect would not be produced. I don't actually know if anyone's run this experiment using split-brain subjects, but I would *expect* that if only the RH *watched* someone mouthing the sound "Ga ga," and only the LH *heard* someone utter the sound "Ba ba" (this would have to be set up as a dichotic listening task in order to work), the subject would *say* (LH) that she'd heard someone say, "Ba ba."

find disorienting the clash of sight and sound, for instance, unless the sight and sound are experienced *together*. Otherwise it is just sight and just sound, and no conflict.

So despite the no doubt deep and important conceptual and causal and perhaps even constitutive relationships between co-consciousness and (at least some elements of) what I have identified as conscious unity, co-consciousness and conscious unity also remain importantly distinct in certain respects. And meanwhile my sense is that a great deal of the literature on the puzzle of “split-brain” consciousness would benefit from drawing this distinction; thinkers (both philosophers and neuropsychologists) seem to talk past each other (or sometimes even past themselves) switching between stating that the two hemispheres *obviously* are associated with distinct streams of consciousness and puzzling over the fact that a split-brain subject’s consciousness sure doesn’t *seem* so disunified. I believe that a split-brain subject’s consciousness is dual and fairly unified, especially outside of experimental situations. (A “normal” subject’s consciousness is nonetheless no doubt more unified, given the kinds of integration of conscious contents that the corpus callosum allows.)

That said, one of the key things that makes our consciousness feel *unified* is *co-consciousness*: the fact that there is something it is like to experience lots of experienced things *together*. In fact the phenomenon of co-consciousness may be what is most puzzling about conscious unity. This is difficult to say without first being clear on all the elements of conscious unity, but co-consciousness is probably more puzzling and mysterious than at least *some* plausible aspects of conscious unity, such as the way consciousness relates to action and agency, or the coherence of mental contents. And I take it that what is most troubling about the split-brain cases,

where consciousness is concerned, involves co-consciousness in particular: both right and left hemisphere of a split-brain subject appear subject to conscious experiences, as evidenced through their right hemisphere-controlled and left hemisphere-controlled behavior; and yet right hemisphere experiences do not seem co-conscious with left hemisphere experiences.

3.5 Consciousness and Subjects of Experience

One standard description of (phenomenal) consciousness is this: a state is phenomenally conscious when there is something it is like for a subject to possess that state. This formulation links the concept of consciousness to that of a subject of experience. The concept of a stream of (phenomenal) consciousness or of co-consciousness may presuppose a subject of experience to an even greater extent.

The close conceptual link between co-consciousness and subjects of experience is made clear when, for example, in the course of defending the apparently dramatic thesis that disunity (multiplicity) of phenomenal consciousness is impossible, Bayne and Chalmers (2003) qualify that they can still accept, if they choose, Puccetti's view that split-brain subjects have two streams of consciousness apiece. For Puccetti also attributes two *subjects of experience* (and in fact two persons) to split-brain subjects, and as it turns out, Bayne and Chalmers only commit themselves to the claim that a single *subject of experience* can have no more than one stream of consciousness at a time. But if this is the case, then their "phenomenal unity thesis" is much less dramatic. For any time one has evidence (assuming there could be such evidence) that a creature possessed multiple streams of consciousness, one

could simply—in fact, one would have to, on Bayne and Chalmers’ view—accept that the creature was also associated with multiple subjects of experience.

As Dainton (2000) points out, Bayne’s and Chalmers’ is the standard way of individuating experiences: the token identity of an experience is standardly thought to depend upon the time at which it occurs, its phenomenal character, and its subject. But Dainton also says that while the first two criteria are both sound and *helpful*, the third criterion is “not very informative” (Dainton 2000: 25). This seems right, and perhaps the *reason* that the “consubjectivity thesis” (the thesis that co-conscious experiences belong to a single subject of experience) is not very informative is that our concept of a subject of experience is so tightly linked with our concept of a stream of consciousness. It is part of our concept of a subject of experience that that subject has no more and no less than a single stream or center of consciousness (at a time).

I suggest that our concept of a subject of experience is so tightly linked with our concept of (especially phenomenally) conscious of experience that we can’t really get any purchase on the idea of a subject of experience who has an experience that is conscious—but that the subject of experience herself or himself or itself can’t experience. We have to say, in that case, either that the experience belongs to some other subject of experience, and not that one, or else that the experience is in fact just a non-conscious representation. Therefore, if there is *nothing* it is like to possess two experiences or streams of consciousness at once, then those two experiences or those two streams of consciousness cannot belong to a single subject of experience. Which means that there must be *two* subjects of experience, one possessing each set of

experiences. Thus the consubjectivity thesis may be true, but true on the basis of our conceptual schema. While this is interesting, if correct, it will not help us *individuate* experiences: if you knew how many subjects of experience a given creature possessed or was constituted by, you would *already* know how many streams of consciousness that creature had. We have no hold on the concept of a subject of experience, that is, apart from our hold on consciousness and co-consciousness.

It is less clear if a single stream of consciousness can be associated with *more* than one subject of experience. I am imagining a case of telepathy in which two creatures temporarily appeared to share, to have equal access to, the same series of token conscious experiences. This might constitute a case in which *two* subjects of experience possess the *same* center or stream of consciousness. This is not clear, however. There appear to be two subjects of experience if subjects of experience are relatively enduring things; if instead there is a new subject of experience at each moment (as Strawson (1997) has argued), then perhaps during the telepathic time period there is but a single subject of experience. (It might also be argued that subjects of experience should be individuated in terms of their purely physical, nonfunctional properties, but this restriction might be unprincipled for telepathically communicating creatures.) Alternatively, perhaps such creatures would not really share a *token* stream or center of consciousness to begin with. In any event, while interesting, these questions don't appear to be essential to resolving matters concerning the structure of consciousness in the sorts of creatures we know of. I will therefore assume that the concept of a subject of experience is one according to which

such a subject does not possess more, or less, than one stream or center of consciousness at a time.

This, of course, still does not say what a subject of experience *is*. While I can't say anything that's illuminating on this question, I will assume that it is an important part of our concept of a subject of experience that *if* such a subject is presented with more than one experience, those experiences all belong to one stream of consciousness. We could, of course, revise the concept—but it's unclear what kind of *work* the concept would do if we did revise it in this way. What work is done by the concept of a subject of conscious experience who has conscious experiences she or he or it can't experience the relations between? Still, perhaps this is mistaken; in any event, subjects of experience enter little into what follows, and I don't try to individuate experiences or streams of consciousness by referring to the subjects of experience that possess them.

Finally, I distinguish a subject of experience from a person, at least conceptually—even for a creature who is associated with both a person and a subject of experience. The concept of a person is first of all richer than that of a subject of experience. Many non-human animals, for instance, may be subjects of experience, without being persons. One possibility is that a person is a subject of experience that/who also possesses additional properties. (E.g., self-consciousness, moral agency...) Or a person may be an *animal* that/who is a subject of experience and that/who also possesses additional properties.

A distinct possibility, however, and one that may be counter-intuitive but that arguably has a certain amount of appeal especially in the context of the split-brain

phenomenon, is that a person could, in theory, be identified with *more* than a single subject of experience, especially for limited periods of time, but perhaps even perpetually. It isn't clear that there is such a close conceptual link between persons and subjects of experience that a person couldn't have or be associated with more than one of the latter, particularly if the subjects of experience were subject to very much the same sorts of experiences, and assuming that the person's mental and behavioral life overall was characterized by a high or moderately high degree of integration. Of course, whether this were possible in fact would depend in large part upon the place of conscious experiences and streams of consciousness in cognitive architecture.

I am still trying to avoid personal identity questions to the extent possible in this work; this explains why I have, throughout, spoken of split-brain *subjects*, without meaning to presuppose any claims about how many persons or subjects of experience they are; in some of what follows I may speak of a (single) split-brain *subject* as being associated with or possessing two *subjects of experience*. Perhaps the best account of personhood is one according to which a person is a subject of experience, or perhaps not; again I wish to remain neutral on these sorts of questions to the extent possible.

3.6 What is Co-consciousness?

One of the several reasons that I accept the rough adequacy of the global workspace theory of consciousness is that I believe that it suggests a plausible account of co-consciousness as well. According to global workspace theory, two representations are

(access) conscious in virtue of both being available to global consumers. They may or may not be available to the same token global consumers: if the representations in question are mine and yours, then they obviously won't be available to the same token global consumers.

Presumably, for two representations to be access co-conscious is just for them to be jointly accessible or jointly globally broadcast. For instance, if an access conscious representation is one that can play a role in your practical reasoning, or one whose content you can report, then two access co-conscious representations can be *jointly* used in your practical reasoning, or you can report *both* of their contents.

I suggest that for two representations to be *phenomenally* co-conscious—for there to be something it is like to have both of those representations at once—is *also* just for them to be jointly accessible or jointly globally broadcast. This joint availability is what makes it the case that there is something it is like for their subject to experience both of them together—that there is a phenomenology of the two states together that differs from the phenomenology either of them has (or would have) alone.

This account needn't require that accessibility explains or confers phenomenal properties. Something else, besides global accessibility, may be necessary to account for the phenomenology that experiences like E1 and E2 have. . . . But their joint accessibility may nonetheless account for their *joint* phenomenology. So, again, for two representations to be co-conscious with each other—either access co-conscious or phenomenally co-conscious—is simply for them to be broadcast or available to the same set of (token) global consumer systems.

I should note that what makes a set of (token) global consumers constitute a set in the relevant sense here is not arbitrary; the set comprised of all of your global consumers and of all of my global consumers is not a set, in the relevant (psychological) sense, because there are no representations that are jointly available to your global consumers and to mine. Thus when I say that a split-brain subject, S, has two *sets* of global consumers—a right hemisphere set and a left hemisphere set—the division of all of S’s global consumers into two distinct sets of global consumers, a right hemisphere set and a left hemisphere set, is not arbitrary, and is not on the basis of purely physical features of those consumers somehow. The division is functional: RH conscious representations are broadcast to RH global consumers, LH conscious representations are broadcast to LH global consumers. So, we can use the term “f-set” (functional set) to refer to the set of global consumers that share access to the same representations. Note also that there is a sort of circular relationship between individuating token representations and individuating token global consumers (or sets of token global consumers). That is, one reason to identify right and left hemisphere, content-bearing neural events with *distinct* conscious experiences, is that right and left hemisphere neural events make their contents available to distinct global consumer systems. And one reason to identify right and left hemisphere consumer systems as distinct sets of consumer systems is that they have access to different conscious representations.

If this account of co-consciousness is correct, then phenomenal (and access) co-consciousness is a product of very much the same forces and mechanisms as those that produce access consciousness itself. Not all (conceivable) cognitive architectures

that can support consciousness as the global workspace theory describes need be capable of supporting co-consciousness, for we can conceive of a global workspace whose contents are always limited to a single representation at a time. But any architecture in which the workspace does take more than one representation at a time is ready to support not just access consciousness, but phenomenal co-consciousness as well.

This account of co-consciousness makes a prima facie case for conscious duality in split-brain subjects. The results of the split-brain experiments suggest that global broadcast is not inter-hemispheric in split-brain subjects; right hemisphere experiences appear broadcast to right hemisphere global consumers, left hemisphere experiences to left hemisphere global consumers. There is, of course, no reason to think that broadcast somehow becomes inter-hemispheric outside of experimental conditions. The conscious duality model for split-brain subjects will be defended against competing models of their consciousness in the next chapter.

I have assumed thus far that co-consciousness is an all-or-nothing affair—that multiple representations either are or are not co-conscious. If the account of co-consciousness that I have accepted is correct, however, then someone might argue that phenomenal co-consciousness can be *partial* in the following sense. There may be greater or lesser degrees of co-consciousness to the extent that there are greater and lesser degrees of *global broadcast*. And there may be greater or lesser degrees of global broadcast if broadcast can be to more or to fewer global consumers *within a set* of global consumers—not as a function of architecture, but as a function of more contingent matters (e.g., attentional focus). In the split-brain case, certain left

hemisphere representations are not broadcast to right hemisphere global consumers, and *could not be* so broadcast, not because of any attentional deficit, but simply because of a split-brain subject's mental architecture. But in some instances a representation might not be broadcast to a full set of global consumers not because of architectural limitations but only because of more contingent factors.

Of course, consciousness and co-consciousness would necessarily be an all-or-nothing affair in a creature that just possessed a single token "global" consumer. Imagine that C just has one system (let's say some "inner sense" faculty), availability to which makes a mental state conscious. (Though this would be a creature for whom or for which the *global workspace* theory would not apply, since conscious representations aren't broadcast globally but only to a single consumer system. Some other theory of consciousness would be necessary for this creature—like a higher-order representation theory. Of course this could have implications for the proper theory of *co-consciousness* in this creature as well, but let us ignore these points for the moment, since we are not such creatures in the first place.) Then if E1 and E2 are available to the inner sense faculty and E3 isn't, it is not just the case that E1 and E2 aren't co-conscious with E3; E3 simply isn't conscious. In such a creature, any conscious experience will necessarily be co-conscious with every other conscious experience.

We are not such creatures, however; the global workspace theory of consciousness posits *multiple* global consumer systems. So perhaps for us consciousness itself can come in degrees, greater degrees of consciousness corresponding to availability to a greater number of global consumers. (Something

like this idea can be found in Dennett (1991); Dennett's 1991 account of consciousness as a "Joycean stream" is not a version of the global workspace theory, but some of its elements are similar.) So imagine a creature with a whole suite of global consumers: a higher-order thought faculty, a practical reasoning system, a language system, an emotional system or systems, a memory system or systems, and so forth. In this creature, two representations that were available to *all* the same token global consumers would be perfectly co-conscious; two representations that were available to an overlapping but still distinct set of token global consumers would be partially co-conscious. So, if access consciousness comes in degrees, and if the account of phenomenal co-consciousness that I have accepted is correct, then phenomenal co-consciousness can come in degrees also.

Note that if access consciousness and thus co-consciousness comes in degrees in the way I have just suggested, co-consciousness could theoretically fail to be a transitive relation. As I said before, intuitively co-consciousness *is* transitive: if E1 and E2 are co-conscious, and E2 and E3 are co-conscious, then E1 and E3 are as well. But now imagine that E1 and E2 are co-conscious by virtue of both being available to two overlapping but distinct sets of (token) global consumers, and that E2 and E3 are similarly co-conscious by virtue of both being available to two overlapping but distinct sets of global consumers—but that the sets of global consumers to which E1 and E3 are both available do not overlap at all.

If access consciousness and thus co-consciousness regularly does come in degrees, then our consciousness is less structured, it would seem, than we might think. Instead of breaking into *discrete* sets of experiences among which perfect co-

consciousness holds, and between which co-consciousness does *not* hold, human consciousness may be more like a messy web of experiences, with some (co-consciousness) links strong and some weak and some simply absent.

But that consciousness *does* come in degrees in the way Dennett (2001, 1991) believes is uncertain; I take it that there is some evidence that broadcasting really is pretty much an all-or-nothing affair; a representation can be available to one or a few consumer systems. . . . Or it can be available to all of them. (See Dehaene et al., 2006.) Of course it is possible that a representation that is broadcast to all global consumers at once will only be *processed* by a couple of these systems, but this need not make it less *conscious* (though such limited processing may have the consequence of making it conscious for less time). If global broadcast does not come in degrees, then on my account, co-consciousness will not come in degrees either. Or at least not for *this* reason—for there is a different (though related) sense in which co-consciousness might be said to be partial, one that is particularly relevant to the split-brain cases. I won't assess the likelihood of this second form of partial co-consciousness until the next chapter, however, keeping the discussion deliberately simple for now.

Finally, as I said earlier, I deal only with synchronic co-consciousness in this work. But perhaps roughly the same sort of account could be given for both synchronic and diachronic co-consciousness. Consciousness takes time; an experience is broadcast to and processed by global consumers not just for an instant, but for a longer period. (Of course it is not clear how long, and of course it no doubt varies from experience to experience.) Its processing may be ongoing when a new

experience is broadcast. Meanwhile, the *represented* times of these experiences may of course overlap or be adjacent as well. Perhaps two experiences represented as occurring in quick succession are diachronically co-conscious when *they* are jointly broadcast to the same global consumers.

Of course, even a rough sketch of that sort of account faces numerous difficulties. I nonetheless suspect that something like the account of synchronic co-consciousness that I adopt could be made to work to explain diachronic co-consciousness also. (Among other things, the account of co-consciousness offered explains both why synchronic co-consciousness is (or tends to be) transitive, while diachronic co-consciousness is not.) But again, in this work I confine myself to a treatment of the structure of consciousness at a time.

3.7 Preliminary Conclusions and Assumptions

The next section of this chapter concerns the resources available to the defender of the conscious duality model to meet the challenge posed by split-brain subjects' normally unified behavior. Since a great many of claims, assumptions, and distinctions have just been laid out, however, let me recap them before moving on.

A preliminary distinction:

- A stream of consciousness is a set of mutually co-conscious experiences. A unified consciousness is a consciousness that meets a set of normative criteria especially concerning consistency, coherence, and perhaps agency.

Assumptions—I will not argue for these, at least for the most part:

- Both hemispheres of a split-brain subject are subject to some conscious experiences.

- Phenomenally conscious experiences are also access conscious.
- Global workspace theory offers an adequate account of access consciousness. (It may or may not offer an adequate account of consciousness, period.) For a representation to be access conscious is for it to be available or globally broadcast to *some* set of (token) global consumers.
- There is one stream of consciousness per subject of experience and vice versa.
- Knowing how many *streams* of consciousness a creature has does not yet tell us how unified the creature's consciousness is (and, albeit to a lesser extent, vice versa). A subject could have two streams of consciousness that nonetheless (together) exhibit a high degree of unity, or a single stream of consciousness that was nonetheless fairly disunified.

Implicit argument—while I don't explicitly argue for this claim, I hope that in using it fruitfully, I will provide some support for its truth:

- Global workspace theory offers an account of (access and phenomenal) co-consciousness. For two representations to be co-conscious is for them to be available to the same set of (token) global consumers.

Explicit argument—this claim will be defended in this and the following chapter.

- Split-brain subjects have two streams of consciousness, though these streams of consciousness are no doubt more unified (with each other) outside than inside of experimental conditions.

4 *The Conscious Duality Model and the Challenge from Unified Behavior*

There are two major models of the structure of split-brain subjects' consciousness.

The first is a "singularity model" of consciousness, according to which split-brain subjects have a single stream of consciousness, at least most of the time. Several different versions of this model are possible. In one such model, for instance, split-

brain subjects always have one stream of consciousness, associated with only their left hemispheres (Eccles 1965; MacKay 1966). In another version of this model (Marks 1980, Tye 2003) split-brain subjects *normally* have a single stream of consciousness, associated with both hemispheres, but *sometimes* have two streams of consciousness, one associated with each hemisphere, for brief moments during the split-brain experiment; in this model, the possession of the same conscious contents suffices for the hemispheres to share a single stream of experience.

The second major model of the structure of split-brain subjects' consciousness is what I have called the "duality model." In this model, a split-brain subject has two streams of consciousness, whose contents will be less or more similar as a function of whether the subject is inside or outside of the split-brain experimental situation, respectively. This model appears more favored by those in the neuropsychological community. (See for example Sperry 1990, 1968, 1966; Sperry, Zaidel, and Zaidel, 1979; Gazzaniga and LeDoux 1978; Bogen 1990, 1985, 1969; Uddin, Rayman, and Zaidel 2005.)

The major evidence in favor of the *singularity* model of split-brain consciousness is simply that split-brain subjects behave in such an integrated fashion under most circumstances. The major evidence against such a view is the dissociated behavior they sometimes display (usually during the split-brain experiment although sometimes at other times), in combination with the best *explanation* for that behavioral dissociation. The best explanation for the dissociated behavior split-brain subjects exhibit during the split-brain experiment is that they do indeed have two streams of consciousness. And because the best explanation for their having two

streams of consciousness during these experiments refers to the permanent embodiment of their mental and conscious lives, their neuroanatomies, it is difficult to believe that the split-brain experimental paradigm itself alters the structure of split-brain subjects' consciousness. There is better reason to believe that, their high degree of behavioral integration notwithstanding, split-brain subjects *always* have two streams of consciousness.

The duality model of split-brain subjects' consciousness is simply one according to which the two hemispheres of a split-brain subject are associated with two distinct streams of consciousness—two sets of token experiences, within each of which there are at least a great deal of co-consciousness relations and between each of which there are at most few co-consciousness relations. The duality model holds that a split-brain subject's two streams of consciousness are explained in part by her neuroanatomy: a split-brain subject lacks the neural pathways necessary to sustain interhemispheric co-consciousness.⁵⁶

That the duality model refers to neuroanatomical facts in this way, or at least that its application to split-brain subjects is partially explained by neuroanatomical facts, probably accounts for the popularity of the duality model within the neuropsychological community (in contrast to the philosophical community). Neuropsychologists are more likely to accept—or to simply assume—some kind of isomorphism between the structure of consciousness and the structure of the brain. As I have emphasized previously, this isomorphism is really between co-consciousness

⁵⁶ This formulation remains uncommitted as to the proper model of consciousness for acallosal subjects; perhaps in an acallosal subject some other structure, besides the (absent) corpus callosum, sustains interhemispheric co-consciousness. This formulation may imply that a “normal” subject has a single stream of consciousness, since it refers to “split brain” neuroanatomy in particular, but I actually wish to remain uncommitted as to the structure of consciousness in a “normal” subject in this work.

(and co-mentality) and *functional* neuroanatomy; it is largely because of observations made during the split-brain experiments themselves that the corpus callosum came to be viewed as a mechanism for co-consciousness (and co-mentality). Granted, if sectioning the corpus callosum had not led to any observable conscious dissociation, then neuropsychologists might still have insisted upon an isomorphism between conscious and neural structure—they might simply have hypothesized the anterior commissure as the crucial mechanism of interhemispheric co-consciousness, for instance. But, since the findings of the split-brain experiment were what they were, the corpus callosum instead was identified as that crucial mechanism. And *once* it was so identified, a conscious duality claim had to be maintained for split-brain subjects in extra-experimental situations.

Of course it is not universally accepted that, absent a corpus callosum, a split-brain subject lacks a means of achieving interhemispheric co-consciousness (conscious singularity). In the previous chapter we looked at an argument (the singularity-through-redundancy position) to the effect that mere *duplication* of conscious contents across the two hemispheres suffices for interhemispheric co-consciousness. And in the next chapter we will look at one model of split-brain subjects' consciousness according to which it is at least partially unified, or (to use the language of conscious singularity rather than unity), according to which a split-brain subjects' two hemispheres are at least partially co-conscious. We will also in that chapter look at a model of split-brain subjects' consciousness according to which a split-brain subject at least usually has only a single *intra*-hemispheric stream of

consciousness at a time—one sometimes located in the right hemisphere, and sometimes in the left.

But before we consider these alternative models of split-brain subjects' consciousness, and the challenges *they* pose to the conscious duality model, let us first examine the most obvious objection to the conscious duality model: the challenge posed by the generally integrated-seeming nature of split-brain subjects' behavior. For while the conscious duality model seems *prima facie* compatible with neuroanatomical evidence, it might also seem *prima facie* incompatible with behavioral evidence, for two reasons.

To begin with, the duality model holds that split-brain subjects' consciousness is dual both in and outside of experimental conditions, and yet split-brain subjects (in contrast to “normal” subjects) *behave differently* in the two types of circumstance. Since the subjects' neuroanatomies obviously remain constant, the duality model must also offer *some* account for subjects' behavioral changes between the two types of circumstance.

The duality model can easily meet this first challenge. The split-brain experimental paradigm is highly artificial, presenting numerous constraints not present in daily life. The absence of these constraints may allow subjects to compensate (not necessarily effortfully), to a large degree, and by a variety of means, for the lack of integration of conscious processes between the two hemispheres.

But the duality model must also explain why it is that, particularly outside but even *inside* experimental conditions, split-brain subjects really don't behave in *so* disunified a fashion, most of the time. This is a greater challenge. But as I attempt to

show in the next section, a defender of the duality model has adequate resources to meet this second challenge as well. For there are many mechanisms that may serve to integrate the behavior produced by the two hemispheres even if these mechanisms do not suffice to make the hemispheres co-conscious. The duality model can accommodate the behavioral data, in other words, even if accepting it means accepting that conscious *singularity* in and of itself does not play as large a role in producing unified behavior as one might have thought. This role may instead be played in large part by conscious *unity* as I have defined it, as well as by a variety of other non-conscious (and even non-mental) mechanisms and factors. I provide this argument in the rest of this section.

4.1 Duplication of Contents

Perceptual redundancy and resulting duplication of conscious contents across the two hemispheres have been favorite mechanisms for behavioral unification for duality theorists, partly because they would seem to explain why a split-brain subject's behavior changes between experimental and non-experimental conditions. The defining methodological feature of what I am calling the "split-brain experiment" is the attempt to *lateralize* perceptual information. Outside of these highly artificial experimental set-ups, the hemispheres have access to the same perceptual information (either at the same time or in very quick succession, due to saccades, etc.) to a significant degree. This is of course largely because they're in the same body, and always at the same place at the same time, and because they possess the same sensory modalities. It is also partly due to overlap in attentional mechanisms. For example, "The unity of the eyeball as well as the conjugate movements of the eyes causes both

hemispheres to automatically center on, focus on, and hence probably attend to, the same items in the visual field all the time” (Sperry, Gazzaniga, Bogen, 1969: 286). And they have equal (or similar) access to physiological states: those associated with hunger, with intense anger, with exhaustion, and so forth.

Note that perceptual redundancy is greatest outside experimental conditions, but impossible to eliminate entirely even within experimental conditions. Perceptual redundancy, then, offers a partial explanation both for the fairly unified behavior of split-brain subjects within experimental conditions, and the even more unified behavior of split-brain subjects outside of experimental conditions.

The hemispheres don't have *identical* perceptual access. Take vision, for instance. The right hemisphere is the only hemisphere to (consciously) represent most of the left visual field, and the left hemisphere is the only hemisphere to (consciously) represent most of the right visual field. It does appear that both hemispheres of a split-brain subject do learn to consciously represent an overlapping area at the visual midline, but this area is no more than two degrees wide (Fendrich, Gazzaniga, 1989; Fendrich, Wessinger, Gazzaniga, 1994). Within that narrow strip of the visual field that is represented by both hemispheres, meanwhile, “the signals conveyed to each hemisphere from the contralateral hemiretina appear to be weak or degraded” (Gazzaniga, 2000). So at any given moment, the two hemispheres will differ with regards to those qualities represented visually in conscious experience.

Granted, split-brain subjects do (when permitted) frequently move their eyes about the visual scene, resulting in the hemispheres acquiring “identical” visual information to some extent. It is still true that hemispheres often won't have the same

information at any one moment; instead, there is a staggering effect: the left hemisphere gets information about the forest while the right hemisphere gets information about the girl playing; then the left hemisphere gets information about the girl playing while the right hemisphere gets it about the shore; then the left hemisphere gets it about the shore while the right hemisphere gets it about the boat at the center of the lake, etc.⁵⁷ But it is also true that our conscious representing of visual scenes in general may be built up over many moments, by a series of saccades, and so maybe over the course of several seconds the hemispheres will receive the same visual input from the world.

Even if the conscious *visual* contents of both hemispheres were identical, the hemispheres would still differ in terms of things like their conscious *tactile* contents⁵⁸. When S's blindfold is removed, both hemispheres can visually represent that S is holding both a flashlight and a book of matches, but only the RH also *feels* the flashlight, in a detailed way, and only the LH also *feels* the matches, in a detailed way. And S will not have any rich tactile representation of the flashlight and the matches at once. Temperature (Gazzaniga 2000) doesn't appear to be experienced in precisely the same way by the hemispheres; Dimond (1980) even says that some split-brain subjects report having burned their left hands for several moments before "they" (their left hemispheres) realized it.

Finally, the two hemispheres also of course differ in linguistic capacity. While there is still significant debate about the linguistic capacity of the disconnected right

⁵⁷ In making this point I am not conflating time of representing and time as represented: the claim isn't merely that hemispheres represent, at different times, the same qualities, but that they will also represent the same qualities as existing (or being experienced) at different times.

⁵⁸ Since stereognostic or "active touch" information for the hands is sharply lateralized in split-brains (Gazzaniga 2000; Gazzaniga, Bogen, Sperry 1963).

hemisphere, and about the connected right hemisphere's contributions to language, it is less controversial to state that split-brain subjects' left hemispheres are generally responsible for the production of propositional speech. Since the right hemisphere is nonetheless often capable of producing non-propositional vocalizations (Code, 1997), the most reasonable hypothesis concerning the *inner* speech of split-brain subjects is that the left hemisphere engages in it and the right hemisphere does not. Inner speech of course has a phenomenology.

Inner speech isn't something out in the world to be perceived, of course; it's a product of the hemispheres' perceptual (and conceptual) systems. This suggests that, while the hemispheres no doubt do enjoy a good deal of perceptual redundancy, especially outside but also inside experimental conditions, it is important to distinguish between identical (or highly similar) perceptual *access* and identical (or highly similar) perceptual *contents*. For two hemispheres (or for that matter, two creatures) to have identical perceptual access is for them to be able to access, via perception, information about the same portions of the world. (And also they should be able to access this information via the same sensory modalities; it isn't the case that I and a bat in my attic have identical perceptual access to the moths I see and the bat hears (or echolocates).) Thus you and I have the same perceptual access if we're both free to look about the same room. The split-brain experiment denies the two hemispheres equal perceptual access to the world: it displays to the right hemisphere what it conceals from the left. Outside of the split-brain experiment, however, a subject's two hemispheres will have similar (albeit not identical) perceptual *access*.

Indeed in some cases the hemispheres may acquire the same perceptual contents even when they lack the same perceptual access. Sperry (1990) notes that “By the laws of perceptual completion and closure each hemisphere automatically tends to fill in its half stimulus across the midline to form a whole bisymmetric percept on each side” (Sperry 1990: 375, citing Trevarthen 1976). By presenting a split-brain subject a (bilateral) chimera, the subject can “be made to perceive two quite different things occupying the same position in space at the same time, something rejected, of course, by the normal brain” (ibid)—but such chimeras (or their equivalents) are rare in nature, and thus this perceptual completion provides one means by which contents are duplicated in the two hemispheres, even *without* identical perceptual (or sensory) access.

Still, this point, that perceptual contents aren’t only a matter of sensory input, cuts both ways. The hemispheres appear to process sensory information of various sorts somewhat differently (and also appear capable of maintaining somewhat different focuses of attention).

For instance, there seem to be some differences between the two hemispheres with regards to emotional perception and experience⁵⁹. At the very least, to the extent that the cognitive component of emotion (Schacter, Singer 1962) involves language and hypothesis-formation (as to the cause of the physiological element of emotion for example), the hemispheres will almost certainly differ with regards to the cognitive

⁵⁹The preponderance of studies in this area suggest that there is *some* difference in the way the hemispheres perceive, process, and express positive versus negative emotions, but there is not yet a clear conceptual model of this difference. Very roughly, though, the data suggest the right hemisphere has some advantages for some aspects of the conscious perception of negative emotions, and vice versa. See for example Lee et al. (2004); Smith et al. (2004).

content of emotional experience⁶⁰. But the hemispheres may differ even with respect to the non-cognitive aspect of emotion; there is some evidence that (at least in many people) the right hemisphere experiences more negative affect. (See Schiffer 1998 for a very dramatic version of this thesis.)

It seems very likely that the visual processing differences between the two hemispheres result in phenomenal differences as well. That the right hemisphere excels at representing (at least some sorts of) spatial relationships between visually perceived objects, for example, suggests that the right hemisphere's phenomenal experience of the visually-perceived world may differ from the left hemisphere's. The left hemisphere also appears more sensitive to local visual configurations and the right to global physical patterns. (See Hellige, 1993, for discussion of hemispheric asymmetries in the visual and visuospatial realm. Also see Zaidel 1990 suggesting that hemispheric differences in the storage and retrieval of long-term semantic memory may underlie some of the perceptual asymmetries between the hemispheres.)

These sorts of differences in perceptual processing presumably have consequences for phenomenal contents. So even identical perceptual access would not necessarily mean identical perceptual contents. This point has probably not been sufficiently acknowledged by supporters of the conscious (and mental) duality models.

Finally, all that said, it is again important to emphasize that (contra the singularity-through-redundancy position discussed in Chapter Three) even if the

⁶⁰ Schiffer et al. (1998) even elicited different answers to emotionally salient questions from the right and left hemispheres of two split-brain patients, although it is of course possible that the hemispheres actually experienced the same emotions but reported them differently.

hemispheres were subject to identical conscious contents, this still would not mean identical perceptual *states*. As Sperry said:

Many unifying factors can be enumerated that tend to make at least components of the inner experience of the two disconnected hemispheres similar or identical in orientation and content especially during ordinary unrestricted behavior. The bilateral sensory projection systems of the brain, like those for cutaneous sensibility of the face, ensure a bilateral reduplication of symmetric sensations in both hemispheres. Thus, with conscious attention focused on facial, auditory, or other stimuli that get bilateral representation, both hemispheres presumably develop a full bilateral percept with no vertical split between right and left aspects. Scanning movements of the eyes yield a similar duplication with respect to vision. *The overall effect in some respects is thus more like a twinning or doubling of the conscious domain of the self rather than midline division.*

(Sperry 1990: 380, emphasis added)

This last point reflects why I have, throughout, spoken of the conscious *duality* model rather than of the conscious *disunity* model of split-brain subjects' consciousness—and it also reflects why Sperry begins the passage quoted above by speaking of *unifying* factors. There is no question but that, where content is concerned, the two streams of consciousness are to a large degree similar, and to an even larger degree coherent. Note that these are two different things: not every difference or inconsistency in contents need signify incoherence. When the conscious contents of the two hemispheres differ, they may do so, in many cases, simply because the two hemispheres are conscious of different portions of the world, or different aspects of the same portion of the world. There *are* cases in which their contents will be incoherent; think again about the chimera case. But again, such things are rare. Normally, by each independently practicing “perceptual completion”

of midline stimuli, the hemispheres will generate representations of midline visual stimuli with similar contents.

4.2 Attention

The literature on attention in split-brain subjects is complex and at time seemingly contradictory. This may in part be because of the difficulty in distinguishing between various types of attention. I will try to quickly give the major findings here before discussing in slightly more detail a couple findings that illustrate the need to distinguish between integrated or adaptive-seeming behavior and behavior that stems from mental and conscious singularity.

The hemispheres compete for processing resources: though they can perform different tasks at the same time (this is true to some extent in “normal” as well as in split-brain subjects), increasing the difficulty of one hemisphere’s task will impair the performance of the other. (Holtzman and Gazzaniga, 1982; Reuter-Lorenz et al., 1996.) While this sharing of (competition for) processing resources won’t serve to integrate the activities of the two hemispheres per se, it may still encourage unified behavior, insofar as it makes it difficult for one hemisphere to “run away with” a task while the other is also deeply involved with a task.

At the same time, the division of resources between the hemispheres can actually aid their performance for some tasks. Holtzman and Gazzaniga (1985) gave split-brain and “normal” subjects a complex spatial memory task, in which information critical to the task was presented in both visual fields. For “normal” subjects, the visual information was combined, and as a result the hemispheres confronted one big problem; for split-brain subjects visual information from the two

fields was of course not combined, with the result that each hemisphere had to solve a problem only half the size. The split-brain subjects outperformed the “normal” subjects.

The two hemispheres may engage in visual search independently; a split-brain subject is twice as fast as scanning, say, eight stimuli when the stimuli are divided across the two visual fields as she is when all eight stimuli are in the same visual field (Luck et al., 1989, 1994). At the same time, Holtzman et al. (1983) found evidence suggesting that a split-brain subject cannot divide her spatial attention across the two hemispheres—directing attention to two different points, one in each visual field, at once, for instance. Using a very similar paradigm, however, Mangun et al. (1994) obtained the opposite result: split-brain subjects *did* appear able to direct spatial attention to both visual fields at once. Arguin et al. (2000) argued that this discrepancy could be explained by the fact that in the Holtzman et al. (1983) study the onset between cue and target was 1500 msec—a fairly long interval, especially since that is a length of time sufficient to allow information for the control of visuo-spatial attention to transfer between hemispheres. (Either hemisphere can direct attention to a point in either visual field.) This may have prevented the hemispheres from operating as autonomously as they otherwise could. And indeed with shorter interval times Arguin et al. (2000) again found that at least most of the split-brain subjects they used in their study could divide their attention between two visual fields, while their “normal” subjects could not.

There has also been some suggestion that voluntary attentional orienting involves inter-hemispheric competition even in the split-brain subject (Kingstone et

al., 1995), with attention lateralized to the left hemisphere, while automatic attentional orienting can proceed somewhat independently in the two hemispheres (Luck et al. 1989, 1994). (Note, though, that Arguin et al. (2000) believe the results of their study shows that even voluntary attentional orienting can proceed independently in the two hemispheres. In the Kingstone et al. (1995) study suggesting the opposite conclusion, the task required not only voluntary directing of attention but strategic directing of attention; perhaps this accounts for the difference?) And there is some evidence (not just from split-brain subjects) that the right hemisphere is dominant for visual spatial attention insofar as it attends to the whole visual field, while the left hemisphere attends just to the right visual field.

It is likely that none of the findings just cited are, individually, secure at this point. Nonetheless we may be able to draw, with some confidence, the conservative conclusion that the hemispheres compete for attentional resources to some extent, while possessing (or at least operating) some attentional mechanisms independently. (Of course, it is possible that *attention* is being used too broadly in the literature on attentional mechanisms at this point, and that some of what looks, now, like competition for attentional resources will eventually be conceptualized as competition for something else, and, similarly, that some of what looks to be the independent exercise of attention now will come to be conceptualized as the independent exercise of something else.) Obviously to the extent that the two “disconnected” hemispheres do share and compete for attentional mechanisms, this could play some role in explaining split-brain subjects’ apparently integrated behavior.

More interestingly, behavior that seems integrated—smooth, coordinated, adaptive—could in part be the product of a *lack* of integration of attention. Let me illustrate this idea using two further studies as an example.

Ellenberg and Sperry (1980) designed a study in which split-brain and “normal” subjects were asked to engage in a variety of sorting tasks, sometimes unimanually and sometimes bimanually, to see how well they could perform different tasks with the two hands (and thus presumably the two hemispheres) at once. All four conditions involved taking small objects from a central container, and then putting each object onto either a top or a bottom shelf, right of the central container for the right hand, left of the central container for the left hand. The objects used in the study were cylindrical beads (C), spherical beads (S), wing nuts (W) and hex nuts (H).

For instance, in the unimanual task subjects sorted the objects in the following way, one hand at a time:

C	S
S	C

In the bimanual-same condition, subjects were asked to sort using both hands at once, but the sorting task was the same for both hands, for example:

S	S

C	C
---	---

In the bimanual-reverse task, they might be asked to sort as follows, using both hands at once:

C	S
S	C

And, finally, in the bimanual-different task (in which the objects were drawn from two central containers, one for the right and one for the left hand), they might be asked to perform this sorting task, again using the two hands at once:

S	H
C	W

Ellenberg and Sperry made several interesting observations. To begin with, in “normal” subjects (and also subjects who had undergone only partial callosotomy), the two hands always moved at the same rate in all the bimanual conditions. At least several of the full callosotomy (split-brain) subjects, however, began sorting at different rates for different hands.

“Normal” (and partial split-brain) subjects were (interestingly) particularly impaired at the bimanual-reverse task, even compared to the bimanual-different task, at least at first; with practice their performance on the two sorts of tasks equalized. But they were always significantly slower at both of these tasks than they were at the unimanual or the bimanual-same task, and their hands always moved in synchrony, and, moreover, whenever they made a mistake with one hand, they tended to simultaneously slip up with the other. (Ellenberg and Sperry speculate that, since “normal” subjects get faster at the bimanual-different and especially the bimanual-reverse sorting tasks, but since bimanual synchrony was never broken, “normal” subjects might improve not because they get better at lateralizing attention, but rather because the task becomes more and more automated with practice, requiring less attention.)

It did not matter to the (full) split-brain subjects which sorting task they were asked to perform,: they performed at about the same rate, and with the same number of errors, regardless of condition, and, again, their hands sometimes sorted at different rates, with no ill consequences. As Ellenberg and Sperry note:

The present findings support the idea that attention is normally limited to a unitary focus and that the cerebral commissures act to keep the left and right hemispheres working together in a single unified attentional system. This bilateral integration of attention appears to be a dominant organizing principle in cerebral function such that the intact hemispheres have great difficulty executing two simultaneous tasks independently despite lateralization of stimuli and responses.

(1980: 416)

So, split-brain subjects perform tasks that “normal” subjects really can’t, and that are therefore said to involve conflicting motor plans (suggesting that motor plans and spatial representations of movements are isolated to each hemisphere) (Franz, Ivry, Gazzaniga, 1996). For instance, a “normal” subject can draw two pictures at once, one with each hand, only if the figures to be drawn are either identical or mirror reversed. (So, a normal subject would not have much difficulty drawing an “O” and an “O” at once, or a “b” and a “d,” but would have trouble drawing a “t” and a “y” at once.) But a split-brain subject can draw two different figures at once.

Now, I accept that the explanation for this difference has to do with some kind of internal, inter-hemispheric unity or integration in the “normal” subject that is missing in the split-brain subject. The first study, for instance:

support[s] the idea that attention is normally limited to a unitary focus and that the cerebral commissures act to keep the left and right hemispheres working together in a single unified attentional system. The bilateral integration of attention appears to be a dominant organizing principle in cerebral function such that the intact hemispheres have great difficulty executing two simultaneous tasks independently despite lateralization of stimuli and responses.

(Ellenberg and Sperry, 1980: 416)

And there is furthermore *one* sense in which the behavior of a “normal” subject during the sorting tasks would look “more unified” than the behavior of a split-brain subject: in the “normal” subject, the movements of the hands would remain coupled (synchronized), even if their sorting tasks were different.

In some other ways, though, I suspect that the split-brain subject’s behavior would look just as unified. Each hand would go about its task smoothly—as the

subject, by the way, sat still and participated cooperatively—and an error by one hand wouldn't disrupt the whole process. Of course, if you ran the experiment on a “normal” subject first and *then* observed the split-brain subject's uncoupled (de-synchronized) hand movements, you might find the latter eerily “disunified.” But I think you would only find it so because, observing the behavior of the “normal” subject, you constructed a theory of the internal process and architecture that would result in both hemispheres halting when one slipped. On its *surface*, though, the behavior of the split-brain subject would look perfectly adaptive—even, in a sense, coordinated: coordinated with the requirements of the tasks and the aims of the subject (as a whole).

Or consider the second study. Imagine telling two subjects to draw a house with one hand a tree with the other, at the same time. One subject sits down and does it, without too much difficulty. The other subject stops and starts a few times, then fails to follow instructions by letting his hands “take turns,” until, when gently chided for this, he says in frustration, “But this is impossible.” Note, again, that the performance of the *first* subject might strike you, *on its face*, as being more smooth, coordinated, adaptive, integrated. It's the second subject who seems to be coming undone.

Of course, it isn't the surface that should interest serious psychology: it's the first subject and not the second who has two minds and two streams of consciousness. The point is just that the day-to-day behavior of a split-brain subject, behavior that looks, again, adaptive, coordinated, smooth, and so forth, may result from inter-hemispheric division and independence rather than unity and integration.

4.3 Cross-cuing

One series of experiments by Sperry, Zaidel and Zaidel (1979, “Self recognition and social awareness in the disconnected minor hemisphere”) illustrated the fairly extraordinary capacity of the disconnected left hemisphere, at least, to observe and correctly interpret right-hemisphere controlled behavior, thus bringing the two hemispheres into some kind of “agreement” about what subjects had experienced even if only their *right* hemispheres had experienced something to begin with. (Though, admittedly, the “mind-reading” exhibited in this study relied not only on the right hemisphere’s cross-cuing behavior but also on affective states generated in response to emotionally significant right hemisphere experiences.)

Bogen (1990) gives several fascinating examples of cross-cuing in the patient L.B. for instance:

1) A pipe was placed in his left hand (RH); LB turned it around in that hand until it was in a normal pipe-holding position; he then lifted it up to (but not into) his mouth, and then said (LH), “Oh, that’s a cigarette.” Feeling it with his right hand (LH) he then correctly identified it as a pipe.

2) When a pair of glasses was put into his left hand (RH), L.B. opened the arms of the glasses and then had to be prevented from sliding them onto his head. He was told to try to identify the object without manipulating it further, and couldn’t. Then he snapped the arms closed and quickly said, “Are those your glasses?”, presumably from the auditory cue.

3) A handkerchief was placed in his left hand (RH). He squeezed it, turned it over, and then “smoothly reached backwards and slipped it into his left hip pocket. At this point he said, ‘Oh, sure, that’s a comb.’ When told he should feel it with his right hand, he did and then shook his head with a chagrined look and said, ‘A handkerchief.’” (Bogen 1990: 218)

4) Bogen (1990) also cites a case in which Butler and Norrsell (1968) tested L.B.'s ability to name objects felt with his left hand (RH), using a wooden sphere, a wooden cube without sharp edges, and wooden pyramid. While L.B.'s hands were hidden from his sight behind a screen the experimenters would place one of these items in his left hand and ask him which he was holding. (Presumably L.B. knew beforehand that these were the choices.) To their surprise, he identified them correctly most of the time. At first Butler and Norrsell suspected L.B.'s hemispheres were exploiting some kind of subcortical communication, but then they noticed that their subject was looking around the room in a systematic way during the experiment. Whenever they put the sphere in his left hand he would look at the clock before saying, "It's the round one", and whenever they put the cube in his left hand he would look at the door before saying, "It's the square one." When they put the pyramid in his left hand he would look up at the ceiling for a few moments before saying, "It must be that triangular shape." (Bogen 1990: 218). Once blindfolded, his performance fell to chance.

Of course, not all cross-cuing is so deliberate-seeming. When a split-brain subject is asked a question to which only her right hemisphere knows the answer, and the subject (LH) makes an incorrect guess, she will often frown, immediately after which she will attempt to correct herself. In such cases the frown is presumably initiated by the RH, but surely not as a deliberate attempt at communicating its displeasure to the LH.

On a loose understanding of "cross-cuing," split-brain subjects could be engaging in it all the time. Given the hemisphere's equal (or nearly equal) access to proprioceptive information, given either hemisphere's ability to see what every part of the body is doing in most circumstances, *any* behavior that is initiated by one hemisphere and that the other hemisphere then gives some intentional interpretation,

could qualify as a form of cross-cuing. But this could mean that the hemispheres were “cross-cuing” each other most of the time.

4.4 Affect... and Aura?

Affect constitutes another potential source of behavioral unification in split-brain subjects; the hemispheres have some kind of shared access to affective information. As noted in Chapter Three, it is not clear whether the hemispheres, by virtue of intact subcortical structures, actually share *token* affective (etc.) states, or whether instead, by virtue of their mutual connectedness to intact subcortical and non-cortical structures, the hemispheres are merely caused to enter into similar (types of) affective states. As Sperry says:

Whether the neural cross integration involved in the foregoing as, for example, that mediating emotional tone, constitutes and extension of a single conscious process or is better interpreted as just a transmission of neural activity that triggers a second and separate bisymmetric conscious effect in the opposite hemisphere remains open at this stage. (1990: 380)

It is easy to think of an affective state as something somehow discrete, sharp, or specific: a wave of anger, a pang of regret, a burning desire, a thrill of excitement, and so forth. But much more *diffuse* (and difficult to characterize) states, with some kind of emotional or affective valence to them, seem to transfer between the hemispheres as well. Sperry notes that “cross-integration systems of the intact brainstem. . . that mediate a prompt bilateralization of emotion generated unilaterally” presumably also involves “mood, alertness, and perhaps subtle shades of these as in the more elemental dimensions of mental sets and attention” (Sperry 1990, 380).

At the same time, the inter-hemispheric transfer of these phenomenally diffuse states seems to sometimes co-occur with or perhaps include the transfer of certain contextual, categorical or conceptual information. This is information that one would expect to have some kind of emotional valence, perhaps—but one wouldn't necessarily expect this information to be conveyable in the form of an affective state. Thus See Sperry, Zaidel, and Zaidel (1979) refer not just to the inter-hemispheric transmission of affect but also of *aura*; other times people speak of “general mental set” (see for instance Sperry 1990, p. 380) or gist. The exact or even fairly rough nature of this “aura” isn't at all clear; among other things it is quite possible that there isn't *one* mental mechanism or property that's being picked out by terms such as “aura” and “mental set” and “gist.”

In cases where the transfer really is of a state with some emotional valence, the anterior commissure may be involved; Gazzaniga and LeDoux (1978) explain that “the human anterior commissure derives its fibers from the temporal lobe and from subcortical ‘limbic’ structures, in particular, the amygdala, and projects to the same regions in the other hemisphere” (153, emphasis added). These are structures involved in emotion, of course, but they are also involved in emotional or affective components of memory, and in memory generally, since the temporal lobe contains the hippocampus, and since it is implicated in semantic/conceptual knowledge.

The transfer of affect and aura probably plays some role not just in unifying the behavior of split-brain subjects, but in providing their two streams of consciousness a unified character, at least to some extent. As mentioned in Chapter Three and in Section 4.1 above, the hemispheres “often differ in their emotional

propensities” (Doty 1989: 70). But even if they are often not in the same emotional state, this does not mean that they can be in wildly different emotional states. In part, emotional states are physiological (or at least many emotional states are) states, with physiological markers: racing hearts, shallow breaths, tense muscles, and so forth. And of course the hemispheres are located in and have similar perceptual access to the same body. So they would be likely to enter into similar affective states even if only via a process that supervened in part on viscera, muscles, etc.

On the heels of discussing sub- and non-cortically mediated mechanisms that generate similar affective states and arousal levels and so forth in the two hemispheres, Sperry cautions that these “bilateralizing and unifying mechanisms are largely of the nonfocal general background category, whereas conscious awareness tends on the other hand to be correlated predominantly with attentional and focal aspects of cerebral function” (1990: 380). In part given the inherent difficulty of investigating such diffuse, non-specific states, however, it would not be surprising if we currently fail to appreciate how deeply and how constantly these diffuse background states color those states at the forefront of our consciousness and attention.

4.5 Same Body

There is of course a limit—a hard one—as to the degree to which a single-bodied creature can behave in a disunified manner. There are numerous types of constraints posed by shared embodiment. Though the hemispheres may initiate distinct actions simultaneously, these must at least share an overall postural set: one can’t jump and

bend down at the same time, for instance. The two eyes and of course the two retinal hemifields move together. And so forth.

Ultimately, of course, however many minds a creature has, those minds must realistically converge and be converged upon by pathways to and from the same body. Conflicting motor/behavioral plans will thus need to be resolved at *some* point prior to the actual initiation of action (at least in cases where the actions are physically, not just intentionally, incompatible), and one mind will receive a much richer degree of sensory information about the actions initiated by the other/s than is possible between multiple creatures. Bogen (1990) for instance points out that, “if one hemisphere initiates motion, the other hemisphere receives considerable *proprioceptive* feedback” (217; emphasis added).

And both hemispheres can also control the whole body; both hemispheres can initiate walking or stopping, bending or stretching, facial expressions and eye movements. The two hemispheres may not function entirely identically with respect to control of the body, but each has sufficient control over most of the body such that actions initiated by one hemisphere can look perfectly “normal” i.e. not disunified. Even for the parts of the body where the hemispheres are least functionally identical with respect to motor control, i.e. the hands, there is a pretty good amount of ipsilateral motor control is still possible (Kingstone and Gazzaniga, 1995; Miller and Kingstone, 2005.) (This may be more true in split-brain than in “normal” subjects, of course.)

Most simply, but probably also most profoundly, *bodies* are spatiotemporally defined, and the two streams of consciousness of a split-brain subject arise from two

minds that have always been together. As Sperry, Gazzaniga, and Bogen once put it, “these two separate mental spheres have only one body so they always frequent the same places, meet the same people, see and do the same things all the time and thus are bound to have a great overall of common, almost identical, experience” (1969: 286). While each hemisphere may have its own mental life, to which the other hemisphere does not have (for example) introspective access, the split-brain *subject* still has but one life to lead.

This is no doubt of profound importance to explaining the degree to which the hemispheres of a split-brain subject are at peace with each other. For it is probably true of all of us that the integrated nature of our behavior, when it is so integrated, owes in large part to constancies and consistencies in our lives, and our surroundings. There is all sorts of evidence from social psychology, for instance, that suggests that a person’s behavior, on a moment-to-moment basis, is shaped by situational factors to a much greater extent than one might believe; in fact a person’s behavior on a moment-to-moment basis might be shaped *more* by situational factors (and of course general psychological abilities, ones common to people generally) than by any idiosyncrasies of her mind/s or her person.

While shared embodiment clearly plays an important role in accounting for the degree of behavioral unification split-brain subjects exhibit, shared embodiment exists both inside and outside experimental situations, of course, and therefore can’t account for the *change* in subjects’ behavior between the two types of circumstance. It is instructive in this regard to note (as Nagel 1979 long ago did) that split-brain subjects really behave in a fairly integrated manner even within experimental

situations. All the mechanisms referred to in this section are available at least to *some* degree during the split-brain experiment itself. So, however disunified split-brain subjects' behavior appears during the split-brain experiment, that behavior is still even at those moments subject to numerous unifying forces—forces that nonetheless, I have argued, don't suffice to create a single stream of consciousness.

4.6 Same Brain

In his early work on the split-brain phenomenon, Sperry (1964) found that:

although deep surgical bisections are possible experimentally that include the roof of the midbrain, the supramammillary commissure, and even the cerebellum, it was sufficient merely to cut the forebrain commissures that mediate cross communication between the hemispheres proper to prevent interhemispheric transfer of perceptual learning and memory.
(Sperry, 1990: 371)

But even very “deep” levels of the nervous system—all the way down to pattern generators in the spinal cord—will play a significant role in integrating behavior. The two hemispheres of a split-brain subject remain connected to the same peripheral nervous system and spinal cord, the same cerebellum, the same cranial nerves, the same brainstem and midbrain and diencephalic structures. This is not to say that those structures serve as a means of inter-hemispheric mental integration; many of them don't. But even *when* they don't there are numerous ways in which the hemispheres' mutual connection to those structures may unify split-brain subjects' behavior.

Bogen, for instance, notes that the two hemispheres share the same blood supply and the same cerebrospinal fluid. This means, among other things, that if one

hemisphere causes a particular hormone to be secreted, the other hemisphere will also feel its effects. Hormone transmission or communication may be relatively slow, but nonetheless “important for mental states” (1990: 217). And presumably because of the brainstem ascending reticular activating system, the two hemispheres will be on the same sleep-wake cycle. (I only know of one report by a split-brain subject of one hemisphere being awake while the other slept (Dimond 1980). And, even in this case, the restless right hemisphere apparently soon slapped the patient’s left hemisphere awake.)

4.7 Single Hemisphere Control

Gazzaniga and LeDoux point to evidence that “potential homolateral pathways may lie dormant or undeveloped in the normal brain” (1978: 36) but develop and become more functionally active following callosotomy. These pathways might make cooperation between the hemispheres less necessary, by allowing both hemispheres greater individual control over (and perceptual access to) the whole body. They write, of these “shifting circuits” and of cross-cuing, that, “It is as if the brain demands integration, and in the absence of the interhemispheric pathways, less efficient ways of achieving mental unity are employed” (Gazzaniga and LeDoux 1978: 39). But what is most clearly achieved, when the hemispheres achieve greater degrees of motor control over ipsilateral pathways, is behavioral unity. Perhaps some mental unity, in the sense of coherence or cooperation, is achieved as well—but not mental or conscious *singularity*.

It is also possible that non-cortical structures become better, over time, at coordinating the outputs of the two hemispheres (even if this simply means inhibiting

or ignoring one set of outputs). And meanwhile Wolford, Miller, and Gazzaniga (2004) write that that over time working with split-brain subjects one gets the sense that one hemisphere often simply “defers” to the other hemisphere. In fact they suspect that it is difficult to get one hemisphere to even try very hard on a task for which the other hemisphere is (or is at least “believed” to be) superior or dominant.

Sperry (1990) believes that activity in one hemisphere may at times even *depress* activity in the other. (Perhaps because the hemispheres compete for some attentional resources?) At the same time, he notes that this is most likely under experimental conditions involving “prolonged use of a single hemisphere with deprivation of focal input to the other” (380).

In ordinary unrestricted behavior, on the other hand, it is rare that conditions would thus selectively restrict sensory input or central processing to one hemisphere for an extended period. Thus, typically, the two disconnected hemispheres appear to be actively, but separately, conscious in parallel, each working and contributing in its own way to the performance on which attention is focused. (Sperry 1990: 380-381)

Gazzaniga (1985) is a fiercer advocate of general left hemisphere dominance, claiming that when the right hemisphere has no language, it also displays:

a striking inactivity that borders on behavioral tedium. This is not to say these right hemispheres do not have specialized systems. They may have, but it is next to impossible to demonstrate their existence in a brain system that is so unable to behave overtly.

(Gazzaniga 1985: 70)

Others (Levy, 1983), just as adamantly contest the claim of general right hemisphere “passivity.” But even if neither hemisphere is *passive* much of the time, one hemisphere could still *dominate*, to some extent, much of the time. It could be that the left hemisphere usually dominates. Or, again, it could be that the hemispheres take turns dominating, depending on what sorts of tasks the subjects are engaged in.

Note that if it *were* the case that, usually, in a split-brain subject, only one hemisphere (or only one hemisphere at a time) dominates behavior, and this was partly responsible for the subject’s generally integrated behavior, mental (and conscious) singularity would *in a sense* be the explanation for that behavior: the behavior in question would be integrated (in part) because it was the product of a single mind and a single stream of consciousness. But this would not mean that the subject only possessed one mind or one stream of consciousness. It would just mean that at any given time the existence of one mind and one stream of consciousness was hidden from us.

4.8 Owing Actions

One interesting area of research and debate in the broad area of consciousness studies in recent years concerns conscious will (see Wegner 2002 for instance), and, relatedly, the *manner* in which people come to posit (correctly or incorrectly) the mental causes of their behavior. One element of this discussion concerns human beings’ “mindreading” capacities, and to what extent our self-knowledge is simply result of our turning those capacities inwards, in a form of self-interpretation. Carruthers (forthcoming), for instance, argues that we have no introspective or conscious access to our own propositional attitudes—beliefs, desires, judgments, and

so forth—but only to our perceptual states. In defending this position he draws upon some of the split-brain data. This is data suggesting that, instead of acknowledging ignorance about the cause of some (RH-controlled) behavior, the subject (LH) will often instead simply invent a reason, apparently swiftly and smoothly and with no sense of engaging in fabrication.

Gazzaniga (1985) has argued for years that the left hemisphere houses an “interpreter” that strives to tell a unified story of the self, a story about a unified, and rational, conscious and mental agent—even when there *is* no such thing:

The same split-brain research that exposed startling differences between the two hemispheres revealed that the human left hemisphere harbours our interpreter. Its job is to interpret our responses—cognitive or emotional—to what we encounter in our environment. The interpreter sustains a running narrative of our actions, emotions, thoughts, and dreams. The interpreter is the glue that keeps our story unified and creates our sense of being a coherent, rational agent. (Gazzaniga 2000: 1320)

The effect of telling this story—often a fiction—may be to make the story come partially true: self-interpretation, that is, may serve to make a subject’s behavior, across time, more integrated (or at least integrated-looking). Recall an example from Chapter Three: our split-brain subject, S, who has been reading, suddenly stands, because his RH is bored with reading. His LH could (or could at least try to) force him to sit back down with the book again. This would look like ever-so-slightly non-integrated behavior. But instead of trying to initiate sitting back down again, S’s left hemisphere may instead “go along with” this RH-initiated behavior, perhaps inventing its own reason for having stood. Perhaps LH’s

hemisphere will say to itself (i.e. perhaps S will say to himself, in inner speech), “Time for lunch.” And once he *does* say that, he may very well go make lunch.

Most callosotomy patients experience intermanual *conflict* within the early recovery period. This behavior subsides with time, although may reappear occasionally in brief episodes for years. More commonly, “Even many years after operation, patients will occasionally be surprised when some well-coordinated or obviously well-informed act has just been carried out by the left hand” (Bogen, 1985: 314). “Patients sometimes complain that their left hand behaves in a ‘foreign’ or ‘alien’ manner, and they routinely express surprise at apparently purposeful left-hand actions” (Zaidel, Iacoboni, 2003). “They may quickly rationalize such acts, sometimes in a transparently obvious way” (Bogen, 1985: 314).

I have spoken so far as if only the left hemisphere has a mindreading system or interpreter, and I have furthermore equated possessing a mind-reading system with engaging in self-interpretation (of the sort that leads to confabulation). In fact I suspect that the right hemisphere also supports or engages in mindreading; it has a sense of humor, anyway, and apparently shifts attention automatically to where someone else is looking, while the LH doesn’t (Kingstone, Friesen, and Gazzaniga, 2000). Some RH-initiated cross-cuing in split-brain subjects looks like it might require some beliefs about minds and communication.

The left hemisphere could still be the only hemisphere associated with an “interpreter” in Gazzaniga’s sense, however. Perhaps, even if the RH does have sophisticated mind-reading capabilities, it doesn’t turn them on itself. Even if the right hemisphere *doesn’t* have an interpreter, though, the very fact that the left

hemisphere does could serve to integrate a split-brain subject's behavior, particularly over time.

4.9 Single Streams versus Unified Behavior

I have just come from suggesting that one hemisphere's tendency to self-attribute beliefs and desires and percepts (even falsely), and to rationalize the subject's behavior, can be a source of behavioral integration.

This leads to a sort of broader point about, and perhaps a limited criticism of, the challenge from unified behavior. To some extent, the challenge assumes that two minds will necessarily be in some degree of obvious, visible competition or conflict with each other. Call this the "agent assumption," for it is the consequence of two other assumptions: that minds are agents, and that agents are individuated on the basis of competition and coordination of action. The latter assumption might rest upon some questionable conceptions of agents. But more relevantly, to this work, a mind and an agent are arguably different things. Clearly an agent *has* (or is partly constituted by) at least one mind, but a mind can lack agency. (One can become totally paralyzed for instance without loss of consciousness or cognition.) Agency requires mind, body, and appropriate interactions between the two. (This is not a dualist claim—I'm using "body" to exclude the brain. A functioning brain isn't enough for agency either, in other words.)

Some degree of shared agency is literally forced on however many minds share a body; e.g., one body can only go one place at one time More interestingly, a still further degree of shared agency is not strictly forced upon those minds, but made highly appealing. One hemisphere of a subject may loathe the clothes the other

hemisphere selects, the food it brings to the mouth, its smoking habit, its preferred leisure activities. . . . But what is that first hemisphere to do? Change the subject's outfit—again? Then he'll be late. Yank the cigarette from his own mouth and throw it on the ground, knowing that soon enough he (the other hemisphere) will just be lighting another one anyway?⁶¹ You could burn through a whole pack that way pretty quickly. This is the sort of competition it might not be worth trying to win.

That the challenge from integrated behavior does make this “conflict assumption,” or that the assumption is questionable, might not be obvious. Someone might say, “Look, of course two people can get along splendidly, without ever fighting with each other. But they will still wake and retire to bed at different times, still watch different programs on television, still exhibit talent for and interest in some different things. One will still prefer vanilla and the other chocolate, one conservative and the other liberal principles, even if they have the good sense not to come to blows about such things. Because their minds are different, their behavior, as peaceful as it is, will be visibly different. And if their two minds were in one body, *it would look like it.*”

But, first off, again, having a single body makes it simply impossible to exhibit much behavior that will obviously *look* disunified. Conjoined twins must at least rise from and retire to bed at the same time, though they surely have two minds; the bodily sharing is even more complete in the case of a split-brain subject, who has just one, discrete (rather than two partly merged or shared) bodies. This again forces a high degree of behavioral unity upon such subjects. You can't eat and not eat vanilla

⁶¹ Though, admittedly, one split-brain patient apparently *did* complain that his left hand (RH) wouldn't let him smoke; it kept taking cigarettes from his right hand (LH) and from his mouth and putting them out. Cited in Joseph, 1990, p. 29.

ice cream at literally the same moment (this would be physically impossible), and as a matter of law, you only get one vote. Some sort of compromise is not just essential; it's inevitable. And then it often doesn't look like compromise.

If you took two minds, each from a different body, and placed them in the same body, you would thereby give them incentive to cooperate. Or rather—since that isn't quite the way of putting it; neither mind would necessarily even know the other was there—you would thereby create a situation in which they would naturally cooperate with each other not deliberately, but *by intending to behave rationally*. (Or, by intending for their subject to behave rationally.) Merely by each “acting” in accordance with many of their own interests, especially important background interests, the minds would manage a kind of cooperation-in-effect. And they would do this, they would at least attempt to do this, no matter how limited their capacity for intra-mind interaction with each other.

Finally, and relatedly, we should note that it is entirely open to empirical question to what extent co-consciousness (and for that matter conscious unity more broadly understood) plays a role in integrating behavior in the first place. Again, split-brain subjects behave in a fairly unified fashion even at their most disunified, during the split-brain experiment. These are the same times at which even defenders of the conscious singularity model like Marks (1980) and Tye (2003) believe that split-brain subjects have two streams of consciousness. I have just offered a list of factors and mechanisms that I believe integrate split-brain subjects' behavior without providing a means of inter-hemispheric co-consciousness. If I am right, then having a single stream of consciousness in and of itself might not play too large a role in

explaining integrated behavior. The relationship of that stream of consciousness to a creature's body, the environment the body is located in, the particular contents of the stream, its functional location within a larger mental architecture, the cognitive capacities of the creature—*these*, instead, may play the larger part.

5 Conclusion

The account of co-consciousness that I have adopted makes a prima facie strong case for the conscious duality model in split-brain subjects. I am assuming that the right hemisphere of a split-brain subject has some conscious experiences. This means that either the right hemisphere has its own set of global consumers, or that its experiences are globally broadcast to the global consumers of the left hemisphere. The results of the split-brain experiment suggest that the latter is not the case. Therefore each hemisphere of a split-brain subject has its own set of global consumers to which its own set of experiences are made available. Co-consciousness no doubt holds *within* each hemisphere, but not (at least by and large) across the hemispheres.

Chapter 5: Models of Consciousness in the Split-Brain Subject

1 Introduction

In Chapter Three I argued against a version of the singularity model of split-brain subjects' consciousness, in favor of the conscious duality model, which I defended in Chapter Four. In this chapter, I contrast the model of split-brain subjects' consciousness that I defend with two other, less standard models. The "partial co-consciousness model" says that the two hemispheres share *some* conscious experiences, and rejects the view that co-consciousness is transitive. The "dynamic singularity model" says that a split-brain subject possesses a single *intra*-hemispheric stream of consciousness, one whose "location" shifts from one hemisphere to the other with a shift in motor intentions. I reject these alternative models, and also show that the conscious duality model has the resources to meet the biggest challenge it faces, which is that of accounting for the generally integrated nature of split-brain subjects' behavior.

2 The Partial Co-consciousness Model

Among those models of split-brain subjects' consciousness that ascribe conscious phenomena to the "disconnected" right hemisphere, the two major competitors are the conscious duality model, and a conscious singularity model that says that the two hemispheres jointly realize a single stream of consciousness.

But it might be said that there is a *third*, alternative model of split-brain subjects' consciousness that is superior to both the singularity and the duality model. According to Lockwood's (1989) *partial unity model* (or as I will call it the "partial

co-consciousness model”), split-brain subjects do not have either a single stream of consciousness within which all experiences are mutually co-conscious, nor two discrete streams of consciousness, between which no two experiences are co-conscious. Lockwood instead suggests that we drop the *transitivity thesis*, according to which co-consciousness is a transitive relation. The partial co-consciousness model that results from dropping this transitivity thesis, Lockwood argues, is not only more compatible with the behavioral evidence than is the duality model of split-brain subjects’ consciousness, and more compatible with the neural evidence than is the singularity model; it is also more compatible with the behavioral evidence than is the singularity model, and more compatible with the neural evidence than is the duality model. And there are *no* reasons to reject it, besides a stubborn and unscientific adherence to the pre-theoretic assumption that co-consciousness is necessarily transitive.

Others, however, have argued that there is no good evidence for a partial co-consciousness model of split-brain subjects’ consciousness (Hurley 1994, 1998, 2003), or have even suggested that there is something impossible or even incoherent about the consciousness described by the model (Bayne and Chalmers, 2003). In this section I describe and offer a limited defense of the partial co-consciousness model; a stream of consciousness characterized by non-transitive co-consciousness relations is in fact *prima facie* plausible. I suggest that the model nonetheless does not offer the best characterization of split-brain subjects’ consciousness. Indeed my arguments imply that for any creature C such that we would have some reason to apply

Lockwood's model of split-brain subjects' consciousness to it, there is in fact greater reason to identify C with two streams of consciousness.

2.1 A Case of Partial Co-Consciousness?⁶²

As Lockwood (1989) notes, we intuitively think of co-consciousness as a transitive relation. We believe, that is, if A and B are co-conscious, and B and C are co-conscious, then A and C are co-conscious, as well. If co-consciousness is transitive in this way, then each experience that occurs within a stream of consciousness is co-conscious with every other experience in that stream, and is not co-conscious with any experience outside that stream.

Now what happens if you acquire evidence that, in apparent contradiction of this transitivity principle, experience A and C are *not* co-conscious with each other? The principle can be saved by concluding that you earlier erred in believing that there was just a single mental token with B's content, and by accepting that there are instead two mental tokens, two experiences, B and B', with the same content, and that B is co-conscious with A and B' is co-conscious with C. Thus one resolves *prima facie* violations of the transitivity principle by revising one's initial individuation of experiences. Lockwood, however, suggests the opposite move: keep the individuation of entities as is—just drop the *principle*.

For some split-brain experiments *do* appear, *prima facie*, to provide examples of failures of transitivity. Consider for instance an experiment by Sperry, Zaidel and

⁶² When I speak of transitivity and failures of transitivity in this chapter, I mean to refer to transitive co-consciousness and failures of co-consciousness to be transitive. Also, I refer to the "partial co-consciousness model" and to "partial co-consciousness," though strictly speaking I should refer to the "[only] partially transitively co-consciousness model" and to "partially transitive co-consciousness."

Zaidel (1979) in which a four-photograph array was projected in the LVF (i.e. to the right hemisphere) of split-brain subject N.G., who was then asked to point to each photograph in turn, using her left hand (under dominant control of her right hemisphere). In one instance the array consisted of three photographs of unknown men, and one photograph of her adult son. N.G. began pointing to each photograph, and then paused when she got to the photograph of her son. The following conversation between N.G. and the experimenter ensued:

- N.G.: Hey, wait a minute. That's L____ [her daughter]. No that's me. No, wait a minute.
Ex: Do you recognize one of these?
N.G.: Yes, that one right there. [She points to the photograph of her son.]
Ex: OK, how do you feel about this person?
N.G.: Good, good, good. Me, when I was younger. . . or L____ [her daughter] or, or B____ [her husband]. . . or I don't know. [She laughs loudly.] That's it. That's gotta be. [She laughs again.]
Ex: Whatever it is, it is OK, eh?
N.G.: Yea, it's fine, it's beautiful.
Ex: You don't see any others there that you recognize?
N.G.: No. Just that one. [She points again to the picture of her son.] The best looking one there.

(Sperry, Zaidel, Zaidel, 1979: 158)

Let me describe what makes this scenario puzzling—first, from an intuitive perspective, and then in a way that will motivate, and illustrate, the partial co-consciousness model.

It would appear that N.G. visually recognizes her son in the photograph: the photograph of her son evokes a positive emotional response, and reminds her of her family. Moreover these appear to be *conscious* responses: N.G. *reports* her positive emotional reaction, and begins *naming* family members. But there is evidence that at the same time N.G. does not even *see* her son; she can't even tell the *gender* of the

person photographed. And this not-seeing appears conscious also, in the sense that N.G. does not describe being blind, or having a blind spot. So she appears to have a “unified consciousness” in *one* respect: the positive emotional response evoked by the photograph appears “unified with” (co-conscious with) whatever thoughts and experiences are generating her stream of speech. But she appears to lack a “unified consciousness” in another respect: again, she must have *seen* and recognized her son, but that visual experience and recognition does not seem co-conscious with whatever thoughts and experiences are generating her stream of speech.

Now let me describe this scene more carefully, and in a way that depicts it as a case of partial co-consciousness. Clearly both of N.G.’s hemispheres had access to *some* information about the stimulus, and clearly the two hemispheres did not have access to *all of the same* information about the stimulus. N.G.’s right hemisphere, presumably, recognized the photograph as being of her son, a recognition that generated in her a positive affective state. N.G.’s left hemisphere also experienced a positive affective state—she expresses, via her verbal left hemisphere, positive feelings towards the person pictured—but lacked access to purely *perceptual* information about the stimulus; again, she apparently could not tell if the pictured person was male or female, for instance. And while her left hemisphere had some category information about the depicted subject—she knew she was seeing a member of her immediate family—it lacked stimulus *identity* information: she (LH) didn’t know she was seeing her *son*, specifically.

Moreover, it seems as if both hemispheres may have been subject to *some* of the same experiences as N.G. looked (RH) at the photograph. Both hemispheres

seemed to experience the same happy, fond emotional response. And that emotion was no doubt co-conscious with the visual experience that produced it, N.G.'s RH visual experience of the photograph. But N.G. was presumably also having a non-overlapping set of visual experiences in her left hemisphere at the same time (she didn't cry out (LH) that everything had gone black), and this set of visual experiences was presumably co-conscious with the affective state evoked by the photograph of N.G.'s son as well, since with her left hemisphere N.G. verbally expressed her positive feelings towards the photographed subject. And yet N.G.'s right and left hemisphere visual experiences were not co-conscious with each other; N.G. could not with her left hemisphere identify the subject of the photograph. In other words, N.G. had a *positive emotion* evoked by the photograph, and this positive emotion was co-conscious with her right hemisphere visual experiences, and with her left hemisphere visual experiences, and yet her right and her left hemisphere experiences were not co-conscious with each other.

So appears to be some mental state—an affective state say—co-conscious with the conscious visual experiences of both N.G.'s right and her left hemisphere. And yet there is still at the same time clearly a *failure* of interaction between some right and left hemisphere experiences. So co-consciousness both appears to hold and *not* to hold across the hemispheres. In the face of this (from his perspective) contradiction, Nagel (1979) suggested that the entire concept of conscious unity (co-consciousness) might be unscientific—might even be an illusion. But Lockwood argues that it is the apparent *contradiction* that is an illusion—or, at least, that it goes

away if we drop the assumption that conscious “unity” is always perfect, i.e. that co-consciousness is necessarily transitive.

2.2 Transitivity and the Partial Co-consciousness Model

Nagel (1979) worried that if streams or centers of consciousness aren’t discrete, then perhaps the whole notion of conscious *unity* had best be dropped altogether. For Nagel, that is, transitivity was an essential property of the co-consciousness relation, and therefore essential to the very concept of a unified consciousness or a stream of consciousness: such concepts are essentially of something *wholly unified*, something necessarily *not* partial. Is this correct? Or might streams of consciousness still exist—even if they are not (or not always) structured in quite as neat a way Nagel believes?

Largely dropping the language of *streams* of consciousness, Lockwood creates the term *phenomenal perspective* to refer to a set of experiences that are all perfectly co-conscious with each other. He then drops the transitivity principle. When we do this, he says, “a remarkable possibility immediately opens up; that of simultaneous, *overlapping* phenomenal perspectives. And it is in these terms, I believe, that we should seek to understand the split-brain cases” (1989: 89; original emphasis).⁶³ Within the partial co-consciousness framework, we can, if we wish, still speak of a split-brain subject’s consciousness using the language of streams of

⁶³ The language of *overlapping phenomenal perspectives* leaves somewhat ambiguous whether the overlap involves mental types or tokens. The most natural connotation of an *overlap in perspective* is that of an overlap in *what is seen*. Thus, my sister’s perspective on the 2008 presidential primaries overlap with mine (we both agree that Thomson would make a bad president, but we disagree over whether or not Huckabee would be worse). But Lockwood does not mean that two phenomenal perspectives overlap when they both include one or more of the same experience *types*, as this would just be a version of the duality model. (At least, according to him, as well as to me, of course.) Rather, two phenomenal perspectives overlap when they each encompass or include one or more of the same *token* experiences. To avoid this potentially confusing terminology, then, I will not speak of overlapping phenomenal *perspectives*; instead I will continue to speak of co-consciousness and transitivity and failures of transitivity.

consciousness. We could just say that a split-brain subject has neither two perfectly discrete streams of consciousness nor one stream of consciousness within which co-consciousness is perfectly transitive. Whereas one who accepts the transitivity principle will say that what makes an experience part of a stream of co-consciousness is that it is co-conscious with *every other* experience in that stream, Lockwood can just say that what makes an experience part of a stream of consciousness is that it is co-conscious with *some* other experiences in that stream.

Lockwood says that if we instead insist upon the transitivity of co-consciousness, then we will have to accept that split-brain subjects have two streams of consciousness, since at any given moment a split-brain subject's inter-hemispheric experiences won't be *perfectly* transitively co-conscious. Unlike Nagel (1979), Marks (1981), and Tye (2003), that is, Lockwood doesn't seem to find the conscious duality claim (or even the two minds claim) particularly difficult to reconcile with the behavior of split-brain subjects. For he says:

ostensibly, it is the two-minds theory that has the most going for it. Indeed, Nagel's arguments for rejecting this alternative are hardly decisive. Nagel places great stress on the apparent normality of split-brain patients when they are not in the experimental situation. But it is possible to explain this normality, consistently with there being two streams of consciousness. . . . the two hemispheres of a split-brain patient will, under normal circumstances, be receiving broadly the same information, and certainly consistent information.
(1989: 85)

Moreover, “Quite apart from that, there is, contrary to what Nagel suggests, a certain amount of evidence for dissociated behavior in split-brain patients, even outside the experimental situations” (1989: 86).

Unlike Marks’ or Tye’s singularity models of split-brain subjects’ consciousness, then, Lockwood’s partial co-consciousness model is not driven by the perceived implausibility of the conscious duality model of “split brain” consciousness. Instead he is drawn to the partial co-consciousness model for two, more principled reasons. One reason concerns the relationship between neural and psychological tokens, as we will see in a moment. The second concerns the relation between a more empirically/theoretically based individuation of psychological entities and what he sees as pre-theoretical, pre-empirical commitments about the nature and relations of those entities. There is, he says:

something deeply unsatisfactory about a philosophical position that obliges one to impose this rigid dichotomy upon the experimental and clinical facts, either we have just one centre, or stream, of consciousness, or else we have two or more, entirely distinct from each other. (1989: 86)

Everyone will agree, he seems to suggest, with Nagel’s original observation that in split-brain subjects the two hemispheres seem co-conscious with each other in some ways but not in others. Why insist that nonetheless they either *are entirely* or *are not at all* co-conscious with each other?

There is an intuition here, a general intuition about philosophy, folk science and psychology, and the natural world, that is appealing: we may tend to *think* in whole numbers and discrete units, we may want and tend to see the world in terms of

dichotomies, sharp boundaries, clear categories and binary distinctions; but the concepts that draw such bold lines may not neatly map onto a complex and messy world, much of the time. Of course, this may always be the case. (At least in situations involving concepts of ostensibly sharply delineable things—concepts like ‘planet’ and ‘mouse’; there are many other concepts—‘justice’, ‘cloud’—we would expect to have looser application criteria.) Surely this does not show that *all* those concepts require modification or outright elimination. Sharp conceptual distinctions can sometimes help us get a grip even on messy or complex phenomena that those concepts don’t always capture neatly. But in some cases, the application of such sharp distinctions may not be even epistemically helpful; it may instead only make the shape or structure of a phenomenon of interest *more* difficult to grasp than it should be. In such cases, rather than simply shrugging that while the ways we have of thinking about the world are not always in lock-step with that world, they’re still the only ways of thinking about the world that we’ve got, why not try to reject and revise and create new concepts, that make *better* sense of our world? This is the first motivation for Lockwood’s partial co-consciousness model.

A second and related motivation for Lockwood’s partial co-consciousness model of split-brain consciousness concerns what I have called the isomorphism thesis, following Hurley (1998)—the thesis that there is necessarily some sort of isomorphism between the structure of the mind (and of consciousness) and neural structure. (Hurley, recall, rejects this thesis.) If the co-consciousness relation supervenes on some physical pathway or pathways, then co-consciousness must come in degrees, since there are obviously degrees of physical connectedness and

disconnectedness. (Descartes reasoned similarly: if the mental (the conscious) is physical, then consciousness isn't essentially unified, since matter isn't essentially unified. Since Descartes believed that the mental was necessarily unified, he concluded that the mental wasn't physical. Lockwood believes that the mental *is* physical, and therefore rejects the claim that the mental is necessarily unified, i.e. in this instance that co-consciousness is necessarily perfectly transitive.) Lockwood asks us to imagine cutting the corpus callosum one tiny section at a time (1989, 1994). The first cut leads to a very small bit of conscious dissociation; the last cut results in the greatest degree of conscious dissociation associated with the split-brain phenomena. Would conscious duality advocates, Lockwood asks, maintain that the very first cut creates two streams of consciousness? Or only the last cut? Or some small cut in between? All of these look like unattractively arbitrary positions to try to defend.

One might be tempted to object that Lockwood commits a version of the "heap" fallacy here. There of course isn't a single neural fiber that, once cut, suddenly makes a single stream of consciousness dual. But differences in degree can still be significant, and the difference between conscious singularity and conscious duality may sometimes be a matter of degree, even if between the two far ends of the spectrum there *is* a middle space that we recognize as containing borderline and indeterminate cases. (Compare with two *actual* streams, whose banks merge for several feet before a boulder divides them, ultimately sending them down different faces of a mountain.) Again, still setting aside, for the moment, how and even whether this is possible, it would seem that if there *were* occasional or some small number of experiences co-conscious with two otherwise distinct sets of experiences,

there would still be two streams of consciousness, co-conscious with each other to some very small degree. For Lockwood to disagree, I think, would be to commit the same error of abstemiousness that he believes those who insist upon the transitivity of co-consciousness commit. Of course, this objection is compatible with the fundamental ideas behind the partial co-consciousness model, and I pursue different objections to the model here.

The partial co-consciousness model ostensibly enjoys the strengths of both the singularity and the duality model, while not suffering from their weaknesses. Where the advocate of conscious singularity referred to conscious singularity to explain behavioral integration, Lockwood can refer to co-consciousness across the hemispheres. Where the advocate of conscious duality referred to conscious duality to explain dissociated behavior, Lockwood can refer to the lack of co-consciousness across the hemispheres. And the partial co-consciousness model is again ostensibly a better fit with the neuroanatomical data (or, at least, with *some possible* neuroanatomies, even if one thinks that the duality model for instance is a better fit with the neuroanatomy of a split-brain subject in particular). The two hemispheres of a split-brain subject are physically connected to each other and causally interact *via* subcortical and non-cortical structures. To the extent that the singularity model is supported by neuroanatomical facts, then, the partial co-consciousness model appears better supported; to the extent that the duality model is supported by neuroanatomical facts, the partial co-consciousness model may again appear better supported.

In fact before examining the crucial question of how and even whether failures of transitivity between co-conscious experiences is possible, I wish to say

three further things in defense of the partial co-consciousness model. First, the isomorphism thesis that I take Lockwood to be defending—a thesis I described in Chapter Three—makes perfectly good sense. Of course, this is in part because that thesis is basically (though not, I don't think, viciously) circular; as Revonsuo says about phenomenal consciousness for instance, “there *must* be isomorphism between one specific level of organization in the brain and phenomenal consciousness, simply because *these boil down to one and the same thing*” (Revonsuo, 2000: 67; original emphasis). Lockwood would not believe that the neuroanatomy of a split-brain subject provided support for his partial co-consciousness model over a conscious duality model, for instance, if he did not already believe that some non-cortical (including subcortical) structures could bear conscious experiences.

Second, the kind of *attitude* Lockwood brings to bear on the split-brain cases is a good one. Certainly there is a good deal of evidence, an increasing amount, that introspection can lead us astray in theorizing about consciousness. This is not to say that introspection is without value or always incorrect, for it obviously isn't. But it is to say that we need more reason for insisting upon the perfect transitivity of co-consciousness other than just the *feeling* that our experience is perfectly transitively co-conscious. This is especially the case because introspection seems particularly incapable of telling us anything about the *structure* of consciousness. Subjects of experience are what have phenomenal perspectives, and there is one subject of experience for each stream of (transitively co-conscious) experiences; even Lockwood would seem to acknowledge this; recall his definition of a “phenomenal perspective” (but see the conclusion of his 1994). Everyone (every subject of

experience!) will *feel* as if they have one stream of consciousness (at a time), in other words—regardless of how many streams of consciousness are actually associated with their brains.⁶⁴ Lockwood’s openness to amending our understanding of conscious unity (co-consciousness), to allow for failures of transitivity for example, is wise. For our pre-theoretic ideas about co-consciousness no doubt are *wrong* in many respects and will require (at a minimum) substantial revision as neuropsychology develops.

Finally, it is, at first glance, easy to imagine how a consciousness could come to be “partially unified” on the account of co-consciousness that I have offered. A single creature could have two distinct sets of (token) consumer systems, and two sets of globally broadcast states. For the most part, each set of global consumers might work with a unique set of experiences. But some experiences, some mental tokens, might be broadcast to *both* sets of global consumers simultaneously. And those experiences would presumably be co-conscious with all the other experiences being broadcast at that time to either or to both sets of global consumers. Again, this all seems right, at a glance—though I will argue in short order that this first glance may be misleading.

⁶⁴ It follows that, where phenomenology is concerned, the partial co-consciousness and the conscious duality model for split-brain subjects tell the same story. There will be two subjects of experience, it would seem, either way. Lockwood apparently denies this, however, saying, “Like Nagel, I am still unable to project myself into the position of a subject with a partially unified and partially disunified consciousness” (1994: 95). My point is that a *subject of experience* is the subject of all experiences that are *transitively co-conscious* with each other—even if that subject of experience belongs to a subject (i.e. a creature) whose conscious experiences are *not* all transitively co-conscious with each other. So, what it is like to be a subject with a “partially unified consciousness” (of the sort under discussion here) is what it is like to be a subject with two streams of consciousness.

2.3 Objections to the Partial Co-consciousness Model

Lockwood would surely (and, I think, rightly) reject a Nagelian objection to the partial co-consciousness model that stated that co-consciousness just *has* to be transitive, simply because this is part of our concept of consciousness, or part of our concept of conscious unity, or something. If the partial co-consciousness model does offer the best characterization of split-brain subjects' consciousness, on phenomenological, functional, and/or neuroanatomical grounds, then our concept of co-consciousness can be revised to not assume transitivity. Given the state of our understanding not just of the structure of consciousness in split-brain subjects, but also of the nature and mechanisms of co-consciousness, and for that matter consciousness, in general, it is impossible to say with any confidence much less certainty which model of conscious structure best characterizes the split-brain phenomenon. Our empirical knowledge concerning interhemispheric transfer in split-brain subjects is too contradictory and incomplete, our theoretical understanding of interhemispheric transfer similarly so, and our understanding of consciousness in general is minimal. I nonetheless *suspect* that neither phenomenological, functional, behavioral, nor neuroanatomical considerations provide strong support for the partial co-consciousness model, and that the conscious duality is more strongly supported.

Let us look at the phenomenological considerations first. Here I argue that while phenomenology can offer some kind of support for an ascription of conscious duality, it cannot do so for an ascription of partial co-consciousness. In making this argument I will refer first to a number of creatures, C1, C2, C3.... These creatures possess experiences *like* E1, E2, and E2', where the numbers refer to types or contents

of experiences, and primes distinguish tokens of content-identical experiences, though only *within* a creature. (I.e. if I say that C1 possesses E2 and C2 possesses E2, this means that C1 and C2 possess distinct token experiences with the same content; if I say that C2 possesses E2' and E2'', this means that C2 possesses two token experiences with the same content.) See Figure 5.1.

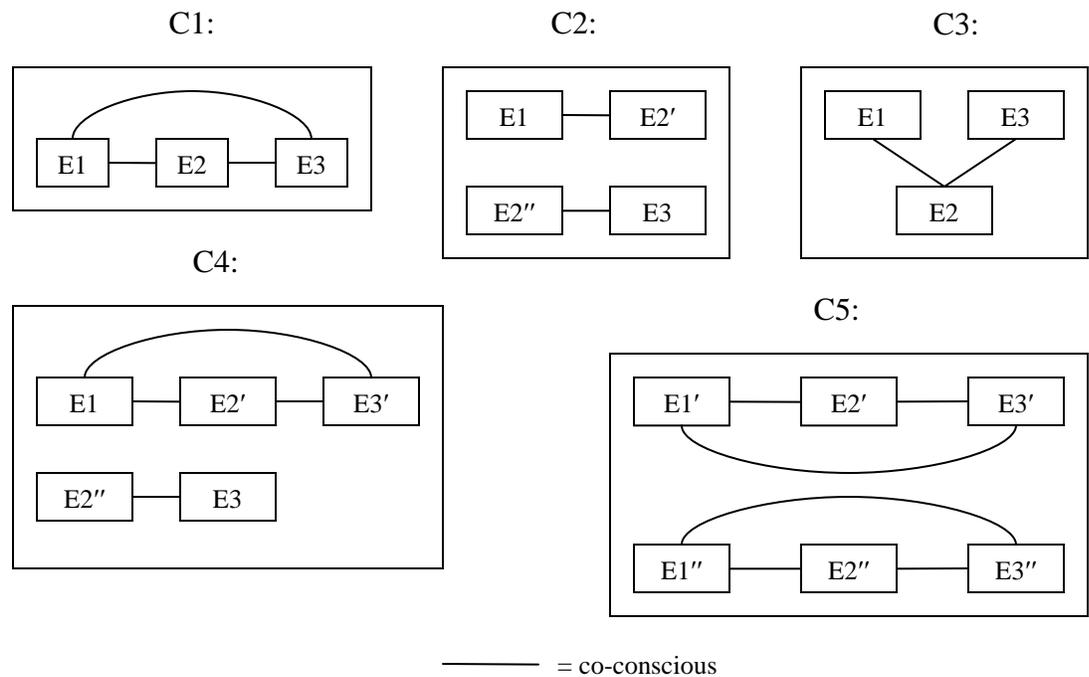


Figure 5.1: Some Kinds of Conscious Structure

C1 has a single stream of consciousness containing or composed of three experiences that are all mutually co-conscious with each other, E1, E2, and E3. Of course there is nothing peculiar it is like to be C1 (at least as so minimally described). In the second creature, C2, consciousness is dual: C2 possesses experiences E1, E2', E2'', and E3. E1 and E2' are co-conscious, and E2'' and E3 are co-conscious, and neither E1 nor E2' are co-conscious with E2'' or with E3. As we have already noted, there is nothing very odd it is *like* for C2 to possess this dual consciousness. With

regards to just the *structure* of their consciousness, in fact, there is no *subjective* difference between being C1 and being C2. There are just two different things it is like to be C2; in fact, it is better to identify C2 with two subjects of experience, S1 and S2, one of whom experiences E1 and E2' together, and one of whom experiences E2'' and E3 together.

(Put differently, if creature C4 is like creature C2 but with a fifth experience, E3', where E3' is co-conscious with E1 and E2' but with none of C4's other experiences, C4's consciousness is transitively co-conscious, and dual, and what it is like for C4 to possess E1, E2' and E3' is exactly what it is like to *be* C1. Of course, what it is like to be C1 still does not capture *everything* it is like to be C4, since there is also something different it is like for C4 to possess E2'' and E3. This is why, as I just said about C2, it would be better to identify C4 with two subjects of experience.)

Now imagine a third creature, C3, to whom one would be tempted to attribute an only partially transitive co-consciousness. C3 possesses the *non-co-conscious* experiences E1 and E3; imagine that there is also some reason or other to believe that C3 has a third, final, and single experience E2, that is co-conscious with both E1 and E3. Now, what would it be like to *be* C3? As Hurley (2003, 1998) notes, being C3 would be no different, from C3's subjective perspectives (plural), from being C2. There is one thing it is like for C2 to possess E1 and E2 simultaneously, and another thing it is like for C2 to possess E2 and E3 simultaneously, and nothing it is like to possess E1 and E3 simultaneously (either with or without E2).

In other words, there is no phenomenological support for attributing to C2 an only partially transitive co-consciousness. It might be said that this says little or

nothing *against* ascribing a partially transitive co-consciousness to C3. After all, I have already noted on multiple occasions, including just above, that having two streams of consciousness feels no different from having one stream of consciousness, for the simple reason that there is *no such thing* as a feeling of having two (synchronic) streams of consciousness, properly speaking. All (conscious) feelings are located *within* one stream of consciousness or another. (And if a conscious feeling occurred outside of a stream of consciousness, it could hardly be a feeling of having that stream of consciousness.) Another way to put this would be to say that things *feel* one way or another only to *subjects of experience*. And transitive co-consciousness really does seem essential to a subject of experience. This is why even Lockwood adopts the language of phenomenal *perspectives*, acknowledging that a split-brain subject would have *two* of these. So a subject of experience will always feel as if he or she or it has a single stream of consciousness (or else a single experience).

Yet I have also argued that there are nonetheless reasons, *non-*phenomenological in nature, not to attribute a single stream of consciousness to a split-brain subject. Phenomenology simply can't tell us much about the *structure* of consciousness.⁶⁵ Thus it would seem that by my own arguments the lack of phenomenological support for attributing only partial co-consciousness (i.e. only a partially transitive co-consciousness) to C3 does not provide any good argument *against* such an attribution, either.

⁶⁵ I put the point this way: phenomenology most directly yields information about the *contents* of consciousness, and not about the structure of consciousness. (Though as I will explain in a moment some facts about conscious structure can be *inferred* from phenomenology.) Hurley (1994, 1998) puts this same point in different terms: phenomenology can reveal to you the *types* of experiences you're having, but not their token identities. We are agreed that objective and not just subjective facts or properties are needed to individuate token experiences and thus streams of consciousness.

While the lack of phenomenological support for attributing to C3 partial co-consciousness is by no means decisive, the comparison between the position of the duality claim relative to the singularity claim, where phenomenology is concerned, and the position of the partial co-consciousness claim relative to the duality claim, where phenomenology is concerned, may mislead. The duality claim *can* be supported by phenomenological evidence over the singularity claim—indirectly. For what phenomenology does reveal is the *content* of a subject's experience. Therefore phenomenology *does* support a conscious duality claim over a singularity claim, for creature C2 above, for example. It does this *not* because there is a phenomenology associated with C2's having two streams of consciousness instead of one; there is no phenomenology associated with having multiple streams of consciousness per se. It does this simply because there is nothing it is like for C2 to have experiences E1 and E3 simultaneously. And because there is nothing it is like for C2 to have experiences E1 and E3 simultaneously, and yet C2 *does* have experiences E1 and E3 simultaneously, we know that C2's experiences, unlike C1's experiences, *cannot* be transitively co-conscious. Unlike C1's phenomenology, C2's phenomenology reveals the *absence* of co-consciousness of E1 and E3. Their phenomenology is different. In this sense, then, phenomenology supports a conscious duality claim over a conscious singularity claim for C2.

In contrast phenomenology *cannot* support a partial co-consciousness claim over a conscious duality claim for C3. There is by necessity a phenomenological difference between having three mutually co-conscious experiences and three only partially transitively co-conscious experiences—between being C1 and C3, in other

words. But there is no necessary phenomenological difference between having a consciousness structured like C3's and having one structured like C2's. Why not just identify C3's E2 with *two* content-identical experiences, then—why not attribute to C3 in other words, E2' and E2''? Then we would attribute to C3 *four* experiences, just as we did with C2, where E1 and E2' are co-conscious and E2'' and E3 are co-conscious (but neither E1 nor E2' is co-conscious with either E2 or E3'').

It is important to be clear on what this first objection to the partial co-consciousness model is saying. It is *not* saying that there is any *positive* phenomenological evidence *against* the partial co-consciousness claim in favor of the conscious duality claim. (That is, if we *assume* that the two hemispheres of a split-brain subject are sometimes associated with qualitatively identical mental e.g. emotional contents. Obviously if they aren't, this provides positive phenomenological evidence against the partial co-consciousness model in favor of the conscious duality model. But let us assume, for the sake of argument, that the two hemispheres of a split-brain subject are sometimes subject to the same conscious contents.) Phenomenology cannot provide this sort of evidence. It is just saying that there isn't any positive phenomenological evidence in favor of the partial co-consciousness claim over the conscious duality claim.

Note, for that matter, that phenomenology doesn't even provide any positive evidence in support a conscious singularity claim over a conscious duality claim for C1. For there is no difference that phenomenology can reveal between C1 and C5, the latter of whom possesses two sets of mutually co-conscious experiences [E1', E2', E3'] and [E1'', E2'', E3''], with no co-consciousness between the two sets. This is an

interesting fact: phenomenology cannot support the partial co-consciousness model because phenomenology can never provide evidence for one mental token rather than for two mental tokens with one content. Phenomenology reveals contents, and thus can provide evidence for multiple conscious tokens by providing evidence of multiple conscious contents, but cannot provide evidence of single conscious tokens, only single conscious contents. Thus something *besides* phenomenology is needed to motivate attributing to C3 the single mental token E2, as opposed to two mental tokens, E2' and E2''.

Some might object, however, that phenomenology can provide at least *prima facie* evidence of conscious singularity, for instance, in a creature whose phenomenology reveals no *contents* that aren't experienced together. I am not sure, though, that a phenomenology that doesn't reveal two sets of not co-conscious contents really does provide evidence of conscious singularity, given the general limits in what phenomenology can reveal about the structure of consciousness. But let us assume that this is right, for the sake of argument: let us assume that the absence of positive evidence for conscious duality (or for the lack of co-consciousness) constitutes *prima facie* evidence in favor of conscious unity (or for the presence of co-consciousness). Phenomenology still can't provide strong support for the partial co-consciousness model—especially since a consideration of another kind of phenomenological entity, the subject of experience, seems to count in favor of the conscious duality model instead. So something *besides* phenomenology is needed to motivate attributing to C3 the single mental token E2, as opposed to two mental tokens, E2' and E2''.

Behavioral integration on C3's part cannot motivate attributing to C3 only a single experience E2, rather than two experiences with the same content as E2. For *both* the duality model and the partial co-consciousness model would explain C3's integrated behavior with respect to whatever E2 represents in *terms* of what E2 represents (in terms of its content, in other words). Now, a *singularity* model, applied to C3, might try to explain C3's integrated behavior with respect to whatever E2 represents not just in terms of E2's content but also, for instance, the content's availability to a *single center of agency and reasoning*. But if the account of co-consciousness that I have adopted is correct, then the partial co-consciousness model does not have this line of explanation available to it. For any failure of transitivity in co-consciousness can only be explained by the presence of at least two f-sets of global consumers.

Lockwood of course looks to "split brain" neuroanatomy to support his model of split-brain subjects' consciousness, which is sound, I think—but the neuroanatomy that matters is functional neuroanatomy. I think Lockwood (1994) is perfectly right, in other words, to respond to Hurley, "it seems to me that the physical basis of the unity of consciousness should be sought in *whatever we have reason to identify as the physical substratum of consciousness itself*" (94; emphasis added). But if the account of co-consciousness that I have offered is correct, then "split brain" neuroanatomy, and in particular the intact subcortical structures that Lockwood believes support interhemispheric co-consciousness in split-brain subjects, explains why some of the same (or similar) *contents* become conscious in the two hemispheres.⁶⁶ But it does not

⁶⁶ Lockwood by the way does not specify which subcortical structures in particular he suspects of playing this role.

suffice to make a single set of conscious representations available to the two hemispheres. Let me explain why, using a familiar example of what could reasonably be seen as a plausible case of partial co-consciousness.

Gazzaniga has just presented to the right hemisphere of the split-brain subject V.P. a “filmstrip from the Cornell Office of Health and Safety, which counsels its employees not to throw fellow employees into fires should they come across one” (Gazzaniga 1985: 76). (He assures us that, “This bit of drama brought. . . by some federal regulation is indeed terrifying.”) The following conversation ensues between V.P. and Gazzaniga:

M.G.: What did you see?

V.P.: I don’t really know what I saw. I think just a white flash.

M.G.: Were there people in it?

V.P.: I don’t think so. Maybe just some trees, red trees like in the fall.

M.G.: Did it make you feel any emotion?

V.P.: I don’t really know why but I’m kind of scared. I feel jumpy. I think maybe I don’t like this room, or maybe it’s you. You’re getting me nervous.

(Gazzaniga 1985: 76-77)

V.P. then apparently “turned to an assistant and said, ‘I know I like Dr. Gazzaniga, but right now I’m scared of him for some reason.’” (ibid)

This scene is a prima facie plausible instance of *partial co-consciousness*: V.P. has some emotional state of fear or anxiety that is co-conscious with both her left and her right hemisphere experiences, for instance, though some of these right and left hemisphere experiences are not co-conscious with each other.

So V.P. is watching an alarming film—a film showing people throwing their coworkers into fires—with her right hemisphere. Her right hemisphere sees people being thrown into fires, on the basis of which a somatic state is entered into. A representation of this somatic state—an affective state—is also formed, in some subcortical structure or structures, say. Perhaps after some cortical elaboration within the right hemisphere, this affective state (or a causal descendant of it) is broadcast, by the thalamus, say, to both the subject's right and to her left hemisphere global consumers. The subject's right hemisphere consumer systems, meanwhile, are also still being broadcast the contents of the film itself. Of course it is difficult to say precisely what right hemisphere emotional experience/s V.P. is having. Is she *distraught* thinking about the likely injuries and scars, perhaps even deaths, of the people thrown into the fire? Is she *angry* at the throwers? Is she simply *upset* by the contents of the film, even though she knows it's a dramatization and is neither concerned for the thrown nor angry at the throwers? But let us just imagine that in or through her right hemisphere V.P. is experiencing the emotional state of distressed concern for the film's victims.

Now V.P.'s left hemisphere consumer systems are not receiving a representation of the film. They are receiving a stream of auditory representations, however; V.P. has been asking herself what sorts of questions Gazzaniga will ask her next. Let us say that in or through her left hemisphere V.P. begins feeling *nervous* about these questions, and soon enough *afraid* of the person who will be asking them of her. Thus the same affective state of negatively-valenced arousal is an aspect of two similar but distinct emotional experiences, and co-conscious with two sets of

experiences—auditory experience of inner speech, and visual experiences of the film—that are not co-conscious with each other.

(In my view the example is not significantly changed if V.P.’s right and left hemisphere are both experiencing the same emotion—e.g., anxiety—because, as I explain in a moment, what is really shared by the hemispheres in either case is a common cause of an emotional state and not the emotional state itself. This is true whether or not the emotion in the two cases is of the same type.)

I have just said that there is a single affective state here, but of course how to individuate affective states is also not given. In fact I think there is some evidence to suggest that the major subcortical components of emotional experience are themselves lateralized, at least to a large extent. (See Doty 1989, for example.) For the sake of argument, though, let us simply stipulate that there is a *single* neural event in the V.P.’s amygdala, say, receiving some kind of information about the film from high level visual and association cortices, and that this event initiates interaction with V.P.’s hypothalamus, which in turn produces an increase in V.P.’s heart rate and other sympathetic nervous system responses. Call this amygdala event the *response-signaling event*, because it is this event that is responsible for generating an autonomic response to the film. Let’s us also simply stipulate that there is some single amygdala event representing the occurring autonomic response (the arousal response), and call the latter a *response-representing event*. Finally, assume that it is in part via access either to the response-signaling event, or to the response-representing event, or to both, that V.P. comes to feel afraid in her left hemisphere.⁶⁷

⁶⁷ Needless to say, while there are things like this “response-signaling” and this “response-representing” event, their neural correlates are not clear. The amygdala has been suggested as the site

Now: is this response-signaling event a token of the type experience—a representation that is a component of a stream of consciousness? And if so, is it a single experience? Or is it (associated with) two experiences? Is this response-representing event a token of the type experience? And if so, is it a single experience, or two experiences? Do the response-signaling and response-representing events together constitute a single token of the type experience?

I don't adopt a particular theory of emotion or emotional experience in this work, although my overall approach is cognitivist insofar as I wouldn't regard either the response-signaling event or the response-representing event described above as constituting a *whole* experience. Nonetheless, even if the response-signaling event described above *is* of the type "experience" or "affective state" (where an affective state is a representational component of an experience), and even if the response-representing event described above *is* of the type "experience" or "affective state," neither the response-signaling event nor the response-representing event should be identified with a single conscious representation.

I have adopted in this work an account of phenomenal (and access) co-consciousness in terms of access consciousness, and I have accepted that some version of the global workspace theory offers an account of access consciousness. According to this theory, a representation is conscious if it is globally broadcast to

of the response-representing event, but so too have the mid-insula (especially in the right hemisphere) and the orbitofrontal cortex independently of the insula. (E.g. Critchley, 2005.) The amygdala is more securely recognized as at least *one* site of response-signaling events, at least for the emotion of fear. (See LeDoux 1990, 1993a, 1993b; LeDoux et al. 1990.) But cortical areas (in particular the orbitofrontal cortex) have also been implicated as the or a site of response-signaling events. (See Bennett, Hacker, 2005; Buttern, Snyder, 1972.) Obviously to the extent that V.P.'s response-signaling and response-representing events are cortically located there is less reason to think that V.P. has only a single emotional or affective experience; I just assume here for the sake of argument, though, that the response-signaling and response-representing events *are* subcortically located.

global consumers. It is the broadcasting that *constitutes* the broadcasted state, that constitutes the experience. That is to say, *the broadcasting event itself is the conscious representation.*

And there are two broadcasting events responsible for V.P.'s emotional responses in the scenario described above. We can see this in two ways. First, just in general, it is apparently possible to broadcast a representation to one hemisphere's set of global consumers and not to the other's; the results of the split-brain experiment show this. Second, and somewhat more subtly, what *makes* an event a broadcasting event is the presence of the global consumers to which the representation is broadcast. But if one hemisphere lacked global consumers, the content carried by the response-representing event in this example would still be broadcast, since it could be broadcast to the other hemisphere's consumer systems. So there are two broadcasting events. It is the broadcasting events, however, and not the response-representing or the response-signaling events in and of themselves, that suffice for experience. And just as there are two broadcasting events, there are two broadcast states or experiences.

But what *are* these two conscious, broadcasted experiences? One might think that they are the two emotions, distressed concern and fright, that V.P. experiences via her right and her left hemisphere, respectively. It is relatively obvious that because there are two broadcasting events, there are two emotions, two emotional experiences. By virtue of the contents of the response-signaling and the response-representing events being broadcast to left hemisphere global consumers, along with other representations being simultaneously broadcast to those consumers, the

experience of *fear* is generated. By virtue of the contents of the response-signaling and the response-representing events being broadcast to the right hemisphere, along with the broadcast of other representations (e.g. visual representations of the film) to those consumers, the experience of *distressed concern* is generated. But these emotional states or experiences aren't the *equivalents* of the broadcasting events. Rather, they are the *outcome* of those broadcasting events, the product of representations being broadcast and then processed by global consumers. *The broadcasted representations are the affective components* of these emotions. And, again, there are two of them. To repeat, there are not just two different emotional experiences (whether of the same or of different types) *generated by* the two broadcasting events. There are two affective experiences, also, *equivalent to* the two broadcasting events.

I have obviously just been assuming some sort of cognitivist theory of emotion that distinguishes between an affective representation of some sort, on the one hand, and on the other hand a representation of the cause or meaning of that affective state. But of course the main point I am making here still holds even if we identify the two broadcasting events not with two affective components of two emotions but simply with two emotions, period. Given what I suspect is the complexity of those affective representations, I lean towards identifying them with conscious emotions. But that does not alter my fundamental argument. What is important to the argument is that neither the response-signaling event nor the response-representing event are themselves conscious or components of conscious experience, as opposed to partial *causes* of conscious experience.

If all this is right, then it has more widespread applicability than just the split-brain case. For it would appear that for any consciousness structured like that of the split-brain subjects, where there are two global workspaces possessing largely different conscious contents, and then a third set of contents making it into both workspaces, the conscious representations carrying those contents are to be identified with the *making it into* the workspaces. I.e., an experience is not an experience because it is *going* to be broadcast. It is not something that exists prior to the broadcast, and which would exist, albeit not as a conscious experience, even if broadcast were not possible. Again, experiences are broadcasting events.

Now by stipulation, there is still a single response-signaling event in the amygdala, and still a single response-representing event in the amygdala. These (jointly or separately) are common causes of the broadcast affective states and emotions. But, again, they are not themselves conscious affective states or emotions. In fact, setting aside for the moment any concerns about consciousness (much less consciousness in split-brain subjects) in particular, and while acknowledging the lack of a broadly accepted account of emotions and of their neural basis, it looks unlikely, at this point, that *any* subcortical activity *alone* will be identifiable with an affective or emotional state. Emotional networks look highly complex, and while they do seem to crucially involve subcortical structures, the activation of subcortical networks involved in emotional experience itself appears inextricably linked (in its inputs and its outputs) with cortical processes.

Finally, recall that in the last major section of Chapter Three I noted that it isn't clear to what extent subcortical (and non-cortical) structures mediate

interhemispheric interaction per se, and to what extent they mediate “horizontal” patterns of brain activity, rather than simply being crucial elements in vertical loops of activity. It is possible, that is, that the two hemispheres of a split-brain subject appear relatively “unified” with regard to emotional experience in large part because of the physiological aspects of emotional experience. Non-cortical structures of course play an essential role in producing and sensing these physiological responses. But they could do this, again, while not directly interacting across the midline. Their interacting would be indirect, via patterns of bodily responses.

My arguments against the partial co-consciousness model are tentative and inconclusive. What I have just suggested about the role of non-cortical structures in providing the two “disconnected” hemispheres with “unified” emotional experience may be incorrect, and moreover, the general account of co-consciousness that I have adopted may be wrong also. It is not possible to say anything definitive against the partial co-consciousness model, and indeed there is no reason to do so. Contra what some critics have suggested (Hurley 1998, 1994; Bayne and Chalmers 2003), the model has some theoretical appeal and may ultimately be easier to work with and provide a more parsimonious model of split-brain subjects’ consciousness.

Still, even if Lockwood’s partial co-consciousness model is the best model of split-brain subjects’ consciousness, split-brain subjects would still have two minds. Imagine that Lockwood is right, and that V.P., for instance, has a single set of affective states and even emotional experiences that are shared by her two hemispheres, which otherwise possess distinct experiences and so forth. If this is correct, then there are “direct” interactions between V.P.’s hemispheres, in the sense

of “direct” first defined in chapter two. But direct interaction and integration at the level of affect and emotion does not alone suffice to make a single mind out of what are otherwise two distinct sets of systems. V.P. still possesses both a right and a left hemisphere mental architecture, each one of which meets the criteria for mindhood; between most of the systems in each architecture, meanwhile, there is *not* direct interaction. Some number of shared mental tokens, token experiences, that is, may suffice for a partially unified consciousness. But they will not suffice for a single mind. Rather, they will simply afford an unusually intimate degree of connectedness between two whole and distinct mental systems.

That said, the fact that the two hemispheres are associated with two different mental systems—indeed, two minds—provides one additional reason to adopt the conscious duality model. Even if we assume that some subcortical structures make the very same conscious affective or emotional contents available to both hemispheres, and via a process that does not supervene on any physiological response, still, the contents are then utilized by two functionally (and of course physically) distinct sets of mental activities. The very fact that a putative mental token is available to two different minds—two minds that aren’t functionally identical at that—provides a reason to question our having individuated only a single mental token in the first place.

3 *The Dynamic Singularity Model*

In *Consciousness and Action* (1998) and other works (2003, 1994), Hurley develops an account of the mind and of mental processes at what she calls the personal and

subpersonal, or conscious and unconscious, level. She does this while criticizing what she calls the traditional or “input-output” model of mind, which she thinks has several objectionable features. First, the traditional model depicts personal-level cognition as sandwiched between perception and action (or intentions to act), and subpersonal-level cognitive processing as sandwiched between sensory input and motor output. But according to Hurley, rather than being “vertically modular” (such that, for instance, perception or sensation could be damaged, leaving cognition and action still more or less intact), the mind is “horizontally modular”: if one particular aspect of perception is impaired, there will be a corresponding impairment in closely linked aspects of cognition and action. (Consider neglect, as an example; neglect patients may not only fail to draw the left sides of objects but fail to engage in leftward movements or actions involving the left side of space.) Furthermore, in Hurley’s work, cognition—as something distinct from perception and action—is largely absent. Like proponents of the “Extended Mind” view (Clark and Chalmers, 1998; Clark, 1997), which will be discussed in Chapter Six, Hurley sees the mind not as something entirely internal to organisms, not as an “interface” between the world-involving processes of perception and action: perception and action *are* the basis of cognition, essentially the entire basis of (at least the personal-level) mind.

There is something else wrong with the input-output picture of mind, too. This concerns the way in which it relates the personal to the subpersonal level. For Hurley, subpersonal-level processes are the vehicle of personal-level mental processes, and I think she thinks this is true of the input-output model of mind, too. But the input-output picture of the mind views sensory inputs as the vehicles of conscious

perception, and views motor outputs as the vehicles of action (or intentions). In actuality, Hurley argues, perception and action do not map neatly onto sensory inputs and motor outputs, respectively. This is in part because subpersonal-level mental processes are best understood as “involving complex relations between sensory inputs and motor outputs and a circular or looping structure of causes, effects, and feedback effects on further causes, with both internal and external feedback loops” (1998: 206-207). Hurley calls this a “*dynamic singularity* in the field of causal flows, centered on but not bounded by a biological organism” (1998: 2007; original emphasis). Subpersonal-level processes meanwhile *are* the vehicles of the personal-level processes of perception and action. But perception and action can’t just be mapped, one-to-one, onto causal (sensory) input and causal (motor) output, respectively, *because* there’s no systematic division of input and output at the subpersonal level. Perception, therefore, maps onto both sensory input and motor output, and action or intention does, too.

Hurley believes that the rejection of the input-output picture of the mind has significant implications concerning the structure of consciousness, among other things. Conscious perception, again, is constructed not just out of sensory input but out of motor output as well. Hurley takes what Bayne (2001) calls an “objectivist” position on the unity of consciousness, or on co-consciousness: she does not think that phenomenal unity can be characterized in wholly phenomenal or indeed any person-level terms. There are person-level constraints on co-consciousness (including of coherence), but there are additionally subpersonal-level, structural constraints on co-consciousness. Hurley does not believe that neuroanatomy, or indeed anything

internal to the organism, can (or at least can necessarily) provide these latter sort of constraints. Rather we must look for a particular sort of *functional* unity to provide them. And it is possible, she says, that this functional unity may include processes that extend outside the brain (and indeed perhaps the body).

The split-brain phenomenon enters Hurley's works at multiple points. As we have already seen in Chapter Three, she uses a comparison between split-brain and callosal subjects to reject Lockwood's "isomorphism thesis," and, relatedly, to support what she calls "vehicle externalism" about consciousness—the view that the structure of consciousness can supervene in part upon actions or behavior. The debate about vehicle externalism is one that I will discuss in the next chapter. But Hurley also uses the results of some split-brain experiments (Trevarthen 1974b, Sergent 1990) to motivate rejecting the input-output model of mind and accepting what I shall call the "intentions-content thesis. This is the thesis that the contents of consciousness can vary, "directly," with motor intentions. She offers a new interpretation of Trevarthen's and Sergent's results, and of the structure of consciousness in split-brain subjects. I call this the "dynamic singularity" model of consciousness. Hurley argues that some of the debate between advocates of partial co-consciousness and of conscious duality can be best resolved by adopting this dynamic singularity model of consciousness. If that is correct, then perhaps, at least at many times, a split-brain subject has only a single, *intra-hemispheric* stream of consciousness.

3.1 Conscious Experience: the Input-Output Picture versus the Intentions-Content thesis

Hurley believes that our intuitive picture of conscious perception is one according to which it is located downstream of sensory input and upstream of motor output.

Hurley provides numerous examples (some quite convincing, and others much less so) meant to test this intuition. If the input-output picture is wrong, she says, then we should accept what I am calling the “intentions-content” thesis. This thesis says that intentions can alter the contents of consciousness “directly” (or “non-instrumentally”)—i.e. even without altering sensory inputs. (That motor intentions can “indirectly” alter perceptual contents is obvious; motor intentions to clap can lead to clapping, which can lead to an auditory experience of hearing oneself clap.)

Hurley concedes that “the idea that the contents of consciousness can change noninstrumentally with our intentions is already familiar and not at odds with widespread assumptions. Consider aspect shifts: we can often if not always intentionally control whether we see a figure as a duck or a rabbit” (1998: 198). She believes, however, that her intentions-content claim is more radical and unintuitive, because we know that aspect shifts involve intentions about *how* to (consciously) perceive something. Hurley’s claim is that all sorts of motor intentions, even those that don’t concern how to consciously perceive something, can in fact directly alter what is conscious perceived. A large part of her work is devoted to defending this thesis.

But Hurley’s work actually contains a multitude of strands and positions that hang together thematically, and that may help support each other, to some extent, but that may also be separately accepted or rejected. For instance, Hurley takes the intentions-content thesis to challenge vehicle internalism. But of course one could accept the intentions-content thesis while remaining an internalist: motor intentions are, after all, internal to the organism. In fact Hurley’s most minimal version of the

intentions-content thesis is at least as compatible with internalism as it is with externalism; the intentions-content thesis, again, says that intentions can alter the contents of consciousness even without altering sensory input—even, that is, without resulting in any kind of *non-neural* activity.

The intentions-content thesis and the input-output model of consciousness may also stand or fall independently of each other, at least on to certain interpretations. Hurley describes the input-output picture of consciousness as stating that consciousness is the *product* of sensory input, and *produces* motor output, i.e. action. (That is, in the input-output model of consciousness, says Hurley, there can be no change in conscious experience without change in sensory input. Or, at least, motor intentions can't make the difference, though perhaps something else (e.g. shifts of attention, or the application of concepts) can.) The intentions-content thesis, if true, challenges *this* input-output model of consciousness.

And indeed I am certain that the input-output model of consciousness just described is *not* true; I am certain, that is, that conscious perception is *not* just a product of sensory input. There is every reason to believe (and Hurley herself makes a strong case that) conscious experience is the product of highly complex processing that is itself sensitive not only to sensory input but also to motor (or motor intentional) factors—and, for that matter, to executive and attentional and conceptual factors as well. In that sense, then, the intentions-content thesis does challenge the input-output picture of mind.

But I am not sure who accepts a characterization of consciousness according to which experience is a product of sensory input alone. This “input-output” model of

consciousness looks something of a caricature. Viable models of consciousness—including some that Hurley herself would reject—should indeed acknowledge that many things, including motor intentions, can alter the contents of consciousness even without altering sensory input.

A more sensible version of the input-output picture of consciousness might just say that conscious contents are perceptual contents. This more sensible version of the input-output thesis could acknowledge that these conscious perceptual contents may themselves be the result of calculations that factor in motor intentions. This version of the input-output picture could for that matter acknowledge that motor intentions play a causal role in determining which perceptual contents become conscious. Thus a sensible version of the input-output model of consciousness is perfectly compatible with the intentions-content thesis—or at least with one version of that thesis.

But there is admittedly more to Hurley's intentions-content thesis than what has been described so far. So far I have described what we could call the *causal* version of the intentions-content thesis, which again just says that conscious contents can vary, directly, with changes in motor intentions. Hurley herself offers a *constitutive* version of the intentions-content thesis, which says that motor intentions are part of the vehicles of conscious experience. On this understanding, motor intentions don't just causally influence perceptual contents. Motor intentions are as much a part of conscious experience as perceptual contents themselves.

The constitutive intentions-content thesis may then challenge even the reasonable version of the input-output model of consciousness. According to the

input-output model of consciousness, reasonably interpreted, the perceptual processing that results in conscious experience may well be sensitive to motor intentions, and use information about motor intentions to make adjustments to its contents and calculations. But this is different from claiming that those motor intentions are *part of* its contents.

Hurley never makes a good case for accepting the constitutive version of the intentions-content thesis over the causal. It is one thing to say that processes at the subpersonal-level are characterized by a multitude of feedback loops between input and output. And it is one thing to say that conscious experience is in part a function of the interaction between causal input and causal output. But it is another thing to say that conscious experience *is*, i.e. is constituted by, both causal input and causal output. Why should we say this, rather than simply saying that subpersonal-level processes—whatever the role that motor intentions and even motor behavior play in them—causally contribute to the generation of conscious perceptual contents? Hurley shows close relationships between causal output (or in some cases just motor intentions, which is arguably not the same thing) and perception. But I'm never convinced that any of these relationships are that of jointly *constituting* conscious experience. I suppose Hurley would just accuse me of being in the grip of the input-output picture of mind, which sees conscious experience as located at the interface of causal input and causal output—but she doesn't ultimately do enough to challenge that picture.⁶⁸

⁶⁸ In characterizing subpersonal-level processes according to the input-output model of mind, Hurley equates causal input with sensory input (and causal output with motor output). But not all causal input *to conscious perception* need be *sensory* in nature. Some of the causal input *to perceptual processing* may be attentional, cognitive, motor-intentional, etc.

This is mostly all just prefatory, to note, again, that various strands in Hurley that might seem to stand or fall together in fact require separate examination and acceptance or rejection. Two strands that are like this, I argue in what follows, are the dynamic singularity model of split-brain subjects' consciousness and the intentions-content thesis (either the causal or the constitutive version, though I will just speak about the causal version). Intentions may well alter the contents of consciousness directly or non-instrumentally, i.e. without altering sensory inputs. They might even in part constitute the contents of consciousness, although again, Hurley provides insufficient reason to think that they do. But this is distinct from the claim that motor intentions in split-brain subjects play *so* significant a role in altering conscious experience that they result in conscious experience, entire, shifting to one hemisphere, leaving the other wholly non-conscious. This is highly unlikely to happen as a general rule. In fact I don't even think it's the best analysis of what's happening in the particular split-brain experiments Hurley discusses.

3.2 The Intentions-Content Thesis and the Dynamic Singularity Model

If motor processes and motor intentions can alter the contents of consciousness, then some of the behavior split-brain subjects exhibit during the split-brain experiment can be reinterpreted as resulting not from conscious duality, nor from partial co-consciousness, but rather from conscious singularity, albeit perhaps a conscious singularity in which only one hemisphere at a time is the site of conscious perceptual contents. I call this the "dynamic singularity" model of split-brain subjects' consciousness, for according to this model a split-brain subject normally has a single,

intra-hemispheric stream of consciousness, but one that shifts from hemisphere to hemisphere, from moment to moment.

Now Hurley herself applies the intentions-content thesis to the split-brain phenomenon largely to question the need to appeal to a *partial co-consciousness* model in particular. Nonetheless, particularly in discussing a commissurotomy study by Trevarthen (1974b), she does suggest that neither the partial co-consciousness nor the conscious duality model need be appealed to in order to explain a split-brain subject's behavioral dissociation under experimental conditions. Rather, she says, we can understand the subject's conscious experience as *moving* from one hemisphere to the other with a change in which hemisphere is motorically active.

Here is how the structure of the rest of this section will go. First, I will discuss "Trevarthen's case", which intuitively makes the best case for the dynamic singularity model of split-brain subjects' consciousness. I suggest, though, that there are other equally plausible interpretations of this case that don't refer to this model. Next, I turn to "Sergent's cases" (Sergent 1990). Sergent's cases actually intuitively make a better case for Lockwood's partial co-consciousness model than they do for the dynamic singularity model, as Hurley herself seems to acknowledge. But a fuller look at the data suggests that Sergent's cases make a stronger case yet for the conscious duality model.

Ultimately I think that the intentions-contents thesis, if not the dynamic singularity model, is plausible, though I suggest that motor intentions may influence the contents of consciousness at least largely through the intervening action of attentional mechanisms. After discussing the cases mentioned above and the

intentions-content thesis more generally, I will then say why it is unreasonable to suppose that split-brain subjects always have a single stream of consciousness, or even that they have a single stream of consciousness outside of experimental situations. For, first of all, it is unlikely that consciousness in one hemisphere ceases altogether even in the absence of any motor intentions in that hemisphere, and, second of all, the two hemispheres seem capable of sustaining (and even acting upon) distinct sets of motor intentions simultaneously. Indeed it is likely that they often do act upon *token-distinct* motor intentions simultaneously.

3.3 Motor Intentions and the Dynamic Singularity Model: Trevarthen's Case

In this subsection I discuss Hurley's application of the intentions-content thesis and the dynamic singularity model to split-brain subjects in two different studies (one by Trevarthen (1974b), and one by Sergent (1990)). The intentions-content thesis may well be supported in the first instance, but it is not clearly supported by Sergent's results. And the dynamic singularity offers the best model of split-brain subjects' consciousness in neither case; the conscious duality model does, instead.

A split-brain subject, N.G., is staring at a central fixation point while holding a pen in her left hand (RH).⁶⁹ An image of an object is projected in her right visual field (RVF/LH). The subject says (LH) she can see the object, unsurprisingly. N.G. is then asked to use the pen, still held in her left hand (RH), to point to the center of the shape. When she attempts to comply, however, she is unable to; "It's disappeared!"

⁶⁹ Actually both hemispheres have a fair degree of motor control even over the ipsilateral hand. The results of this experiment suggest, however, that N.G.'s right hemisphere controlled her hand in this experiment. This may be because of the spatial nature of the task she was asked to perform with that hand; the RH generally dominates at such tasks.

she exclaims (LH). In fact, even as the shape is slid, slowly, towards her LVF (RH), she remains motionless. As soon as the shape slides to or past center, however, her left hand (RH) jumps into action, pointing to the center of the object as requested (Trevarthen, 1974b).

Hurley says that Trevarthen's own interpretation of the case is one involving partial co-consciousness. (I am not sure I agree with Hurley's interpretation of Trevarthen's interpretation, but that isn't really relevant.) Here is how Hurley describes the case according to the partial co-consciousness model (or as she calls it, following Lockwood, the "partial unity" model):

suppose that right- and left-visual-field contents are both conscious. And that while right- and left-field contents were not co-conscious with each other, both are co-conscious with one set of background intentions that are continuous across the discontinuity in perceptual awareness, such as linguistic intentions, intentions to cooperate with the experimenter, and so on. (Hurley 1998: 196)

Different motor intentions (to speak, to point) arise over the course of the experiment, and these are co-conscious with different visual experiences. But visual experiences, motor intentions, and background intentions are all conscious: they are simply not all *co-conscious* with each other.

Unfortunately this study, involving N.G., arguably provides a poor, and somewhat confusing, case for the application of the partial co-consciousness model. I suppose the "background intentions" could be co-conscious with both right and left hemisphere visual experiences. But, first of all, why think this is any more likely than that right and left hemisphere both have (distinct) background intentions to

cooperate? (This is what the “conscious duality (with partially redundant contents) model” would say.) Second of all, *intentions to cooperate* might not be components of streams of consciousness at all. Certainly the major advocates of the partial co-consciousness model—including Lockwood, of course, but also perhaps including Trevarthen himself, —see as paradigm cases of partial co-consciousness those in which the hemispheres appear to share access to at least the same *types* (including contents) of *perceptual* information. (In the same article, for instance, Trevarthen suggests that certain *visual* information—concerning the “ambient visual field”—is inter-hemispherically co-conscious. And Sperry (1986), recall, seemed to suggest that emotional or affective information might be co-conscious with both right and left hemisphere experiences.)

Hurley provides her own model of N.G.’s consciousness during this experiment, however—one according to which (she says) both the partial co-consciousness interpretation and the conscious duality (with some duplication of contents) interpretation “are preempted” (1998: 197). According to this interpretation, which I am calling the “dynamic singularity” model:

before the patient forms the immediate intention to move, background intentions are co-conscious with left-hemisphere perceptual content, not with right-hemisphere motor and perceptual content. Once the patient forms the immediate intention to move, background intentions are co-conscious with right-hemisphere motor and perceptual content.
(1998: 197)

There are five sets of mental events that can be or fail to be co-conscious with each other in this story: background intentions to cooperate with the experimenter,

motor intentions to speak (LH motor intentions), motor intentions to point with the left hand (RH motor intentions), LH visual representation (initially of the object), and RH visual representation (initially of a blank screen). There are also three time periods or stages at which the content and structure of N.G.'s consciousness can be analyzed. At stage one, N.G. is asked to describe what she sees; she hasn't yet been asked to do anything with her hand. The shape is in her RVF (LH), and so she says (LH) she can see it. At stage two, the shape is still in her RVF (LH), but now the subject has been asked to point to the center of the shape (RH); she now says (presumably LH) she can't see the shape. At stage three the shape is moved into her LVF (RH). I will only discuss the first two stages, because these are sufficient to illustrate the dynamic singularity model of what's happening.

Figure 5.2 diagrams the first two stages, according to Hurley's dynamic singularity model. At stage one, a set of verbal motor intentions (LH) are active, and at stage two a set of motor intentions concerning the left hand (RH) are active. Notice that in this diagram, co-consciousness is transitive (as it *wouldn't* be according to the partial co-consciousness model).

At first, background intentions to cooperate are co-conscious with *one* set of motor intentions—motor intentions to *speak*. Since these are left hemisphere intentions, those intentions, and the background intentions to cooperate and so forth, are co-conscious with *left* hemisphere visual experience, and not with right hemisphere visual experience. N.G. therefore reports *left hemisphere* visual experience. Now a *new* set of motor intentions is created, however: motor intentions to draw the object, using the left hand. Assume, as Hurley does, that these motor

intentions are located in the right hemisphere. Now the background intentions to cooperate are co-conscious with this *new* set of motor intentions—motor intentions to draw (using the left hand). Since these are right hemisphere intentions, they, and the background intentions to cooperate and so forth, become co-conscious with *right* hemisphere visual representation, and cease to be co-conscious with *left* hemisphere visual representation. Therefore the subject *intends* to cooperate by pointing (RH) to the object, but cannot, because the shape is not yet in the LVF (RH).

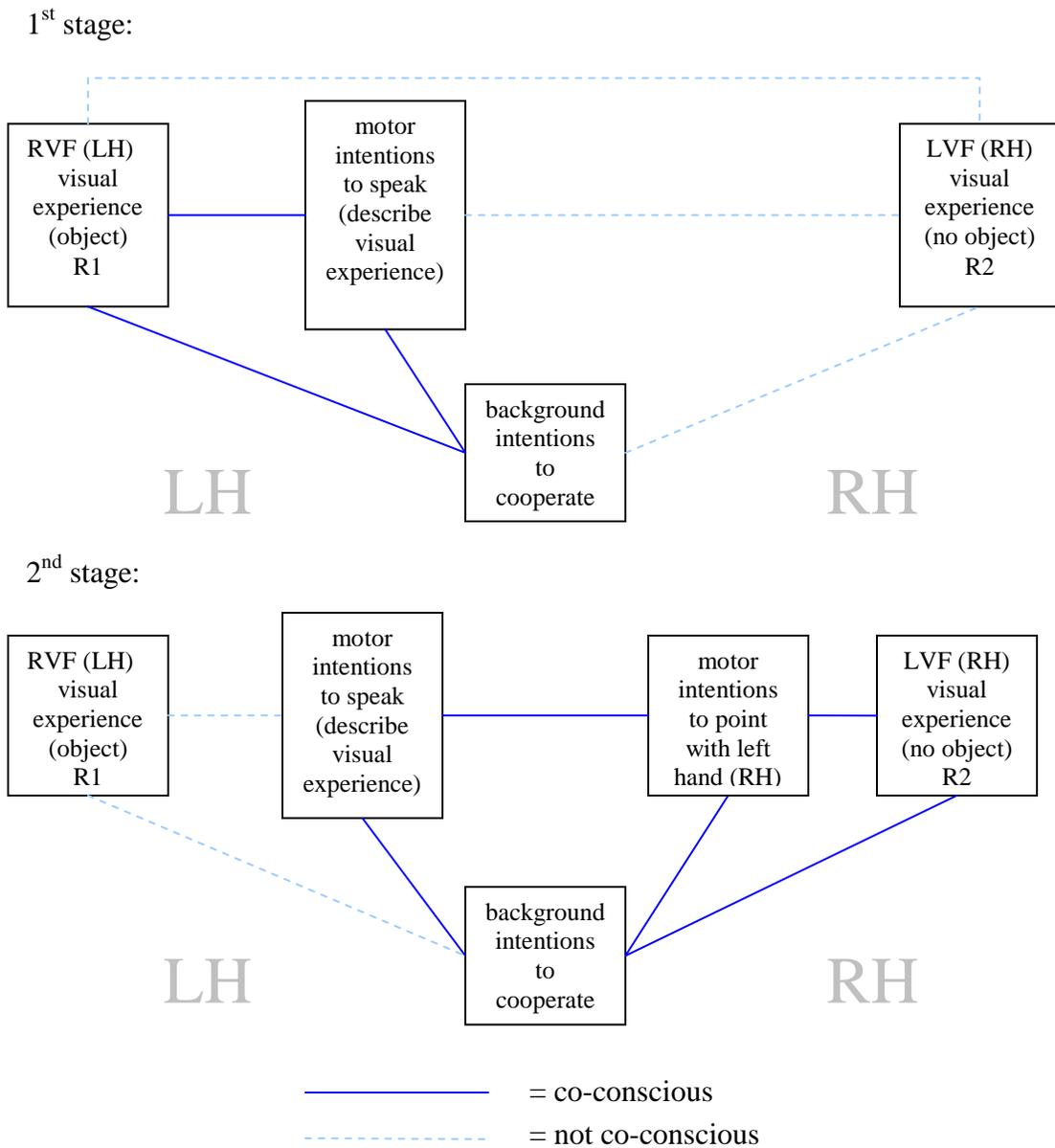


Figure 5.2: The Dynamic Singularity Model: Trevarthen's Case

Of course, the subject doesn't just fail to point to the object at this state two; she also claims that the shape has vanished. So she has a set of verbal motor intentions at this point also. Now perhaps these motor intentions are actually *right*

hemisphere motor intentions. (Some split-brain subjects do develop some spoken language in their right hemisphere.) This is possible but probably unlikely (I return to this below), and Hurley does not believe this. Rather, she says that this case involving N.G. suggests that *left hemisphere verbal intentions* can be co-conscious with either LH or RH visual experience.

(This might seem create a problem for Hurley's own, "dynamic singularity" model of the case, however. Hurley is going to say that the contents of consciousness shift depending on which motor intentions are active. But then why don't left hemisphere *verbal* motor intentions make left hemisphere visual contents conscious? But perhaps this isn't a problem; different sorts of motor intentions might plausibly play a greater or lesser role in altering conscious contents.)

Since the shape is still in the RVF (LH) prior to stage three, one would expect N.G. to claim (LH) that she can still see the object, even if she can't point (RH) to its center. But N.G. claims (LH, by assumption) that she *can't* see the object. Why? Because, Hurley says, her background intentions to cooperate (to point, but also to speak, to respond, to describe her visual experience when asked) are no longer co-conscious with her RVF (LH) visual representation: these background intentions have become co-conscious with her LVF (RH) visual experiences, because of new RH motor intentions (to point).

But of course, I have so far left out of this application of the dynamic singularity model what Hurley considers its most interesting claim. And that there is only one set of conscious visual experiences at each stage. At stage one, left hemisphere visual experiences are conscious. At stage two, right hemisphere visual

experiences are conscious. But there are no other, contralateral conscious visual experiences at either of these stages.

Hurley essentially just questions whether the right hemisphere's visual experience is really conscious at stage one, and whether the left hemisphere's visual experience is really conscious at stage two. Perhaps motor intentions and/or background intentions to cooperate create changes in conscious contents. (Possibly by partially *constituting* conscious experiences.) Hurley believes that it is tempting to view Trevarthen's cases in terms of partial co-consciousness just because "it is natural to pass over the temporal distinction marked by motor intentions, given that relevant sensory inputs are held constant" (1998: 198) In other words we assume that whatever perceptual states of the subject N.G. and of her hemispheres are conscious at stage one, these same perceptual states are conscious at stage two, and the only question becomes one concerning the *co-consciousness* relations between these various conscious experiences. We assume this, Hurley says, because at stage one and stage two, the subject is receiving the same sensory input; the shape is present and in the RVF (LH) at both stages, and N.G. is still staring at the same fixation point.

But what if the shift in motor intentions that occurs between stage one and stage two marks a shift in conscious contents? In that case, there need not be two sets of visual experiences—a visual experience of the shape, and a visual experience of a blank screen—that are conscious at both stages but just not co-conscious with each other. Perhaps at each stage there is one conscious visual experience: of the shape, and associated with the LH, at stage one, and of the blank screen, and associated with the RH, at stage two.

I have just described Trevarthen's case according to the dynamic singularity model. (The case, as Hurley interprets it, also illustrates the "intentions-content" thesis, though note that there is nothing in the description to favor the constitutive over the causal version of that thesis.) And the dynamic singularity model makes *some* sense of the case, although it also requires accepting that when N.G. says (LH) the object has disappeared, she is describing her RH's visual experience. (Hurley apparently just accepts that this is the case, stating that, "Information from the motorically active hemisphere, whether left or right, can be reported by the subject verbally, that is, using the left, linguistic hemisphere" (1998: 195). But then why are there so many split-brain experiments—including other experiments using this particular subject—in which subjects *can't* verbally describe their RH experiences? And why does Sperry (1990) mention that it is not uncommon for the left hemisphere to keep up a stream of chatter only on the aspects of an RH task that the *left hemisphere* has perceptual access to—even when the RH is engaged in a motor task?)

I am not convinced that the dynamic singularity model of split-brain subjects' consciousness offers the clearest interpretation of the results of this study. Admittedly, the data is difficult to explain for any model of consciousness. Among other things, given the spatial nature of the task, it's a little surprising that the subject couldn't perform the pointing task at all, even under dominant control of her right hemisphere, and even if her right hemisphere could not consciously perceive the shape. For this seems like the sort of thing the right hemisphere (although perhaps not the left) would be able to do, in part due to midbrain representations of the visual field. (See for instance Corballis 1995b; or see Ishiai, Kyama, and Furuya, 2001, for a

discussion.) It's surprising, that is, that N.G. had to wait for the shape to enter the LVF (RH) in order to perform the pointing task. Given this fact, I'm somewhat reluctant to speculate about this one, potentially non-representative case. But I do still think that it can be interpreted under the conscious duality model.

While I don't at this point *have* an explanation as to why N.G. could not perform the pointing task until the shape was *consciously* perceived by her RH, I don't see that Hurley has one either. (I am not familiar with the details of N.G.'s case history, but perhaps there is something in that history which would explain this?) But for whatever reason, N.G. couldn't target the center of the shape, at least with her left hand, until the shape was in her LVF (RH). Now, why did N.G. claim that the object had disappeared as soon as she attempted to point to its center—and even though the object was still within the visual field of her speaking hemisphere?

I suppose one possibility is that N.G.'s exclamation that she could no longer see the object was actually a RH vocalization. I am not sure how likely a possibility this is. Trevarthen (1974b) describes N.G. as having said, "It's disappeared!" and as "continu[ing] to report that the stimulus in the right field was not visible, even while it was moved slowly about by the experimenter" (Trevarthen 1974b: 196), and this sounds too sophisticated for an RH vocalization. (See Levy, Nebes, and Sperry, 1971, however.) Note that Sperry, Zaidel, and Zaidel (1979) at one point at least considered the possibility of spoken language—in particular, short verbal exclamations—in this particular subject; I don't know if they later reached some sort of more decisive view about RH language in N.G.

A second, more likely possibility is that N.G. was simply embarrassed to find herself unable to point to an object she had just stated that she could see—a task any child could perform. She may then have simply stated that the shape had vanished, either through deliberate dishonesty or in mere confusion. Alternatively, there is Trevarthen's own interpretation of the experimental result, offered "in terms of interlateral, probably interhemispheric, communication of an inhibitory nature, leading first to erasure or occlusion of the right field percept from consciousness, and then report of loss emanating from the left hemisphere" (1974b: 196). One version of this third alternative explanation is that, in preparation to perform as requested, the right hemisphere diverted significant attentional resources from the left hemisphere, or at least from the left hemisphere's visual areas, making the left hemisphere's percept of the object non-conscious.

According to either of these latter two interpretations, the left hemisphere's percept of the object may have indeed become non-conscious, at stage two—but the left hemisphere was still reporting the contents its *own* stream of consciousness, and not that of the right hemisphere. Note that if this interpretation is correct, the intentions-content thesis is affirmed—motor intentions really *have* made a difference to conscious contents, directly—but the dynamic singularity model is not, for there are still two streams of consciousness, and the left hemisphere is merely reporting the contents of its *own* stream. This is the more conservative interpretation of the case.

There is another result of Trevarthen's (Levy, Trevarthen, and Sperry, 1972) that I don't believe Hurley discusses but that seems to make an equally strong case for the dynamic singularity model (however strong that case ultimately is). In this

study bilateral chimeric stimuli were presented tachistoscopically for brief periods to split-brain subjects, with a different half-object, or a different word or pattern, just on either side of the vertical meridian. When asked to *say* what they'd seen, split-brain subjects described the *RVF* (LH) stimulus. When asked to *point*, with either hand, "to a picture matching the stimulus in form or appearance, the left half of the chimera, seen by the right brain, was chosen" (Trevarthen, 1984: 333). This pattern of results is easily explained by the intentions-contents thesis. When the motor intentions in question are associated with the left hemisphere (i.e. intentions to speak), left hemisphere visual percepts are conscious. But the right hemisphere is dominant for many visuo-spatial, including drawing, tasks; perhaps, then, when asked to *draw* an object, *right* hemisphere motor intentions are generated (or dominant), leading *right* hemisphere visual percepts to become conscious. This pattern of results, again, lends itself to an interpretation in terms of the dynamic singularity model.

Of course it equally well lends itself to an interpretation in terms of conscious duality: whichever hemisphere is dominant for the *response type* indicates *its* conscious contents. As data points for the dynamic singularity model of consciousness, then, both of Trevarthen's cases suffer from the same problem: the dynamic singularity model offers a workable interpretation of the results; there are just other equally good interpretations possible. In fact some of these other interpretations are arguably better, if only because they are supported by a great deal of other experimental results with the same subjects.

3.4 Motor Intentions and the Dynamic Singularity Model: Sergent's Case

Sergent hypothesized that the experimental paradigm that first found evidence of conscious dissociation in split-brain subjects may have overestimated the extent of conscious dissociation in such subjects by requiring them to *name* right-hemisphere lateralized stimuli. Her studies on split-brain subjects exploited different types of paradigms, in order to test whether the degree of conscious dissociation found depended upon the type of question subjects were asked.

Sergent (1990) asked subjects to cross-compare two bilaterally presented stimuli in two different ways and in two different trial types. In both sets of trials, two numbers were flashed simultaneously to a subject, one in each visual field. When asked to push a button indicating whether the numbers were same or different, split-brain subjects were at chance. This is a standard result for split-brain subjects: they can't cross-*identify* objects. Sergent wondered, however, whether they might be able to cross-compare objects in some other way. In a second trial type, subjects were asked to make a motor response indicating not whether the *identity* of the two numbers was the same, but rather which of the two numbers was greater in *value*. The object of this study was to see if the hemispheres were capable of passing some abstract information associated with a stimulus, even if the stimulus identity (either in the form of a name or in the form of perceptual representation of the stimulus) could not be passed. And, as it turned out, subjects performed well above chance (though not perfectly) at this task.

An impressive, even startling, amount of transfer thus seemed to obtain in this study. Perhaps part of what was impressive was the *kind* of information that seemed

to transfer between hemispheres: not information about the identity of the numeric stimuli, but at least information about their quantitative value. Sergent hypothesized that while *perceptual* information about stimuli could not transfer, other, more “abstract” information about stimuli could. Take a trial in which a “3” is presented in the RVF (LH) and a “6” in the LVF (RH). The *identity* of the RH stimulus did not transfer to the LH, nor vice versa; so presumably the *shape* of the RH stimulus did not transfer to the LH, nor vice versa; yet somehow the *value* associated with one hemisphere’s stimulus *had* transferred to the other hemisphere, apparently.

As Hurley notes, Sergent herself did not suggest that the information about the “6” that reached the left hemisphere, say, via some subcortical structures, was *conscious*. (Sergent did not specify which subcortical structures she thought might be playing this role.) It was still possible that it was not conscious. But an interpretation of Sergent’s results in terms of partial co-consciousness is, as Hurley notes, quite tempting. At any one moment, a split-brain subject perceiving “6” in her LVF (RH) and “3” in her RVF (LH) would have a conscious percept of the numeral on the left as a 6, which was co-conscious with a “perception” (as Hurley puts it) of that number as higher (than the other number), which is also co-conscious with a perception of the number on the right as a 3—but the perception of the number on the left as a 6 and the perception of the number on the right as a 3 are not themselves co-conscious. Another way of explaining these results using the partial co-consciousness model would be to say that (conscious) “abstract” or semantic information is inter-hemispherically unified, even if conscious perceptual information is not.

Prima facie, then, Sergent's results seems to make a nice case for the partial co-consciousness model. They do not seem to make a good case for the intentions-contents thesis or the dynamic singularity model. In fact, how, exactly, Hurley views Sergent's pattern of results as supporting the intentions-content thesis is unclear. (See Hurley, 1994.) For it is difficult to see what the different motor intentions are in this case. Hurley believes that there is a "change in intention between the same/different and higher/lower tasks in a Sergent-like case [that] is sufficient for a change in perceptual consciousness despite constant sensory input" (1994: 77). But how have the *motor* intentions changed between the two sorts of tasks? As Hurley herself describes the *two* trial and task types, "Sergent's series of experiments projected pairs of numerals simultaneously, one numeral to each half-brain, and asked patients to compare them. Either hand was allowed to reply by pushing a lever to indicate the correct response in a given task" (Hurley 1994: 52). The motor intentions in the two types of task are the same, then: using whatever hand you wish—and, by the way, it apparently makes no difference which hand you use; subjects can't do the same/different task with either hand, and they can do the greater/lesser task with either hand—push the appropriate lever. *Something* has changed in between the two trial types, of course, and admittedly it isn't the stimuli themselves. But I wouldn't call what differs *motor intentions*. What changes is just the *type* of information about the stimulus that the subject is asked to consider and compare. It is some sort of representation *of the stimulus*, and not of a motor act, that changes. So Sergent's results make a better case for the partial co-consciousness model than they do for the intentions-contents thesis or the dynamic singularity model.

Hurley acknowledges that Sergent's results are readily interpreted in terms of partial co-consciousness. But the problem with accepting such a model of split-brain subjects' consciousness, Hurley says, is that *any* putative instance of partial co-consciousness could also be characterized as a case of conscious duality (or multiplicity) with some redundancy of contents. Hurley offers the intentions-content thesis in part as a way out of this "indeterminacy" problem.

I wouldn't call it "indeterminacy," however, and I don't see it as a problem (certainly not for the conscious duality view). Even if there is no behavioral evidence that could ever *rule out* conscious duality with some redundancy of contents anywhere partial co-consciousness appeared a plausible interpretation also, or vice versa, other factors could come into play in our choice of a model of consciousness. For instance, as Lockwood notes, we could turn to neuroanatomy. (Though concededly Hurley believes that we *can't* turn to neuroanatomy, as already mentioned; I have said something already about why I think that we can turn to neuroanatomy, and in the next chapter I give a defense of vehicle internalism or mind-brain supervenience that will support this move even more.) For that matter, we could turn to considerations of simplicity, as I also believe Lockwood also implies (see Lockwood 1989, but also Lockwood 1994, responding to Hurley 1994). If a subject has a hundred experiential contents, for instance, and it appears that two of these contents are not co-conscious with each other, and the rest are, partial co-consciousness looks much more appealing than does conscious duality with almost entirely redundant contents (particularly if the subject is a "normal" rather than a split-brain subject).

In any event, though, once we get beyond first glances, Sergent's results actually don't make a compelling case for the partial co-consciousness model, either; they make a compelling case for conscious duality. Sergent herself noted a methodological flaw with the trials involving the higher/lower task: only the numbers one through nine were presented, and the subjects knew this, and so either hemisphere could employ a simple guessing strategy and be successful a statistically significant 78% of the time—which was about the level of accuracy split-brain subjects in fact achieved. (I.e. 78% correct could be obtained by either hemisphere, or by both hemispheres separately, employing the following rule: if you see a number greater than 5, guess that *that* number is “greater”; if you see a number less than 5, guess that *that* number is “lesser”; if you see a 5, guess at random.) In fact when Seymour, Reuter-Lorenz, and Gazzaniga (1994) redid the study but using digits that were never more than one digit apart (so that subjects were always asked to compare a 2 and a 3, or a 5 and a 6, a 7 and a 7, etc.), subjects did well on the within-field condition (i.e. where both digits were presented within the same visual field) but were at chance on the between-field condition—presumably because the within-hemisphere guessing strategy just described could not be employed. (In a different set of trials in which the two digits presented could be more than one number apart, subject J.W. was between 72.7% and 77.8% correct (depending on response hand) for the between-field conditions, and again, perfect use of the guessing strategy would predict 78% correct on between-field conditions. And in fact, J.W. described himself as using that guessing strategy.)

My conclusion therefore is that these two split-brain experiments with which Hurley tries to motivate the dynamic singularity model do no such thing. While at first glance one of the experiments in question might seem best accounted for by the partial co-consciousness model, upon closer examination, the partial duality model provides the best explanation of both sets of results.

3.5 A Qualified Defense of the Intentions-Content Thesis

Although Hurley's attempts to apply the dynamic singularity model to split-brain subjects are not highly successful, the intentions-content thesis motivating it does not seem implausible, or even (pace Hurley) particularly counterintuitive. (And indeed, at other points in *Consciousness in Action* she provides some persuasive empirical evidence for the thesis.) Perception itself, certainly conscious perception, involves extraordinarily rich feedback, much of whose nature and significance we probably do not know. Among other things, to become conscious, a perceptual state must, it seems, win more than one competitive processes. Some of the competition is local, i.e. some of it involves a competition within the visual system between multiple visual interpretations, for example, and some of it is global: some of it involves a competition for general processing and attentional resources. To respond, deliberately, to a perceptual stimulus, no doubt involves a complex pathway that *does* leave room for feedback at multiple points, and perhaps some of this feedback might alter the strength of the perception, or even make the difference between its becoming conscious or not.

So there is some compelling evidence that at least the causal version of the intentions-content thesis is correct. It is interesting to consider whether motor

intentions might alter the contents of consciousness via the intervention of, or because of their relation to, processes of (especially spatial) attention. Brain imaging studies, for instance, have suggested substantial overlap between structures involved in spatial attention and structures responsible for at least some behavioral responses—including those involving hand and eye movements. (I have not read anything linking the structures of spatial attention with structures for verbal responses.) Some have hypothesized that the “extrastriate body area” integrates “visual, spatial attention, and motor signals for the dynamic updating of the observer’s body representation. Astafiev et al. (2004) considered this process related to the updating of visual space for eye and hand movements that takes place in the posterior parietal cortex” (Praagstra, Boutsen, and Humphreys, 2005: 771). And many people believe that shifts of attention can directly alter the contents of consciousness. Perhaps, then, the generation of motor intentions can have some influence on the contents of attention which can in turn influence the contents of consciousness. (Note that this means of influencing conscious contents would still be “direct” in Hurley’s sense, i.e., not dependent on actual motor behavior resulting in changed sensory input.) This is mere speculation though, and of course the direction of influence could go the other way around only, with attention priming or inhibiting motor intentions.

Still, let me quickly say why, even if the intentions-content thesis is true (in either its causal or its constitutive version), split-brain subjects would still have two streams of consciousness.

3.6 The Conscious Duality Model, Defended

In order for motor intentions to unify a split-brain subject's consciousness, or, as I have been putting it, in order for motor intentions to create a single stream of consciousness in split-brain subjects, one of two conditions must hold:

Condition 1: Motor intentions are only present or dominant in one hemisphere at a time, in such a fashion as to produce a single intra-hemispheric stream of consciousness at any one moment.

Condition 2: Motor intentions are integrated inter-hemispherically in such a fashion as to produce a single inter-hemispheric stream of consciousness.

Note that only the truth of condition 1 would vindicate Hurley's dynamic singularity model, as I understand it, but if condition two held instead, I imagine that she would be nearly equally pleased, as the model of consciousness so resulting would be quite consistent with a number of the themes and arguments in her work.

There is excellent evidence that neither condition 1 nor condition 2 can hold at *all* times; indeed, I won't even bother to review some of this evidence. For we have already seen, in the context of other discussions, numerous instances in which each hemisphere has evidenced a unique set of motor intentions that are not integrated with those of the other—and in which each hemisphere has evidenced this *simultaneously* with the other.

Given that neither of these conditions hold at *all* times, the question becomes whether one of these conditions might still hold at *most* times. Now I will discuss the two cases separately, starting with condition 2.

It might be said that there is some compelling evidence for condition 2, since split-brain subjects generally behave in so integrated a fashion, as if their hemispheres' motor intentions *were* coordinated inter-hemispherically. But the first thing to note about condition 2 is that it assumes a thesis that goes beyond anything that's been presented thus far: the thesis that integration of motor intentions *suffices* for conscious unity (co-consciousness). Granted, Hurley does argue that unity of agency (which includes but need not be exhausted by the integration of motor intentions) is *necessary* for the unity of consciousness, relying in particular upon a discussion of "Marcel's case" (Marcel, 1993). (See especially the discussion of "Marcel's case" in Hurley 1998.) While I am skeptical that Hurley draws the correct conclusion from this case, we should note that even if hers *was* the correct interpretation, this would only show that disunity of motor intentions can *disunify* consciousness—not that integration of motor intentions can *unify* consciousness.⁷⁰

⁷⁰ This was a complicated study with a large number of trial types and conditions, but essentially Marcel found that the *way* in which ("normal") subjects were asked to indicate perception of a stimulus influenced how likely they were to indicate that the stimulus had been perceived. In particular, winking was more accurate than pressing a button, and both were more accurate than verbal response. When subjects were just asked to *guess* whether a stimulus had been presented or not, all response types were more accurate. I question Hurley's and Marcel's interpretation of the study insofar as I question whether the performance difference between report and guess conditions really reduces to a difference between conscious and non-conscious perception. For a number of reasons, however, it seems more likely to me that responses on verbal affirmation trials were most likely to be reports of conscious experience, and that superior performance on winking and button-pressing trials is explained by these motor modes having greater access to non-conscious perceptual information.

The split-brain cases seem to show that integrated behavior can result from two streams of consciousness—and they also show that behavior and motor intentions can appear integrated even when they arguably are not. It is surely true that the motor outputs of the two hemispheres are coordinated non-cortically somehow—in the basal ganglia, the reticular formation, the cerebellum, etc.—to some, possibly great, degree. But that sort of coordination is of motor intentions that probably *don't* shift the contents of consciousness. For as I mentioned in Chapter Four, it appears that a good deal of the representations that guide our actions aren't conscious to begin with. It is the selection of actions to perform, motor intentions framed as the fulfilling of goals, and so forth, that seem most closely linked to conscious perception. But these are cortical motor processes.

Keep in mind, again, that it is not always *obvious* when behaviors *aren't* the result of intimate, inter-hemispheric coordination of motor intentions. Split-brain subjects may even perform “remarkably free of noticeable clumsiness or incoordination” and may display “fine motor precision” on many “unimanual and bimanual tasks” (Zaidel and Sperry, 1977: 196). . . even if their performance is quite slow. Severe impairments may be most obvious only for a particular *sort* of bimanual task: that requiring “interdependent bimanual movements”, i.e. in which “the movements of one hand *depended* upon the sequencing of movements in the opposite hand” (Zaidel and Sperry, 1977: 199, emphasis added; see also Preilowski, 1990). It

is when the two hands *cannot* each just “do their own thing” separately, that is, to contribute to a movement, that performance breaks down.

Motor coordination in split-brain subjects, then, may frequently be a result of cooperation between two streams of consciousness. In other instances, meanwhile, the kind of inter-hemispheric coordination of motor output that’s really at play may be a simple inhibition of one hemisphere’s outputs. Motor control may switch from one hemisphere to the other. (It might be slightly misleading to even call this “coordination”: it results in coordinated *behavior*—but still behavior originating from only one hemisphere. The coordination of motor outputs in this case is at most the inhibition of one hemisphere’s output by the other’s.) Meanwhile, as Sperry points out:

[Various] kinds of evidence further confirm that, while one hemisphere is performing, the nondominant less-active hemisphere, though overtly passive in not exerting control over the motor system, may nevertheless be alert and consciously cognizant of what is going on externally. This is indicated, for example, in disgusted shaking of the head or irked facial expressions triggered from the minor hemisphere after it has heard its speaking partner making what it knows to be an incorrect answer. Also, it is not uncommon, while the informed right hemisphere is performing, for the vocal hemisphere to make remarks like, ‘Now, why did I do that?’ ‘What’s the matter with me anyway?’ (Sperry 1990: 375)

Indeed, Sperry notes in this same article that it is in fact not uncommon for the left hemisphere to maintain “a running commentary based on those aspects of the

situation not restricted to the mute hemisphere” (ibid 374) when the right hemisphere is engaging in some kind of perceptual, cognitive, and behavioral task.

This is not to say that the drawing of attentional resources to one hemisphere can't to some extent depress the conscious activities of the other. This may well happen, to some degree, at some times (as Sperry suggests in the same article, 1990). Of course, just because the left hemisphere shuts down behaviorally doesn't necessarily mean that it literally loses all consciousness of what's going on. (I've never heard, for instance, of a split-brain subject suddenly blinking, after being deeply absorbed in a RH task, and asking (LH), “What just happened? Was I asleep?”, or saying (LH) that for the past few moments, everything was dark.) But in part because of a shift of attention to one hemisphere, the other hemisphere's consciousness may be reduced or at least made into a more passive sort.

But now we have shifted to talking about condition 1: whether it is likely that, at most times, only one hemisphere of a split-brain subject will be associated with a stream of consciousness at all. The split-brain experiment is often set up in such a fashion so as to *require* the simultaneous behavior of both hemispheres. Outside of experimental conditions, though, perhaps only one hemisphere at a time has active motor intentions—perhaps just switching consciousness from one hemisphere to the other, but not sharing it between hemispheres at a time.

This seems implausible on its face. Hurley herself does not seem to believe this, and she willingly acknowledges that the input-output picture “is about right for many cases” (1998: 289) and that perceptual experience does not *always* depend upon motor intentions. Not only does it seem unlikely that competition for attentional

resources would routinely drain one hemisphere of *all* conscious experience, but also, split-brain subjects' performance on many motor tasks in fact seems to require active motor intentions in both hemispheres simultaneously. They're impaired on those tasks requiring that the intentions be integrated in a particular way: when the *precise* intention the LH should form *depends on* which one the RH has formed or will form. Since the hemispheres can't integrate their intentions in this way, they have to rely on sensory feedback: one hemisphere has to look at the actual motor consequences of the other hemisphere's intentions before it acts, resulting in slow, halting, and somewhat ragged-looking performance.

In an experiment by Preilowski (1990), for instance, "normal" and split-brain subjects were each given an Etch-a-Sketch-like device and told to draw a line that stayed between two given parallel lines. Obviously the task is quite simple when the two given lines are either vertical or horizontal; all you need to do is spin a single knob as quickly as you can. When the two given lines are diagonal, the two knobs must be turned in a coordinated way, and the rates of turn for each knob will not necessarily be equal; this will vary based on the slope of the two given lines. But apparently "normal" subjects begin to draw smooth, straight lines, and very quickly, with a little practice. Split-brain subjects draw staircases, painstakingly struggling just to stay within the two given lines.

At the same time, they appear to perform many other kinds of motor tasks successfully (if unusually slowly) using both hemispheres at once. Indeed recall that while they have trouble with motor tasks that require inter-hemispheric *integration* of motor intentions, they are better than are "normals" at *decoupled* motor activity. They

can draw two pictures with two hands at once, or engage in two different sorting tasks with two hands at once, with no obvious interference, as discussed in Chapter Four. Their hands, that is, seem to work perfectly well at the same time, independently. So at least at those moments, they have two streams of consciousness. Even though, as I mentioned in Chapter Four, their behavior might actually look, in a sense, perfectly coordinated. So there is no need to refer to “dynamic singularity” to explain the unified-seeming behavior of split-brain subjects, as I have previously argued.

Motor intentions may have a direct influence on conscious contents, even without effecting any change in sensory input. In some cases intentions may have this effect via the intervention of attentional mechanisms. Because the hemispheres share and compete for at least some attentional mechanisms, active motor intentions may depress the conscious contents of the other hemisphere at some moments. It is unlikely, however, that one hemisphere would be rendered *completely* non-conscious during such moments; at many other moments, meanwhile, motor intentions are either not particularly active in either hemisphere or quite active in both. The kind of integration of motor intentions that does occur between the hemispheres, meanwhile, does not seem to suffice to integrate their conscious contents—to make the hemispheres co-conscious, that is.

4 *Conclusion*

In this chapter I’ve defended the conscious duality model from two interesting and alternative models of split-brain subjects’ consciousness: the partial co-consciousness model and the dynamic singularity model. The latter, I’ve argued, is simply unlikely

empirically. But the former provides a good fit with the empirical data, and indeed it, and the conscious duality model, should make basically equivalent predictions with respect to the behavioral data. I suggested that there may still be reasons to prefer the conscious duality model, for functional considerations of several sorts seem to weigh in favor of characterizing interhemispheric transfer of conscious affect, for example, as involving the generation of the same conscious content but different conscious experiences in the two hemispheres of a split-brain subject. But these arguments are admittedly tentative, and perhaps split-brain subjects have a single stream of consciousness, spanning two hemispheres and shared by two minds, within which co-consciousness is not perfectly transitive.

Chapter 6: Functionalism and the Boundaries of the Mental

1 Introduction

At several points throughout this dissertation the issue of the *supervenience base* of mental phenomena has been raised either explicitly or implicitly.⁷¹ For instance, the claim that something about the split-brain experiment itself changes the structure of split-brain subjects' consciousness is ultimately plausible only if the *vehicles* of consciousness—conscious representations, and not just conscious contents—may be located outside the brain. And Hurley's rejection of the isomorphism thesis—the thesis that there is some kind of isomorphism between the structure of cognition and consciousness and the structure of the nervous system—seems to rest most fundamentally upon her rejection of mind-brain supervenience.

Up until this point, however, I have merely assumed the truth of some kind of mind-brain supervenience claim. In this chapter I finally defend *vehicle internalism*: the view that mental states and minds themselves are internal to the organism, from the criticisms of proponents of *vehicle externalism* (or the “extended mind” hypothesis). This defense supports many of the arguments made and conclusions

⁷¹ *A supervenes on B—or B is the supervenience base of A—if there can be a change in B without a change in A, yet there can be no change in A without a change in B. Constitution is one way that supervenience can come about—if B constitutes A, then A supervenes on B—though constitution is arguably a stronger relationship. (So one might believe that the aesthetic properties of, say, a dress—its flamboyance—supervene on the dress's basic perceptual properties—its shapes, colors, textures—without being constituted by those basic perceptual properties. Thanks to Jerrold Levinson for this example.) In this chapter I argue that minds supervene on brains. Consider my total state of mind right now—call this *M*. This state could not change without *some* change in the actual state of my brain right now—call this *B*. But in fact *B* could change without *M* changing; slight changes in glucose metabolism or in neurotransmitter levels or in oxygen levels, the death of a single glial cell, etc. . . . Any of these things might occur without a change in *M*.*

reached in earlier chapters. Insofar as it unifies several of these arguments, it also serves in part to conclude the dissertation.

Let me define some relevant concepts and distinctions and make a few clarificatory remarks before continuing. First, instead of contrasting the claim of mind-brain supervenience with the extended mind hypothesis, I refer instead to the debate between vehicle externalism and vehicle internalism. This is properly speaking a debate about the supervenience base of the vehicles of cognition and consciousness. By *vehicles* of cognition and consciousness I simply mean token mental (sometimes conscious), representational states, i.e. states carrying content. So, let us say that, if I consciously examine a bottle of ketchup, I have a single representation of a bottle of ketchup. (I may instead have numerous representations of the bottle of ketchup, but make it one representation for simplicity's sake.) The content of that representational state represents the bottle as a bottle of red stuff that is ketchup, and the representation itself is the *vehicle* of that content. But I don't just mean to defend internalism about the vehicles of consciousness and cognition; I mean also to defend internalism about whole minds, the systems that mental representations are states of.

Vehicle externalism and vehicle internalism are normally defined as follows. Vehicle internalists, it is said, are those who believe that minds supervene on brains, and that mental states (percepts, emotions, etc.) and mental events (perceiving X, feeling Y, etc.) supervene on neural states and events. Vehicle externalists, in contrast, may accept that minds center on brains in some sense, but don't accept that the supervenience basis of a mind necessarily *ends* at the borders of a brain. Minds can extend beyond brains, in other words, and even beyond organisms. So, to the

vehicle externalist, not every mental state or event is a neural state or event. For reasons that will emerge shortly, however, I think it is more helpful to describe vehicle internalists as believing that mental states supervene on states of *nervous systems*, and to describe vehicle externalists as believing that mental states can supervene on things outside of nervous systems as well.

I want to distinguish the debate about the supervenience base of the *vehicles* of consciousness and cognition from debates about the supervenience base of the *contents* of consciousness and cognition. Content externalists (Putnam (1975), Burge (1986, 1979), Dretske (1988), Tye (2000, 1995)) believe that the *contents* of mental states aren't or aren't always or entirely determined by what is inside the brain. So, the content of a subject's mental state, *M*, could change even if nothing inside that subject's brain changed, simply by changing something about the subject's world. Content internalists (Segal (2000), Botterill and Carruthers (1999)) instead believe that the contents of a subject's mental states are determined entirely by internal states of the subject. The debate about content externalism is not unrelated to the debate about vehicle externalism; some of the considerations that weigh in favor of vehicle internalism also weigh in favor of content internalism, for instance. But it is nonetheless fully possible to be a *content* externalist and a *vehicle* internalist. (Indeed, content externalism is probably a significantly more popular position among philosophers than is vehicle externalism.) And it is only the debate about *vehicle* externalism that is really relevant to my purposes, since I am interested in individuating the vehicles of contents (as well as whole minds), and not in individuating or characterizing their contents themselves. So, henceforth, if I refer

simply to internalism and externalism, I mean to refer to vehicle internalism and vehicle externalism.

The specific critics of internalism that I address myself to in this chapter are Hurley (1998), who advocates externalism in part in the context of discussing the split-brain phenomenon, and Clark and Chalmers (1998), who defend externalism (or as they call it the “extended mind” hypothesis) in a different context. Again, the defense mounted against these particular critics ties together a number of strands that have emerged over the course of this work.

In the next section I briefly explore Hurley’s implicit and explicit claims about the vehicles of *consciousness* in split-brain subjects and in acallosal subjects in particular. When we think carefully about what it would mean for behavior to serve as a mechanism of co-consciousness, however, it seems unlikely that behavior could play such a role. This suggests a more general diagnosis of one major difficulty with vehicle externalism, which is just that (at least most) non-neural events don’t seem capable of serving the same *fine-grained* functional role that some neural events play. This point is illustrated in Section Three in which I discuss Clark and Chalmers. The major distinction between vehicle externalists and vehicle internalists—when vehicle internalism is properly defined—isn’t that the latter draw an arbitrary, unprincipled distinction between what’s inside the skull and what’s outside of it. The major difference lies in the way they conceive of functional roles. For the externalists, apparently, these roles are not very fine-grained, whereas for the internalists, they are. A properly formulated vehicle internalism in fact seems to be uniquely compatible

with psychofunctionalism—or at least so an examination of scientific psychology at *this* point would suggest.

2 *What are the Vehicles of Consciousness?*

In Chapter Three I argued that neural facts constrain the individuation of psychological entities. I have agreed with Lockwood that mental structure is isomorphic to neural structure in at least one, albeit minimal and sort of obvious way. And in fact I have suggested more and less explicitly throughout this dissertation that neuroanatomical facts provide support for the mental and conscious duality models for split-brain subjects.

Yet all of these positions might seem to require, or to amount to, or to at least be importantly motivated by, some sort of mind-brain supervenience claim, or by a commitment to internalism about the vehicles of consciousness and cognition. Else, why think *neuroanatomy* in particular offers important constraints about something as functionally complex as co-consciousness? Why allow that apparent mental unity can be shown to be merely apparent in part on the basis of *neural* facts? Why even think that there is any distinction between intra-mind and inter-mind interactions that can be given in psychological terms—why think, that is, that there is some sort of hard *line* we can draw between events occurring within a mind and events occurring outside of it? The vehicle externalist raises all of these challenges.

Hurley's (1998) views on vehicle externalism in split-brain and in acallosal subjects were already described once in Chapter Three, in the context of a discussion of isomorphism between psychological and neuroanatomical structure. Hurley does

not think that split-brain subjects and acallosal subjects differ significantly at the level of neuroanatomy. As I noted at the time, this is actually not clear; split-brain subjects and acallosal subjects may differ significantly at the *relevant* neuroanatomical level, for “the same” neuroanatomical structures—structures that are the same in terms of their (gross) morphology, epigenetic origin, etc.—may subserve different functions in the two groups of subject. Their functional neuroanatomies may differ, in other words. (Though these functional differences may come with morphological differences as well, for example, a larger anterior commissure in acallosal subjects.)

Hurley’s more interesting point, though, is that even *if* we assume that acallosal subjects and split-brain subjects really don’t differ at the neuroanatomical level, this does not show that they don’t differ at the level of conscious structure. Even if acallosal and split-brain subjects behave similarly outside of the split-brain experiment, and even if their behavioral integration is achieved through the same means, acallosal subjects might still have a single stream of consciousness while split-brain subjects had two streams of consciousness.⁷²

Hurley’s reasons for thinking that the lack of a corpus callosum says more about the structure of a split-brain subject’s than an acallosal subject’s consciousness were given in Chapter Three and I won’t discuss them again here. I wish instead to consider the more general claim that cross-cuing and access movements *could* by themselves create a single stream of consciousness across two hemispheres that would otherwise each possess a unique stream consciousness. Hurley states that (at

⁷² Clearly, then, Hurley is not guilty of the sort of verificationism that *may* lurk in some of Marks’ (1980) and Tye’s (2003) arguments for mental and conscious singularity in split-brain subjects; Hurley does not confuse behavioral evidence concerning the structure of split-brain subjects’ consciousness with the actual structure of their consciousness, in other words.

least in acallosal subjects) “to the extent either external or internal mechanisms of integration function reliably, there is no reason not to regard them as part of the vehicles of co-conscious contents and of a unified consciousness” (1998: 191).

Access movements and cross cuing *transmit* information from one hemisphere to the other, thereby *carrying* contents (information), thereby constituting mental vehicles of content—mental representations.

In Chapter Four I argued that the global workspace theory of consciousness (Baars, 1997, 1988; Baars and Franklin, 2003; DeHaene and Naccache, 2001) can account for (access and phenomenal) *co-consciousness*. In this account, for two experiences to be co-conscious is just for them to be available to the same set of (token) global consumer systems. If we accept this account of co-consciousness, determining the structure of a split-brain *or* an acallosal *or* a “normal” subject’s consciousness is a matter of determining how many sets of global consumers the subject has, and which perceptual states they have access to.

From this perspective, Hurley’s willingness to go externalist about the vehicles of consciousness or at least co-consciousness, amounts to accepting that cross-cuing and access movements can provide the global consumer systems of *one* hemisphere access to experiences in the other hemisphere. In defense of this kind of claim, Hurley writes that:

in principal, an external mechanism of integration could be part of a causal system that supports the very unity of consciousness itself. There is nothing magical about this possibility: it appeals to a system of causes and effects in a perfectly naturalistic way, even though some causal paths go external. Indeed, the point depends on recognizing that there is no magic causal boundary around the brain that makes it impossible in principle for the vehicles of a unified consciousness to spread beyond it. . . we

can think of the subpersonal basis of a unified consciousness as a kind of *dynamic singularity* in the field of causal flows. . . Such a dynamic singularity is centered on the organism and moves around with it, but it does not have sharp boundaries. (2003: 81; original emphasis.)

Hurley twice asserts in this passage that vehicle externalism isn't committed to the existence of magic, which is correct—but of course I don't think vehicle internalism is committed to anything magic, either. The internalist could agree that "There is no reason *in principle*. . . why the unity of consciousness could not be supported by causal mechanisms that pass outside of the central nervous system, so long as they do so in a way that meets the relevant functional criteria" (1994, fn 14: 71; emphasis added). The vehicle internalist simply suspects that the vehicles of consciousness and of co-consciousness and of cognition generally *in fact do not* spread beyond the central nervous system (at least as we know it, in the creatures we know of it, so far).

That said, Hurley is also right to note that the relevant functional criteria—for co-consciousness, or for being a desire, or a memory, or anything else—"remain to be specified!" (ibid) So it is not possible to definitively rule out that cross-cuing could serve as a mechanism of co-consciousness. But my suspicion is that the relevant functional criteria for the mechanisms of co-consciousness will not be ones that cross-cuing and access movements will meet.

Certainly they will not if the account of co-consciousness that I have adopted is correct. According to the theory of co-consciousness I first described in Chapter Four, the mechanism of co-consciousness, that which makes E1 and E2 co-conscious, is simply the joint availability of multiple experiences to the same set of (token)

consumer systems. For cross-cuing and access movements to serve as the mechanism of co-consciousness, then, they would have to make *the same token experiences* available to the same set of global consumers. But access movements—moving the eyes to bring a previously LVF (RH) percept into the RVF (LH), for example—seem capable, at best, of making experiences with the same contents available to RH and LH global consumers. (And even here, recall that in Chapter Four I argued that the functional non-identity of the two hemispheres with regard to perceptual processing make likely inter-hemispheric differences in conscious contents.) And cross-cuing doesn't seem capable of doing even that. Cross-cuing is a means of exchanging information, by creating new experiences—but the new experiences are unlikely to have the same rich contents as the original experiences. For instance, S has two conscious representations of the lighter, one in his RH and one in his LH, and both representations may even be multimodal. (His right hemisphere has a tactile and a visual representation of the lighter; his left hemisphere has an auditory and a visual representation of the lighter, because he's constructed a mental image of a lighter.) But what are the chances the contents of these representations will be identical?

Of course, this account of co-consciousness may very well be wholly incorrect. Even if it is, however, I still suspect that, whatever the *correct* account of co-consciousness, cross-cuing and access movements will not be able to serve as mechanisms of it—even if they contribute to the contents of consciousness in split-brain subjects, as they clearly do.

An analogy to a non-mental instance of drawing such boundaries may be helpful. Imagine a person who regularly drinks ginger ale as a digestive aid. Why

doesn't the carbonation process at the ginger ale plant count as part of this person's digestive process? After all, the carbonated beverage really is providing some important contribution to his digestive process as a whole. So why aren't the steps involved in preparing the carbonated beverage a genuine part of this process? Of course, no one would think that they were, and there is good reason not to think so, which is that none of the explanatory generalizations and distinctions and so forth that make sense of digestion *apply* to whatever goes on at the ginger ale manufacturing plant.

It is, again, *possible* that the generalizations and distinctions and so forth that apply to the generation of co-consciousness within a brain, and that help us make sense of the generation of co-consciousness within a brain, will also apply to cross-cuing and access movements in split-brain subjects; that is, it is not possible to rule out this possibility. It simply seems unlikely that the developed laws of an adequate psychological theory that account for things such as perceptual binding in the brain, will somehow apply also to smiles and frowns and turns of the head. The internalist needn't be committed to some sort of "magic" boundary at the boundary of the brain. The internalist simply believes that there is an explanatorily interesting boundary to be drawn at the brain.

It may be worth noting in passing that while brains aren't magic, they are somewhat impressive. (The brain is impressive in part because it *isn't* magic; how does something whose basic constituents and activities are so simple nonetheless do such complex things?) Neural stuff allows for interactions and entities whose richness and subtlety we probably don't even fully appreciate yet, and certainly don't

understand, and can't yet explain. This again isn't to say that things besides the brain couldn't realize equally rich phenomenal experiences, or couldn't realize equally subtle forms of integrating cognitive and perceptual information. It's just to say that the bar is set pretty high—higher, I submit, than vehicle externalists tend to set it.

3 *Functional Roles and Fineness of Grain*

The most important point of contention between internalists and externalists may be how fine-grained our functional characterizations of mental states should or will ultimately be. Hurley is right and correct to remind us not just that what matters in psychological individuation is functional role, but also that the functional roles of conscious states and of co-consciousness have not yet been specified. I think that the internalist and the externalist have different beliefs about how those and other functional roles should and ultimately will be specified, however.

Let me explain this by talking about vehicle externalism more generally, especially since I have already explained why I think that split-brain subjects have two sets of global consumers and two streams of consciousness. I don't deny the possibility of non-neural mental states and events. Indeed I can imagine hypothetical scenarios in which such things are actually plausible—cases involving digital memory storage devices, for instance, that have not yet been invented. Admittedly in these scenarios, however, the “external” device I am imagining could itself be called an electronic or robotic or digital part of the nervous system. For, whatever it was made out of, the “external” device would be so deeply integrated with the rest of the nervous system, in its functioning, that it would constitute a genuine functional part

of it. So I think that the disagreement between internalists and externalists is better captured by talk of nervous systems, where these are defined largely functionally, than by talk of brains, where these are defined largely physically/anatomically.

I have just suggested that the conditions for being a nervous system are at least largely functional. They won't be entirely functional of course, but at least partly physical, at the very least in the minimal sense that something must have some physical properties in order to be a nervous system, both for standard physicalist reasons, and for reasons laid out in the arguments in Chapter Three that the constitutive conditions for mental tokens are partly neural. (I won't work out the conditions for being a nervous system here, but note that one plausible condition is the very rapid transfer of information within a nervous system—probably much more rapid than communication between other parts of the organism's body, and certainly faster, of course, than the organism's behavior.) Now perhaps it is a further condition for something's being a nervous system that it is made (at least in part) out of biological stuff. (See Rey, 1997, pp. 191-194, for a brief discussion of "anchored functionalism" and "physiofunctionalism." Rey, however, is discussing the constitutive conditions for being a *mind*, and the possibility that some properties of biological matter are essential for consciousness and even cognition.) And even if some *nervous systems* can be non-biological, it could turn out that biological matter is essential for consciousness and even cognition. I simply don't mean to *rule out* now, a priori as it were, the possibility of non-biological nervous systems and indeed non-biological minds.

But now, surely, we must revisit the claim that the constitutive conditions for mental tokens are partly *neural*. This claim, defended in Chapter Three, has played a significant role in my arguments for mental and conscious duality in split-brain subjects. Yet I have just allowed that minds and indeed nervous systems may be constituted by something other than (biological) neurons! But there is no contradiction here; to resolve the apparent contradiction one need merely understand *neurons* as, at a first pass, the physical constituents of nervous systems. In split-brain subjects, these constituents are neurons; in some possible creatures these constituents might be electronic switches, or something. All the arguments for the constitutive criteria for minds and other mental tokens being partly neural—arguments concerning realism and physicalism and the causal nature of the mental—still go through even if we understand “neural” in this broader way.

Whatever the physical basis of a creature’s nervous system, some non-functional, physical properties of it are part of the constitutive conditions for that creature’s mental tokens. Or so I claim. This claim is what is at issue between the internalists and the externalists. Vehicle internalists believe that states of minds supervene on states of nervous systems; vehicle externalists believe that states of minds can supervene on things beyond the nervous system as well.

Now internalists may believe that, at this point in time, nervous systems are constituted entirely of neurons, and many may believe that mental states supervene on certain *parts* of existing biological nervous systems in particular, i.e., on brains (or on cerebral cortices, or whatever). But the internalist could allow that this is just a contingent fact about the state of our medical technology. Properly understood, then,

vehicle internalism isn't committed to it being metaphysically impossible for minds to supervene on anything but brains. This isn't the source of disagreement between internalists and externalists. Rather what is at issue seems to be some sort of nomological possibility: internalists believe that the laws (or models or whatever) of psychology fail to apply to the sorts of non-neural states and events that externalists must claim at least some of them *do* apply to. Internalists deny that the *particular* cases externalists believe show that minds don't entirely supervene on brains (or nervous systems) really do show this.

Internalists deny, for instance, that the case of Otto and his notebook (Clark and Chalmers, 1998) constitutes an exception to mind-brain supervenience. Otto is a person with Alzheimer's who functions adaptively in his daily life due to steadfast reliance upon a particular notebook that he carries with him at all times. In this notebook Otto writes down every new bit of information he learns that he thinks he might need in the future. When he wants to recall where the MoMA is located, or indeed where his own apartment is located, or perhaps even what his own name is, he looks it up in the notebook. Clark and Chalmers, defenders of vehicle externalism, believe that Otto's notebook entries qualify as mental entities: they constitute (at least some of) Otto's dispositional, non-occurrent beliefs.

In a moment I will explain why Otto's notebook entries look as if they will *not* fall under the laws of a developed psychological theory. Before offering this objection to the Otto case in particular, however, it is worth emphasizing the crucial role that the Otto case plays in Clark and Chalmers' arguments for vehicle externalism (or for the "extended mind hypothesis" as they call it). Otto's case isn't

introduced until halfway into their article, after several other mundane and less mundane scenarios meant to test the *intuition* that the mind does *not* extend beyond the brain. One such mundane scenario is the use of pencil and paper to calculate a long division problem; another is the use of the rotate function on a controller to determine whether a particular shape in the game Tetris will fit in a particular slot. Clark and Chalmers argue that pressing the rotate button or writing numerals aren't external to some cognitive process but are rather *components* of cognitive processes. So stated, however, the argument isn't necessarily counter-intuitive; depending on how broadly or how loosely we understand "cognitive process," or what it is to participate in a cognitive process, all sorts of non-mental entities may participate in such processes. (It isn't as if, for example, psychological accounts of people's behavior and of the underlying mental causes of that behavior can never refer to non-mental things to do explanatory work. Explanations of addiction might refer to drugs or to money without implying that drugs and money are mental entities.)

Clark and Chalmers seem to understand this; they acknowledge that readers looking at scenarios depicting cognitive *processes* that involve entities located outside the brain, might shrug: "Perhaps some *processing* takes place in the environment, but what of *mind*?" (1998: 12; original emphasis) After all, "Everything. . . said so far is compatible with the view that truly mental states—experiences, beliefs, desires, emotions, and so on—are all determined by states of the brain. Perhaps what is truly mental is internal, after all?" (ibid)

Otto is their attempt to "take things a step further", by showing that mental states, "*beliefs* [themselves] can be constituted partly by features of the environment,

when those features play the right sort of role in driving cognitive processes. If so, the mind extends into the world” (ibid; original emphasis). Otto is arguably *the* crucial case for Clark and Chalmers’ defense of vehicle externalism, in other words.

And they use Otto’s case to argue that vehicle internalists draw an unprincipled distinction between entities and events inside the brain and those outside it, a distinction that the best functionalist psychological theories will prove irrelevant. Otto’s notebook entries play the same functional role for Otto as ordinary long-term (semantic) memories or beliefs play for ordinary human subjects. So Clark and Chalmers begin by describing a comparison case, a “normal case” involving a “belief embedded in memory” (1998: 12): Inga decides to see an exhibit at the MoMA today, “thinks for a moment and recalls that the museum is on 53rd Street, so she walks to 53rd Street and goes into the museum.” Clark and Chalmers note that surely Inga *believed* the museum was on 53rd Street *before* she began thinking about visiting it today, “before she consulted her memory. It was not previously an *occurrent* belief, but then neither are most of our beliefs. The belief was somewhere in memory, waiting to be accessed” (ibid; original emphasis).

They next contrast Inga with Otto:

When he [Otto] learns new information, he writes it down. When he needs some old information, he looks it up. For Otto, his notebook plays the role usually played by biological memory. Today, Otto hears about the exhibition at the Museum of Modern Art, and decides to go see it. He consults his notebook, which says that the museum is on 53rd Street, so he walks to 53rd Street and goes into the museum.

(Clark and Chalmers, 1998: 12-13)

Just as Inga surely believed that the museum was on 53rd Street even before she consulted her memory, surely Otto believed that the museum was on 53rd Street even before he consulted his notebook, Clark and Chalmers reason:

For in *relevant* respects the cases are entirely analogous: the notebook plays for Otto the same role that memory plays for Inga. The information in the notebook functions just like the information constituting an ordinary non-occurrent belief; it just happens that this information lies beyond the skin.

(1998: 13; emphasis added)

Both Otto and Inga must “consult” their “memory”; Inga’s is just located in her skull, and Otto’s (in part) in his notebook. Otto’s notebook entry interacts (albeit not directly) with Otto’s desire to see the new MoMA exhibit, just as Inga’s memory interacts with Inga’s desire to see the new MoMA exhibit; Otto’s notebook entry allows him to find the MoMA, just as Inga’s memory allows her to find the MoMA. So, for Clark and Chalmers, the functional role is the same: both Otto’s notebook entries and Inga’s memories meet the constitutive criteria for memories. All that differs is the supervenience base.

Now perhaps Otto’s notebook entries play the same role for Otto, in some sense, that Inga’s memories do for Inga. This seems unlikely, but I will assume that this is the case for the sake of argument. Nonetheless, even if Otto’s notebook plays the same role *for him* that Inga’s memories do *for her*, the relevant question for the realist is what role long-term (semantic) memory, or dispositional belief, will play in a psychological theory. Will an adequate psychological theory leave unanalyzed the phrase, “consult one’s memory”, for instance? If so, then perhaps that theory would view Otto’s notebook as part of his semantic memory, which he consults in the same

contexts in which Inga consults hers. For even our *current* (developing and still inadequate) psychological theories don't leave "memory consultation" unanalyzed. There are many kinds of memory, and many (still poorly understood) means of accessing one's memory, and cognitive psychologists are interested in developing separate accounts of all of them.

And quite obviously Inga accesses her memory of the MoMA's address wholly differently from how Otto accesses his "memory" of its address. If this isn't quite obvious in the passages above, it's because in their description of Otto and his notebook Clark and Chalmers (deliberately, it would seem) skate over issues any memory researcher would ask immediately about any putative memory system—questions like, "How is information *organized* in this memory system? In what format is that information is encoded? How does one *locate* a particular entry within the system?" For once we do ask those questions, it is perfectly plain that the theoretical characterization of Inga's memory and Otto's notebook will be wildly different. Are Otto's notebook entries organized alphabetically by the name of the person or place or object he's looking for information on? ("Andy Abrams"; "Betsy Johnson Boutique, San Diego"; "Car Wash, Local".) Alphabetically by category? ("Markets"; "Mosques"; "Museums".) Alphabetically by information type? ("Addresses"; "Fashion Tips"; "Recipes".) Will entries be organized by the frequency with which he will need to consult them? (Not a bad system, but it will require some careful thought to get it up and running, since he won't always know in advance just how often he'll need to consult an entry, and since he won't be able to rely on his *memory* of how frequently a particular entry has been used in the past.) Will entries

be organized by the date on which the information was acquired? (Hopefully not—he might have to read through the entire notebook to search for any piece of information, assuming he can't remember when he acquired it; parts of *Inga's* memory, however, may be organized by date.) Do Otto's notebook consultations reveal "state-dependency"; does he have a tendency to find happy listings (birthdays, lists of books he's enjoyed) when he's in a good mood, and sad entries (reminders of upcoming dental appointments, lists of terrible movies he definitely does not want to rent and watch a second time by accident) when he's feeling down?

There is of course an alternative to accepting that Otto's notebook entries themselves constitute non-occurrent beliefs; as Clarke and Chalmers acknowledge:

The alternative is to explain Otto's action in terms of his occurrent desire to go to the museum, his standing belief that the Museum is on the location written in the notebook, and the accessible fact that the notebook says the Museum is on 53rd Street.

(1998: 13)

But, they insist, "this complicates the explanation unnecessarily" (*ibid*):

If we must resort to explaining Otto's action this way, then we must also do so for the countless other actions in which his notebook is involved; in each of the explanations, there will be an extra term involving the notebook. We submit that to explain things this way is to take *one step too many*. It is pointlessly complex, in the same way that it would be pointlessly complex to explain *Inga's* actions in terms of beliefs about her memory. (1998: 13; original emphasis)

This is question-begging, and I think somewhat confused. It is question-begging because we shouldn't say anything about Otto that it would be "pointlessly complex" to say of *Inga only if* Otto's notebook entries really *are* the equivalent of

Inga's memories. It is somewhat confused because it wouldn't be *pointlessly complex* to explain Inga's successful navigation to the museum in terms of her occurrent desire to go to the museum, her standing belief that the museum is at the location specified in her memory, and the accessibility of her memory that the museum is on 53rd Street (or, the accessible fact that her memory says the museum is on 53rd Street (!)). It would be outright *misleading* to say this. For Inga might not *have* any beliefs about her memory whatsoever; maybe she's skeptical, on her walk down to 53rd, that her memory can be trusted; maybe Inga somehow doesn't even know what memory is. For memories are no doubt often accessed without first forming or accessing any beliefs about memories. (Animals appear to access long-term memories, presumably without possessing beliefs about their memory systems.) Inga, in other words, need have no beliefs at all about her memory system in order to retrieve from memory the MoMA's address. But Otto's beliefs about his notebook are plainly essential to his ability to access and utilize its entries; we could easily manipulate his ability to access and utilize those entries by altering his beliefs about, e.g., his notebook's current location.

It is significant, I think, that, when attempting to find a plausible example of an external mental state, the *type* of mental state Clark and Chalmers come up with is the dispositional belief. It is interesting that it is *this* role that Otto's notebook entries are claimed to meet, for non-occurrent beliefs are not in and of themselves causally efficacious. While I can't offer much in the way of an analysis of this problem, much less a solution to it, I think that what the Otto case illustrates is a certain puzzle (or perhaps a host of puzzles) about the role of merely dispositional, non-occurrent

beliefs in psychology, rather than illustrating the inadequacy of vehicle internalism.⁷³ While Inga's dispositional belief that the MoMA is located on 53rd Street is just as easily described as a semantic memory that the MoMA is so located, there is something puzzling about describing Otto's notebook entries as Otto's semantic memories. Otto doesn't have semantic memories. (Though apparently his procedural memory is intact, since he uses his notebook so well.) That's why he, but not Inga, *needs* the notebook. Of course, Clark and Chalmers would presumably accuse me of begging the question by supposing that Otto's memories aren't just differently realized than Inga's, by supposing that Otto is cognitively impaired in the first place! But this seems an odd response on at least Clark's part, for he has previously and persuasively argued (1997) that we citizens of the modern technological (and even the not-so-modern, pre-technological) age "off-load" so much information into our environments (books, street signs, personal digital assistants), and up-load information we deliberately store in our environments (road maps, telephone books, customer service representatives), precisely *because* there are so many limits to our memory, precisely to *lighten* our own mnemonic burdens.

Meanwhile, to actually cause something, a belief (or any stored memory) would seem to have to be occurrent. How does an entry in Otto's notebook—say, "The MoMA is located on West on 53rd Street between Fifth and Sixth Avenue—become an occurrent belief? Well, to begin with, very differently from the way in which Inga's non-occurrent belief becomes an occurrent belief. Among other things,

⁷³ While the motivation for and usefulness of referring to non-occurrent, dispositional or standing beliefs is clear, there are also certain problems with such explanations. I suspect that our understanding of dispositional beliefs will undergo serious transformation as psychology develops, but this is just a hunch.

Otto has to *read* his non-occurrent belief—literally, physically read it. (Let’s hope that, on top of Alzheimer’s, he doesn’t also have glaucoma, or arthritis in his hands that makes flipping pages difficult.)

And, of course, the representational format carrying the fact the entry represents (i.e., the address of the MoMA), has to undergo a radical transformation. His entries are linguistic entities, sentences in natural language. His semantic memories are stored perhaps in a language of thought. (The very *handwriting* his entries are written in may determine their causal role; if Otto were the paranoid sort, he might distrust any notebook entry written in boxy print as opposed to his own loopy cursive.) Finally, his notebook entries, though they are representations, need themselves to be re-represented in order to become causally efficacious, whereas this is presumably not true of Inga’s dispositional belief or semantic memory representing the MoMA’s being located on 53rd Street. To explain why Otto crosses Central Park on his way downtown, Otto needs to generate a *genuine* mental representation, one instantiated in his brain, of the MoMA’s being located on 53rd Street. In other words, we end up needing to go back to good old internal physical states after all.⁷⁴

4 Conclusion

Cognitive psychology looks for *stable patterns* of functional interactions in constructing models of cognition; it seeks out causes of behavior that are relatively stable, in order to have as much predictive and explanatory power as possible. Clark

⁷⁴ My point in this paragraph is that identifying entries in a notebook with *non-occurrent mental states* constitutes a kind of hedge, a significant qualification of the externalist position. Bartlett (2008) points out that Clark and Chalmers in fact hedge in two further ways as well: they don’t say that Otto’s notebook *contains* some of his beliefs, only that his notebook entries *partly constitute* some of his beliefs. This hedges against the claim that any of Otto’s beliefs exist *entirely* outside his body, and also against the claim that any of Otto’s beliefs are *located* outside his body.

and Chalmers argue that if Otto's reliance on his notebook becomes sufficiently steadfast—if he and his notebook form a sufficiently tightly “coupled” system—then he and his notebook become *a* mental system. So they believe that it would be pointlessly complex to refer to Otto's notebook each and every time we want to explain his successful navigation in the world because “Otto's book is a constant for Otto, in the same way that memory is a constant for Inga; to point to it in every belief/desire explanation would be redundant. *In an explanation, simplicity is power*” (1998: 13-14; emphasis added). Similarly, Hurley believes that the stable causal system that causes and explains a subject's unified behavior may include access movements and cross-cuing, even if such external mechanisms are not part of the stable causal system that causes and explain *all* subjects' unified behavior.

Part of *my* point in response to this is that cognitive psychology doesn't just look for patterns of functional interaction that are stable *within* an individual, or a small group of subjects. It looks for patterns that are stable *across* individuals as well. We aren't just looking for a theory of Otto psychology, for instance: we are looking for a theory of human psychology. And the counterfactuals and generalizations that apply to human memory will not apply to Otto's notebook.

Defenders of vehicle externalism might object that they *also* believe that cognitive psychology looks for patterns of functional interaction that exist across individuals. After all, Clark and Chalmers argue that Inga and Otto both fall under the same patterns of functional interaction, and this is why Otto's notebook entries are dispositional beliefs just as Inga's own semantic memories are. Indeed, I take it that Clark and Chalmers believe that they are looking deeply, just as theoretical

psychologists (and psychofunctionalists) should do; they even claim that “differences between Otto’s and Inga’s cases are all *shallow* differences” (1998: 16; original emphasis). They seem to say, that is, “Of course, on the surface, Otto’s notebook entries and Inga’s long-term memories are differently *realized*, which means there will be *some* small differences in how they are accessed so forth. But look *deeply* and you will see that *central* functional roles of these two different physical entities are the same.”

But this is only plausible if our functional roles are very coarse-grained. For the psychofunctionalist, who believes that the best account of mental states will come from an analysis by ramification of a developed, adequate scientific psychological theory, functional roles are unlikely to be so coarse-grained. For the history of theoretical work even in cognitive psychology so far is one of drawing ever finer functional distinctions: between procedural and semantic and episodic and working and iconic and echoic memory systems, between identification-oriented and action-oriented visual processing streams, between acoustic and phonetic and syntactic and semantic and pragmatic components of language comprehension, between various kinds of somatosensory information—pain versus temperature versus proprioception versus deep touch versus light touch. The patterns of functional interaction that characterize the activities referred to in psychological theories are subtle, and the categories of entity that these give rise to may be very fine-grained—or at least more fine-grained than vehicle externalists appear disposed to recognize.

Carruthers (in conversation) has pointed out that there is nothing *psychologically* interesting in the pencil marks on a piece of paper that were created

in performing an act of long multiplication. Of course, the marks made could reveal to us some psychologically interesting things, like that this four-year old is capable of long multiplication, or that this accountant is suffering from some kind of cognitive impairment; but the pencil marks are not psychologically interesting in and of themselves. It is of course psychologically interesting that most calculus students can do long multiplication in their heads only painfully (and unreliably), while they can do the same problems quite easily provided that pen and paper are allowed; this tells us something about working memory capacity, for instance. But nonetheless all the entities and activities that are psychologically interesting in and of themselves are *internal* to the agent: in the *percepts* of the four and the seven, in the accessing of the memory that four times seven is twenty-eight, in the procedural memory used to nearly automatically write the “8” at the bottom of the rightmost column and write the (“carried”) “2” at the top of the second column from the right. But the *actual* marking of the shape 2 on the paper? That’s just physics.⁷⁵

There are *some* functional similarities, of course, between Otto’s notebook entries and Inga’s dispositional beliefs or stored semantic memories. And no doubt there are some functional similarities between inter-hemispheric information exchange via cross-cuing, and standard cases of intra-hemispheric information

⁷⁵ Similarly, not all of the *internal* events that partly constitute the entire process of solving the problem are psychologically interesting. It is itself an interesting question of scientific psychology and of philosophy of psychology at what *point*, *internal* to the agent, the in-and-of-themselves psychologically interesting things start happening. But it is unlikely that *all* of the entities and events necessary to moving the hand to calculate the long multiplication problem are psychologically interesting, even if they are interesting from some neuroscientific perspectives, e.g. from that of electrophysiology. A single skeletal muscle fiber, for instance, is usually innervated by just a single motor neuron; when an action potential travels down a motor neuron axon, a very large postsynaptic potential results, and this one synaptic potential is generally by itself sufficiently large to trigger an action potential in the muscle fiber. There is no need, to explain this final necessary cause of motor behavior, to refer either to representation or to computation. Physics and chemistry will take it at least from here.

exchange. But there are also enough genuine functional differences between the entities and activities realized by neural stuff and the entities and activities *not* so realized, at least at this point in time, to make the vehicle internalist's way of drawing the border around the mental genuinely principled. External things can carry information, of course; telephone directories and road maps and Stars of David all represent, and calculators and PDAs process representations, and all of these things may play roles in cognitive processes. But these external vehicles of information cannot yet do the things that internal vehicles can do. This is why the borders of the mind can still be located within the borders of the brain. And this is why *neural* facts offer special constraints on the individuation of mental entities—constraints that facts about notebooks and even body parts like fingers and mouths do not offer.

The externalists do not need to make the case that we should look deeply in order to construct the best accounts of the nature of beliefs and memories, experiences and streams of consciousness. They need to make the case that they are looking deeply enough. “In an explanation, simplicity is power,” Clark and Chalmers say—but their explanation ultimately complicates. Consider their response to the objection that Otto's notebook entries, unlike Inga's long-term memories, need to be *perceived* in order to play a role in Otto's adaptive daily behavior: they suggest that “to put things this way is to beg the question,” for if their arguments are correct:

Otto's internal processes and his notebook constitute a single cognitive system. *From the standpoint of this system*, the flow of information between notebook and brain is not perceptual at all; it does not involve the impact of something outside the brain. (1998: 16; emphasis added)

But again, what about the standpoint of psychological *theory*? Must a developed psychological theory now have to say that people can sometimes access the representational content of sentences written in natural language without engaging in any sort of reading (or listening)? Or will it have to say that reading is sometimes not at all a perceptual process? Either way we now need additional psychological accounts of how this is possible. This *adds* a burden onto our psychological theories. Explaining each and every one of Otto's notebook-guided behavior by referencing his beliefs about his notebook does not complicate our psychological theorizing or theories. Denying that Otto's notebook entries are memories may make *giving* explanation after explanation of Otto's individual behaviors slightly more tedious. But the account of memory thus permitted is both simpler and deeper. In fact to the extent that it is simpler, the power of its simplicity arguably comes in part from its depth—from its deeper and more careful characterization of memory processes.

Conclusion

The split-brain studies are fascinating from numerous perspectives: they raise questions about our own self-knowledge, about the epistemic limits of phenomenology and introspection, about the sources of our agency and our sense of free will, about the unity of our mental lives, about personhood, and about the very nature of the human self. Yet this dissertation has been focused much more narrowly on another, humbler, question: what is the proper way to individuate mental tokens in split-brain subjects?

I have argued that split-brain subjects have two minds—and, most likely, two streams of consciousness—albeit minds bearing an unusually intimate relationship with each other, because they belong to one subject. On the assumption that each hemisphere of a split-brain subject (perhaps in conjunction with non-cortical structures) constitutes a competent mind, and given the fact that that the two hemispheres are hooked up to the same body, the integrated-seeming and adaptive nature of split-brain subjects' behavior is not surprising. Two systems can produce integrated behavior without their own mental processes being integrated with each other.

Although the two hemispheres of a split-brain subject interact with each other in various ways, these interactions occur at least in most cases in part via the mediation of behavioral and other non-mental events, rather than via the kinds of interactions that seem to characterize the inner workings of a mind. The results of the split-brain experiments meanwhile suggest that the two hemispheres of a split-brain

subject are clearly not co-conscious with each other, at least with respect to most of their conscious states. Certain conscious information does appear to transfer from one hemisphere to the other— affective information for instance—but that it does transfer still does not show that the hemispheres share any mental tokens, or even, again, that the transfer occurs via a wholly mental process.

As I have acknowledged throughout, however, these conclusions, regarding the structure of consciousness and cognition in split-brain subjects, are necessarily tentative. There is far too much we still do not understand about minds, brains, and consciousness, to have real confidence in our conclusions about this hard case.

Perhaps more important than these specific conclusions, however, is the approach I have tried to take in reaching them. While the split-brain phenomenon is probably a source of insight for many non-scientific inquiries and in many important domains, I have tried to show that *how many minds* split-brain subjects have—and how many streams of consciousness they have, and so forth—are questions that may well ultimately be answerable from a theoretical (i.e. theoretical-empirical) perspective. For even the concept of a mind appears to do some important, albeit perhaps implicit, work in functional, psychological explanation.

As a psychofunctionalist I believe that the constitutive conditions for mental phenomena will come from the ramsification of a developed psychological theory. Functionalism in its various forms is most fundamentally a theory of mental *kinds* or types. It offers a general strategy for discovering the constitutive conditions for various *types* of mental states—a strategy for determining what it is to be *a belief* (any belief) as opposed to *a desire*, or as opposed to being something with no mental

properties at all. But in this work I assume that the constitutive conditions for being a mental token—one belief rather than another, one mind rather than two minds—will also come from a developed psychological theory. And although such a theory is something that we again do not yet possess, this work at least begins to explore some of the constraints that a developed psychological theory is likely to impose on the individuation of mental entities.

In particular I have argued throughout that neural facts constrain the individuation of mental entities. As I first acknowledged in Chapter Two, it is admittedly difficult to say at this point at what level of abstraction away from the physical a developed psychological theory, and thus a psychofunctionalist approach to mental phenomena, will be pitched. But psychology is clearly interested in giving a *causal story* of the mind. And psychofunctionalists who are realists, who believe that there really are things occupying the functional roles implicitly defined by a psychological theory, are similarly committed to the existence of causally efficacious mental tokens, and similarly committed to the causal reality of these tokens. Therefore their functional roles are genuine causal roles. But causal role isn't just necessary to type-identifying mental phenomena: it's necessary to token-identifying such phenomena, to individuating them, as well. The causal properties of a neural token constrain the mental tokens with which we can identify it.

Not all of a neural token's causal or physical properties need matter to its token identity, just as they needn't all matter to its type identity. But some of them do. One thing that *is* relevant to the type identity of a mental token is the kind of causal relations it bears to other instantiated mental types. And one thing that's relevant to

the token identity of a mental token is the kind of causal relations it bears to other particular instantiations of mental types. Thus the *causal distance* between multiple neural events has implications for their individuation as mental tokens. Indeed I have tried to provide some reason to think that that there is a psychologically important difference between the causal processes that occur within each hemisphere of a split-brain subject and the ways in which the hemispheres interact with each other. If these conclusions are correct, then the hemispheres may communicate, and even cooperate, with each other—but this communication and this cooperation occur *between minds*.

Bibliography

- Aboitiz, F. 2003. One hundred million years of interhemispheric communication: the history of the corpus callosum. *Brazilian Journal of Medical and Biological Research* 36: 409-420. Associacao Brasileira de Divulgacao Cientifica.
- Adamec, R., and Morgan, H. 1994. The effect of kindling of different nuclei in the left and right amygdala on anxiety in the rat. *Physiology and Behavior* 55: 1-12.
- Arguin, M., Lassonde, M., Quattrini, A., Del Pesce, M., Foschi, N., and Papo, I. 2000. Divided visuo-spatial attention systems with total and anterior callosotomy. *Neuropsychologia* 38: 283-291.
- Armstrong, D. 1968. *A Materialist Theory of the Mind*. Routledge and Kegan Paul.
- Astafiev, S., Shulman, G., Stanley, C., Snyder, A., Van Essen D., and Corbetta, M. 2003. Functional organization of the human intraparietal and frontal cortex for attending, looking, and pointing. *Journal of Neuroscience* 23: 4689-4699.
- Baars, B. 1997. *In the Theater of Consciousness*. Oxford University Press.
- Baars, B. 1988. *A Cognitive Theory of Consciousness*. Cambridge University Press.
- Baars, B., and Franklin, S. 2003. How conscious experience and working memory interact. *Trends in Cognitive Sciences* 7: 166-172.
- Baddeley, A. 2000. The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences* 4: 417-423.
- Baddeley, A., and Hitch, G. 1974. Working memory. In Bower, G. (Ed.), *The Psychology of Learning and Motivation*, pp. 47-89. Academic Press.
- Baddeley, A., and Logie, R. 1999. Working memory: the multiple-component model. In A. Miyake and P. Shah (Eds.), *Models of Working Memory*, pp. 28-61. Cambridge University Press.
- Banich, M. 1998. Integration of information between the cerebral hemispheres. *Current Directions in Psychological Science* 7: 32-37.
- Bartlett, G. 2008. Whither internalism? How internalists should respond to the extended mind hypothesis. *Metaphilosophy* 29: 163-184.
- Bayne, T. 2005. Divided brains & unified phenomenology: An essay on Michael Tye's "Consciousness and Persons". *Philosophical Psychology* 18: 495-512.

- Bayne, T., and Chalmers, D. 2003. What is the unity of consciousness? In A. Cleeremans (Ed.), *The Unity of Consciousness: Binding, Integration, and Dissociation*, pp. 23-58. Oxford University Press.
- Bayne, T. 2001. Co-consciousness: Review of Barry Dainton's 'Stream of Consciousness'. *Journal of Consciousness Studies* 8: 79-92.
- Beeman, M. and Chiarello, C. 1998. Complementary right- and left-hemisphere language comprehension. *Current Directions in Psychological Science* 7: 763-784.
- Beeman, M., Friedman, R., Grafman, J., Perez, E., Diamond, S., and Lindsay, M. 1994. Summation priming and coarse coding in the right hemisphere. *Journal of Cognitive Neuroscience* 6: 6-45.
- Bennett, M., and Hacker, P. 2005. Emotion and cortical-subcortical function: conceptual developments. *Progress in Neurobiology* 75: 29-52.
- Berlucchi, G. and A. Antonini. 1990. The role of the corpus callosum in the visual representation of the visual field in cortical areas. In C. Trevarthen (Ed.), *Brain Circuits and Functions of the Mind*, pp. 129-139. Cambridge University Press.
- Berridge, K. 2003. Comparing the emotional brains of humans and other animals. In R. Davidson, K. Scherer, and H. Goldsmith (Eds.), *Handbook of Affective Sciences*, pp. 25-51. Oxford University Press.
- Block, N. 2007. Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behavioral and Brain Sciences* 30: 481-548.
- Block, N. 1995. On a confusion about a function of consciousness. *Brain and Behavioral Sciences* 18: 227-247.
- Block, N. 1986. Advertisement for a semantics for psychology. In P. French, T. Uehling and H. Wettstein (Eds.). *Midwest Studies in Philosophy* 10: 615-678.
- Boatman, D., Freeman, J., Vining, E., Pulsifer, M., Miglioretti, D., Minahan, R., Carson, B., Brandt, J., and McKhann, G. 1999. Language recovery after left hemispherectomy in children with late-onset seizures. *Annals of Neurology* 46: 579-586.
- Bogen, J. 1997. Does Cognition in the Disconnected Right Hemisphere Require Right Hemisphere Possession of Language? *Brain and Language* 57: 12-21.

- Bogen, J. 1998. Physiological consequences of complete or partial commissural section. In M. Apuzzo (Ed.), *Surgery of the Third Ventricle*, pp. 167-186. Williams & Wilkins.
- Bogen, J. 1990. Partial hemispheric independence with the neocommissures intact. In C. Trevarthen (Ed.), *Brain Circuits and Functions of the Mind*, pp. 215-230. Cambridge University Press.
- Bogen, J. 1985. Split-brain syndromes. In J. A. M. Frederiks (Ed.), *Handbook of Clinical Neurology, Vol. 1, Clinical Neuropsychology*, pp. 99-105. Elsevier.
- Bogen, J. 1969. The other side of the brain II: an appositional mind. *Bulletin of the Los Angeles Neurological Societies* 34: 135-162.
- Bogen, J., and Vogel, P. 1965. Cerebral commissurotomy in man. *Bulletin of the Los Angeles Neurological Society* 27: 169-172.
- Botterill, G., and Carruthers, P. 1999. *The Philosophy of Psychology*. Cambridge University Press.
- Braun, C., Achim, A., Larocque, C. 2003. The evolution of the concept of interhemispheric relay time. In E. Zaidel and M. Iacoboni (Eds.), *The Parallel Brain*, pp. 237-258. MIT Press.
- Brownell, H., Griffen, R., Winner, E., Friendman, O., and Happe, F.. 2000. Cerebral lateralization and theory of mind. In S. Baron-Cohen, H. Tager-Flusberg, and D. Cohen (Eds.), *Understanding Other Minds*: 306-333. Oxford University Press.
- Burge, T. 1986. Individualism and psychology. *Philosophical Review* 95: 3-45.
- Burge, T. 1979. Individualism and the mental. *Midwest Studies in Philosophy* 4: 73-121.
- Burkland, C. and Smith, A. 1977. Language and the cerebral hemispheres. *Neurology* 27: 627-633.
- Butler, S. and Norrsell, U. 1968. Vocalization possibly initiated by the minor hemisphere. *Nature* 220: 793-794.
- Butter, C., and Snyder, D. 1972. Alterations in aversive and aggressive behaviours following orbitofrontal lesions in rhesus monkeys. *Acta Neurobiologiae Experimentalis* 32: 525-565.
- Butterworth, G. 1995. An ecological perspective on the origins of self. In J. Bermúdez, A. Marcel, and N. Eilan (Eds.) *The Body and the Self*, pp. 87-106. MIT Press.

- Carlson, J., Godlman, R., Bollinger, J., and Wiedle, K. 1974. Der effect von problemverbalisation bei verschiedenen aufgabengruppen und darbietungsformen des 'Raven Progressive Matrices Test'. *Diagnostica* 20: 133-141.
- Carruthers, P. Forthcoming. How we know our own minds: the relationship between mindreading and metacognition. *The Behavioral and Brain Sciences* 32.
- Carruthers, P. Forthcoming.. An architecture for dual reasoning. In J. Evans and K. Frankish (Eds.), *In Two Minds: Dual Processes and Beyond*. Oxford University Press.
- Carruthers, P. 2006. The case for massively modular models of mind. In R. Stainton (ed.), *Contemporary Debates in Cognitive Science*, pp. 3-21. Blackwell.
- Carruthers, P. 2005a *The Architecture of Mind*. Oxford University Press.
- Carruthers, P. 2005b. Conscious experience versus conscious thought. In U. Kriegel and K. Williford (Eds.), *Consciousness and Self-Reference*. MIT Press.
- Carruthers, P. 2004a. On being simple-minded. *American Philosophical Quarterly* 41: 205-220.
- Carruthers, P. 2004b. *The Nature of the Mind: An Introduction*. Routledge.
- Carruthers, P. 2004c. Why the question of animal consciousness might not matter very much. *Philosophical Psychology* 17: 83-102.
- Carruthers, P. 2003. Moderately massive modularity. In A. O'Hear (Ed.), *Minds and Persons*. Cambridge University Press
- Carruthers, P. 2000. *Phenomenal Consciousness*. Cambridge University Press.
- Carruthers, P. 1986. *Introducing Persons: theories and arguments in the philosophy of mind*. Croom Helm.
- Chiarello, C. 1995. Does the corpus callosum play a role in the activation and suppression of ambiguous word meanings? In F. L. Kitterle (Ed.), *Hemispheric communication: Mechanisms and models*, pp. 177-188. Lawrence Erlbaum Associates.
- Church, J. 1997. Fallacies or Analyses? In N. Block, O. Flanagan, and G. Güzeldere (Eds.), *The Nature of Consciousness: Philosophical Debates*, pp. 425-426. A Bradford Book (MIT Press).

- Churchland, P. S. 2002 a. *Brainwise: studies in neurophilosophy*. USA: MIT Press.
- Churchland, P. S. 2002 b. Self-representation in nervous systems. *Science* 296: 308-310.
- Churchland, P. S. 1986. *Neurophilosophy: Toward a Unified Science of the Mind-Brain*. MIT Press/Bradford Books.
- Churchland, P.S 1981. How many angels. . . ? *The Behavioral and Brain Sciences* 4: 103-104.
- Clark, A., 1997. *Being There*. MIT Press.
- Clark, A., and Chalmers, D. 1998. The extended mind. *Analysis* 58: 7-19.
- Code, C. 1997. Can the right hemisphere speak? *Brain and Language* 57: 38-59.
- Corballis, M. 1996. A dissociation in naming digits and colors following commissurotomy. *Cortex* 32: 515-525.
- Corballis, M. 1995a. Visual integration in the split brain. *Neuropsychologia* 33 (8): 937-959.
- Corballis, M. 1995b. Line bisection in a man with complete forebrain commissurotomy. *Neuropsychology* 9: 147-156.
- Corballis, M. 1994. Split decisions: problems in the interpretation of results from commissurotomized subjects. *Behavioural Brain Research* 64: 163-172.
- Corballis P., Funnell, M., Gazzaniga, M. 1999. A dissociation between spatial and identity matching in callosotomy patients. *Neuroreport* 10: 2183-7.
- Corballis, M. and Forster, B. 2003. Water under the bridge: interhemispheric visuomotor integration in a split-brain man. In E. Zaidel and M. Iacoboni (Eds.), *The Parallel Brain: The Cognitive Neuroscience of the Corpus Callosum*. Bradford Books (MIT Press).
- Critchley, H. 2005. Neural mechanisms of autonomic, affective, and cognitive integration. *The Journal of Comparative Neurology* 493: 154-166.
- Cronin-Golomb, A. 1986. Subcortical transfer of cognitive information in subjects with complete forebrain commissurotomy. *Cortex* 22: 499-519.
- Cummins, R. 1983. *Psychological Explanation*. MIT Press.
- Cummins, R. 1975. Functional analysis. *The Journal of Philosophy* 72: 741-756.

- Dainton, B. 2000. *Stream of Consciousness: unity and continuity in conscious experience*. Routledge.
- Damasio, A. 1999. *The Feeling of What Happens*. Harcourt, Inc.
- Damasio, A. 1995. *Descartes' Error*. Picador/Penguin.
- Damasio, A., Adolphs, R., and Damasio, H. 2003. The contributions of the lesion method to the functional neuroanatomy of emotion. In R. Davidson, K. Scherer, and H. Goldsmith (Eds.), *Handbook of Affective Sciences*, pp. 66-92. Oxford University Press.
- Davidson, D. 1982. Paradoxes of irrationality. In R. Wollheim and J. Hopkins (Eds.), *Philosophical Essays on Freud*, pp. 289-305. Cambridge University Press.
- Davidson, D. 1975. Thought and talk. In S. D. Guttenplan (Ed.), *Mind and Language*, pp. 7-24. Oxford University Press.
- Davidson, D. 1973. Freedom to Act. Reprinted in 1980, *Essays on Actions and Events*. Oxford University Press.
- Davidson, R. 1981. Cognitive processing is not equivalent to conscious processing. *The Behavioral and Brain Sciences* 4: 104-105.
- Day, P., and Ulatowska, H. 1979. Perceptual, cognitive, and linguistic development after early hemispherectomy: two case studies. *Brain & Language* 7: 17-33.
- Dehaene, S., Changeux, J., Naccache, L., Sackur, J., and Sergent, C. 2006. Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in Cognitive Sciences* 10: 204-211.
- Dehaene, S. and Naccache, L. 2001. Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition* 79: 1-37.
- Dennett, D. 2001. Are we explaining consciousness yet? *Cognition* 79: 221-237.
- Dennett, D. C. 1998. *Brainchildren: Essays on Designing Minds*. MIT Press.
- Dennett, D. 1997. The Path Not Taken. In N. Block, O. Flanagan, and G. Guzeldere (Eds.), *The Nature of Consciousness: Philosophical Debates*, pp. 417-419. A Bradford Book (MIT Press).

- Dennett, D. C. 1996. *Kinds of Minds: Towards an Understanding of Consciousness*. Basic Books.
- Dennett, D. 1993. The message is: There is no *medium*. *Philosophy and Phenomenological Research* 53: 919-31.
- Dennett, D. C. 1991. *Consciousness Explained*. Little, Brown, & Company.
- Dennett, D. 1981. Where Am I? In D. Hofstadter and D. Dennett (Eds.), *The Minds I*, pp. 217-231. Basic Books.
- Dennett, D. 1978. *Brainstorms*. Bradford Books.
- Dimond, S. 1980. *Neuropsychology: A Textbook of Systems and Psychological Functions of the Human Brain*. Butterworths.
- Doty, R. 1989. Some anatomical substrates of emotion, and their bihemispheric condition. In G. Gainottie and C. Caltagirone (Eds.), *Emotions and the Dual Brain*, pp. 56-82. Springer-Verlag.
- Doty, R., Negra, N., and Ymaga, K. 1973. The unilateral engram. *Acta Neurobiologie Experimentalis* 33: 711-728.
- Dretske, F. 1988. *Explaining Behavior*. MIT Press.
- Dretske, F. 1981. *Knowledge and the Flow of Information*. MIT Press.
- Eccles, J. 1994. *How the Self Controls Its Brain*. Springer-Verlag.
- Eccles, J. 1965. The brain and the unity of conscious experience. *Nineteenth Arthur Stanley Eddington Memorial Lecture*. Cambridge University Press.
- Efron, R. 1990. *The Decline and Fall of Hemispheric Specialization*. Erlbaum.
- Ellenberg, L., and Sperry, R. 1980. Lateralized division of attention in the commissurotomed and intact brain. *Neuropsychologia* 18: 411-418.
- Farber, I., and Churchland, P.S. 1994. Consciousness and the neurosciences: philosophical and theoretical issues. In M. Gazzaniga (Ed.), *The Cognitive Neurosciences*, pp. 1295-1306. MIT Press.
- Faust, M., Kravetz, S., and Nativ-Safrai, O. 2004. The representation of aspects of the self in the two cerebral hemispheres. *Personality and Individual Differences* 37: 607-619.

- Feest, U. 2003. Functional analysis and the autonomy of psychology. *Philosophy of Science* 70: 937–948.
- Fendrich, R, and Gazzaniga, M. 1989. Evidence of foveal splitting in a commissurotomy patient. *Neuropsychologia* 27: 273–81
- Fendrich, R, Wessinger, C., and Gazzaniga, M. 1994. Processing profiles at the retinal vertical midline of a callosotomy patient. *Society for Neuroscience Abstract* 20: 1579.
- Finlay, B., Darlington, R., and Nicastro, N. 2001. Developmental structure in brain evolution. *Behavioral and Brain Sciences* 24: 263-278.
- Finlay, B., and Darlington, R. 1995. Linked regularities in the development and evolution of mammalian brains. *Science* 286: 1578-1584.
- Flanagan, O. 1997. Prospects for a unified theory of consciousness or, what dreams are made of. In N. Block, O. Flanagan, and G. Güzeldere (Eds.), *The Nature of Consciousness*, pp. 97-109. MIT Press.
- Fodor, J., 2000. *The Mind Doesn't Work That Way*. MIT Press.
- Fodor, J. 1975. *The Language of Thought*. Harvard University Press.
- Franz, E., Eliassen, J., Ivry, R. and Gazzaniga, M. 1996. Dissociation of spatial and temporal coupling in the bimanual movements of callosotomy patients. *Psychological Science* 7: 306-310.
- Funnell, M. G., P. M. Corballis, and M. S. Gazzaniga. 2003. Temporal discrimination in the split brain. *Brain and Cognition* 53: 218-222.
- Fuster, J. M. 1995. *Memory in the Cerebral Cortex*. MIT Press.
- Gallistel, C. 1980. *The Organization of Action: A New Synthesis*. Erlbaum.
- Gandevia, S. 1978. The sensation of heaviness after surgical disconnection of the cerebral hemispheres in man. *Brain* 95: 327-346.
- Garson, J. 2001. (Dis)solving the binding problem. *Philosophical Psychology* 14: 381-392.
- Gazzaniga, M. 2005. Forty-five years of split-brain research and still going strong. *Nature Reviews Neuroscience* 6: 653-659.

- Gazzaniga, M. 2000. Cerebral specialization and interhemispheric communication: does the corpus callosum enable the human condition? *Brain* 123: 1293-1326.
- Gazzaniga, M. 1998. *The Mind's Past*. University of California Press.
- Gazzaniga, M. 1987. Perceptual and attentional processes following callosal section in humans. *Neuropsychologia* 25: 119-133.
- Gazzaniga, M. 1985. *The Social Brain*. Basic Books Inc.
- Gazzaniga, M. 1984. *Handbook of Cognitive Neuroscience*. Springer-Verlag.
- Gazzaniga, M. 1983a. Right hemisphere language following brain bisection: A 20 year perspective. *American Psychologist* 38: 525-537.
- Gazzaniga, M. 1983b Reply to Levy and Zaidel. *American Psychologist* 38: 547-549.
- Gazzaniga, M. 1970. *The Bisected Brain*. Appleton-Century-Crofts.
- Gazzaniga M., Bogen J., Sperry, R. 1963. Laterality effects in somesthesia following cerebral commissurotomy in man. *Neuropsychologia* 1: 209-15.
- Gazzaniga, M., Ivry, R., and Mangun, G. 1998. *Cognitive Neuroscience: The Biology of the Mind*. W.W. Norton & Company.
- Gazzaniga, M., and LeDoux, J., 1978. *The Integrated Mind*. Plenum Press.
- Gazzaniga, M., Smylie, C., Baynes, K., Hirst, W., and McCleary, C. 1984. Profiles of right hemisphere language and speech following brain bisection. *Brain and Language* 22: 206-220.
- Gazzaniga M., and Sperry, R. 1967. Language after section of the cerebral commissures. *Brain* 90: 131-48.
- Gazzaniga, M., Volpe, B., Smylie, C., Wilson, D., and LeDoux, J. 1979. Plasticity in speech organization following commissurotomy. *Brain* 102: 808-815.
- Geffen, G., Nilsson, J., Simpson, D., and Jeeves, M. 1994. The development of interhemispheric transfer of tactile information in cases of callosal agenesis. In M. Lassonde and M. Jeeves (Eds.), *Callosal Agenesis: A Natural Split Brain?*, pp. 185-197. Plenum Press.
- Goodale, M., and Milner, A. 1992. Separate visual pathways for perception and action. *Trends in Neurosciences* 15: 20-25.

- Gott, P. 1973. Cognitive abilities following right and left hemispherectomy. *Cortex* 9: 266-274.
- Gould, S. 1977. *Ontogeny and Phylogeny*. Harvard University Press.
- Grabowecky, M., and Kingstone, A.. 2004. Can semantic information be transferred between hemispheres in the split-brain? *Brain & Cognition* 55: 310-313.
- Grice, P. 1975. Method in Philosophical psychology (From the Banal to the Bizarre). *Proceedings and Addresses of the American Philosophical Association* 48: 23-53.
- Grice, P. 1961. The causal theory of perception. *The Aristotelian Society: Proceedings, Supplementary Volume* 35: 121-152.
- Handy, T., Gazzaniga, M., and Ivry, R. 2003. Cortical and subcortical contributions to the representation of temporal information. *Neuropsychologia* 41: 1461-1473.
- Hardcastle, V. 1998. The binding problem. In Bechtel, B. and Graham, G. (Eds.) *A Companion to Cognitive Science*, pp. 555-565. Blackwell Publishers Inc.
- Harman, G. 1987. (Nonsolopsistic) conceptual role semantics. In E. LePore (Ed.), *Semantics of natural language*, pp. 55-81. Academic Press.
- Harman, G. 1982. Conceptual role semantics. *Notre Dame Journal of Formal Logic*. 23: 242-56.
- Harris, L. J. 1995. The corpus callosum and hemispheric communication: an historical survey of theory and research. In F. L. Kitterle (Ed.), *Hemispheric communication: Mechanisms and models*, pp. 1-59. Lawrence Erlbaum Associates.
- Hayes, A., Davidson, M., Keele, S., and Rafal, R. 1998. Toward a functional analysis of the basal ganglia. *Journal of Cognitive Neuroscience* 10: 178-198.
- Heller, W., Nitzchke, J., and Miller, G. 1998. Lateralization in emotion and emotional disorders. *Current Directions in Psychological Science* 7: 26-32.
- Hellige, J. 1993. *Hemispheric Asymmetry: What's right and what's left*. Harvard University Press.
- Hobson, J. 1993. Understanding persons: the role of affect. In S. Baron-Cohen, H. Tager-Flusberg, and D. Cohen (Eds.), *Understanding Other Minds: Perspectives From Autism*, pp. 204-227. Oxford University Press.

- Hochman, E., and Eviatar, Z. 2004. Does each hemisphere monitor the ongoing process in the contralateral one? *Brain and Cognition* 55: 314-321.
- Holtzman, J., 1983. Interactions between cortical and subcortical visual areas: evidence from human commissurotomy patients. *Vision Research* 24: 801-813.
- Holtzman, J. and Gazzaniga, M. 1985. Enhanced dual task performance following corpus commissurotomy in humans. *Neuropsychologia* 23: 315-321.
- Holtzman, J., and Gazzaniga, M. 1982. Dual task interactions due exclusively to limits in processing resources. *Science* 218: 1325-1327.
- Holtzman, J., Sidtis, J., Volpe, B., Wilson, D., and Gazzaniga, M. 1981. Dissociation of spatial information for stimulus localization and the control of attention. *Brain* 104: 861-872.
- Holtzman, J., Volpe, B., and Gazzaniga, M. 1984. Spatial orientation following commissural section. In R. Parasuraman and D. Davies (Eds.), *Varieties of Attention*, pp. 375-394. Academic Press.
- Hughlings Jackson, J. (Ed). 1958. *Selected writings of John Hughlings Jackson* (Vols 1 and 2). Staples Press.
- Humphreys, G. 2003. Conscious visual representations built from multiple binding processes: evidence from neuropsychology. In A. Cleeremans (Ed.), *The Unity of Consciousness: Binding, Integration, and Dissociation*, pp. 114-131. Oxford University Press.
- Hurley, S. Forthcoming. Varieties of externalism. In R. Menary (Ed.), *The Extended Mind*. Ashgate.
- Hurley, S. 2003. Action, the unity of consciousness, and vehicle externalism. In A. Cleeremans, ed. *The Unity of consciousness: Binding, Integration, and Dissociation*, pp. 72-91. Oxford University Press.
- Hurley, S. 1998. *Consciousness in Action*. Harvard University Press.
- Hurley, S. 1994. Unity and objectivity. In C. Peacocke (Ed.), *Objectivity, Simulation and the Unity of Consciousness*, pp. 49-77. Oxford University Press.
- Hurley, S. and Noë, A. 2003. Neural plasticity and consciousness. *Biology and Philosophy* 18: 131-168.
- Ishiai, S., Koyama, Y., and Furuya, T. 2001. Conflict and integration of spatial attention between disconnected hemispheres. *Journal of Neurology, Neurosurgery, and Psychiatry* 71: 472-477.

- Ivry, R. 1993. Cerebellar involvement in the explicit representation of temporal information. *Annals of the New York Academy of Sciences* 682: 214-230.
- Jackendoff, R. 1987. *Consciousness and the Computational Mind*. MIT Press.
- Jeannerod, M. 1997. *The Cognitive Neuroscience of Action*. Wiley-Blackwell.
- Johnston, M., 2003. Human concern without superlative selves. In R. Martin and J. Barresi (Eds.), *Personal Identity*, pp. 260-291. Blackwell Publishing Ltd.
- Johnson, S., Corballis, P., and Gazzaniga, M. 1999. Roles of the cerebral hemispheres in planning prehension: accuracy of movement selection following callosotomy. *Journal of Cognitive Neuroscience* 11 (Supplement): 85.
- Joseph, R. 1990. *Neuropsychology, Neuropsychiatry, and Behavioral Neurology*. Springer.
- Joynt, R. 1981. Are two heads better than one? *The Behavioral and Brain Sciences* 4: 108-109.
- Kagan, J. 1981. *The Second Year: The Emergence of Self-Awareness*. Harvard University Press.
- Kaplan, J., and Zaidel, E. 2002. *The Neuropsychology of Executive Function: Hemispheric Contributions to Error Monitoring and Feedback Processing*. UCLA unpublished doctoral dissertation. University of California, Los Angeles, CA.
- Kingstone, A., Enns, J., Mangun, G., and Gazzaniga, M. 1995. Guided visual search is a left hemisphere process in split-brain patients. *Psychological Science* 6: 118-121.
- Kingstone, A., Friesen, C., Gazzaniga, M. 2000. Reflexive joint attention depends on lateralized cortical connections. *Psychological Science* 11: 159-166.
- Kingstone, A., and Gazzaniga, M. 1995. Subcortical transfer of higher order information: More illusory than real? *Neuropsychology* 9: 321-328.
- Kolb, B., and Whishaw, I. 1990. *Fundamentals of Human Neuropsychology*. W. H. Freeman.

- Kopp, B., and Rist, F. 1999. An event-related brain potential substrate of disturbed response monitoring in paranoid schizophrenic patients. *Journal of Abnormal Psychology* 108: 337-346.
- Koriat, A. 2000. The feeling of knowing: some metatheoretical implications for consciousness and control. *Consciousness and Cognition* 9: 149-171.
- Kreuter, C., Kinsbourne, M., and Trevarthen, C. 1972. Are disconnected cerebral hemispheres independent channels? A preliminary study of the effect of unilateral loading on bilateral finger tapping. *Neuropsychologia* 10: 453-461.
- Lambert, A. 1991. Interhemispheric interactions in the split-brain. *Neuropsychologia* 29: 941-948.
- Lamme, V. 2004. Separate neural definitions of visual consciousness and visual attention: A case for phenomenal awareness. *Neural Networks* 17: 861-872.
- Lamme, V. 2003. Why visual attention and awareness are different. *Trends in Cognitive Science* 7: 12-18.
- Lassonde, M. and Jeeves, M., (Eds.). 1991. *Callosal Agenesis: A Natural Split Brain?* Plenum Press.
- LeDoux, J. 2002. *Synaptic Self*. Viking Press.
- LeDoux, J. 1993a. Emotional memory: in search of systems and synapses. *Annals of the New York Academy of Science* 702: 149-157.
- LeDoux, J., 1993b. Emotional memory systems in the brain. *Behavioural Brain Research* 58: 69-79.
- LeDoux, J. 1990. Emotion circuits in the brain. *Annual Review of Neuroscience* 23: 155-184.
- LeDoux, J., Cicchetti, P., Xagoraris, A., and Romanski, I. 1990. The lateral amygdaloid nucleus: sensory interface of the amygdale in fear conditioning. *Journal of Neuroscience* 10: 1062-1069.
- LeDoux, J., and Gazzaniga, M. 1981. The brain and the split brain: A duel with duality as a model of mind. *The Behavioral and Brain Sciences* 4: 109-110.
- Lee, G., Meador, K., Loring, D., Allision, J., Brown, W., Paul, L., Pillai, J., Lavin, T. 2004. Neural substrates of emotion as revealed by functional magnetic resonance imaging. *Cognitive and Behavioral Neurology* 17: 9-17.

- Leslie, A., and Keeble, S. 1987. Do six-month-old infants perceive causality? *Cognition* 25: 265-288.
- Lewis, D. 1972. Psychophysical and theoretical identification. *Australian Journal of Philosophy* 50: 249-258.
- Levine, J. 2001. *Purple Haze: The Puzzle of Consciousness*. Oxford University Press.
- Levy, J. 1983. Language, cognition, and the right hemisphere: a response to Gazzaniga. *American Psychology* 38: 538-541.
- Levy, J., Nebes, R., and Sperry, R. 1971. Expressive language in the surgically separated right hemisphere. *Cortex* 7: 49-58.
- Levy, J., and Trevarthen, C. 1977. Perceptual, semantic, and phonetic aspects of elementary language processes in split-brain patients. *Brain* 100: 105-118.
- Levy, J., Trevarthen, C., and Sperry, R. 1972. Perception of bilateral chimeric figures following hemispheric disconnection. *Brain* 96: 61-78.
- Loar, B. 1982. Conceptual role and truth-conditions. *Notre Dame Journal of Formal Logic* 23: 272-283.
- Loar, B. 1981. *Mind and meaning*. Cambridge University Press.
- Locke, J. (1690) *An Essay Concerning Human Understanding*: Book 2, Chapter 27. (Many editions now available.)
- Lockwood, M. 1994. Issues of unity and objectivity. In C. Peacocke (Ed.), *Objectivity, Simulation, and the Unity of Consciousness*, pp. 89-95. Oxford University Press.
- Lockwood, M. 1989. *Mind, brain and the quantum*. Basil Blackwell Ltd.
- Luck, S., Hillyard, S., Mangun, G., and Gazzaniga, M. 1989. Independent hemispheric attentional systems mediate visual search in split-brain patients. *Nature* 342: 543-545.
- Luck, S., Hillyard, S., Mangun, G., and Gazzaniga, M. 1994. Independent attentional scanning in the separated hemispheres of split-brain patients. *Journal of Cognitive Neuroscience* 6: 84-91.
- Lycan, W. 1984. *Logical form in natural language*. MIT Press.

- MacKay, D. 1966. Cerebral organization and conscious control of action. In J. Eccles (Ed.), *Brain and Conscious Experience*, 422-440. Springer Verlag.
- Mangun, G., Hillyard, S., Luck, S., Handy, T., Plager, R., Clark, V., Loftus, W., and Gazzaniga, M. 1994. Monitoring the visual world: hemispheric asymmetries and subcortical processes in attention. *Journal of Cognitive Neuroscience* 6: 267-275.
- Marcel, A. 1994. What is relevant to the unity of consciousness? In C. Peacocke (Ed.), *Objectivity, Simulation and the Unity of Consciousness*, pp. 79-88. Oxford University Press.
- Marcel, A. 1993. Slippage in the unity of consciousness, in G. Bock and J. Marsh (Eds.), *Experimental and Theoretical Studies of Consciousness*, pp. 168-179. John Wiley & Sons.
- Marks, C., 1980. *Commissurotomy, Consciousness, and Unity of Mind*. Bradford Books.
- Marr, D. 1982. *Vision*. W. H. Freeman & Co., Ltd.
- McCloskey, D. 1973. Position sense after surgical disconnection of the cerebral hemispheres in man. *Brain* 96: 269-276.
- McCulloch, W. 1949. Mechanisms for the spread of epileptic activation of the brain. *Electroencephalography and clinical neurophysiology* 1: 19-24.
- McDowell, J., 1994. *Mind and World*. Harvard University Press.
- McGinn, C. 1982. The structure of content. In A. Woodfield (Ed.), *Thought and Object*, pp. 207-258. Oxford University Press.
- McGurk, H.; and MacDonald, J. 1976. Hearing lips and seeing voices, *Nature* 264: 746-748.
- Merker, B. 2007. Consciousness without a cerebral cortex: A challenge for neuroscience and medicine. *Behavioral and Brain Sciences* 30: 63-164.
- Metzinger, T., 2003. *Being No One*. MIT Press.
- Metzinger, T. 2000. The subjectivity of subjective experience: A representationalist analysis of the first-person perspective. In T. Metzinger (Ed.), *Neural Correlates of Consciousness*, pp. 285-306. MIT Press.
- Miller, M., Kingstone, A. 2005. Taking the high road on subcortical transfer. *Brain and Cognition* 57: 162-164.

- Milner, A. and Goodale, M. 1996. *The Visual Brain in Action*. Oxford University Press.
- Milner, B., Taylor, L., and Jones-Gotman, M. 1990. Lessons from cerebral commissurotomy: Auditory attention, haptic memory, and visual images in verbal-associative learning. In C. Trevarthen (Ed.) *Brain Circuits and Functions of the Mind: Essays in Honor of Roger W. Sperry*, pp. 293-303. Cambridge University Press.
- Moscovitch, M., Chein, J., Talmi, D., and Melanie, C. 2007. Learning and memory. In B. Baars and N. Gage (Eds.) *Cognition, Brain, and Consciousness*, pp. 255-293. Academic Press.
- Mosenthal, W. 1995. *A Textbook of Neuroanatomy: With Atlas and Dissection Guide*. Taylor & Francis.
- Myers, J. 1955. Neural basis of bilateral perceptual integration. *Science* 122: 877.
- Myers, J., and R. W. Sperry. 1985. Interhemispheric communication after section of the forebrain commissures. *Cortex*: 21, 249-260.
- Myers, R., and Sperry, R., 1985. Interhemispheric communication through the corpus callosum. Mnemonic carry-over between the hemispheres. *Archives of Neurology and Psychiatry*, 80: 298-303.
- Myers, R., and Sperry, R., 1953. Interocular transfer of a visual form discrimination habit in cats after section of the optic chiasma and corpus callosum [Abstract]. *The Anatomical Record* 115: 351-352.
- Nagel, T., 1979. *Mortal Questions*. Cambridge University Press.
- Naikar, N. 1996. Perception of apparent motion of colored stimuli after commissurotomy. *Neuropsychologia* 34: 1041-1049.
- Oakes, L. M. (1994). Development of infants' use of continuity cues in their perception of causality. *Developmental Psychology* 30: 869-870.
- O'Brien, G., and Opie, J. 1994. The disunity of consciousness. *Australasian Journal of Philosophy* 76: 378-395.
- Ogden, J. 1989. Visuospatial and other "right-hemispheric" functions after long recovery periods in left-hemispherectomized subjects. *Neuropsychologia* 27: pp. 765-776.

- Parfit, D., 2003a. Why our identity is not what matters. In R. Martin and J. Barresi (Eds.), *Personal Identity*, pp. 115-143. Blackwell Publishing Ltd.
- Parfit, D., 2003b. The unimportance of identity. In R. Martin and J. Barresi (Eds.), *Personal Identity*, pp. 293-317. Blackwell Publishing Ltd.
- Parfit, D. 1995. The unimportance of identity. In H. Harris (Ed.), *Identity*. Oxford University Press.
- Parfit, D. 1984. *Reasons and Persons*. Oxford University Press.
- Parfit, D., 1971. Personal Identity. *Philosophical Review* 80: 3-27.
- Parvisi, J., and A. R. Damasio. 2001. Consciousness and the brainstem. *Cognition* 79: 135-159.
- Peacocke, C. 1992. *A Study of Concepts*. MIT Press.
- Peacocke, C. 1986. *Thoughts*. Blackwell.
- Perry, J., 1975. The importance of being identical. In Amelie Rorty (Ed.), *The Identities of Persons*, pp. 67-90. University of California Press.
- Pietroski, P. 1996. Experiencing the Facts: Critical notice of John McDowell's *Mind and World*. *Canadian Journal of Philosophy* 26: 613-636.
- Plourde, G., Sperry, R. 1984. Left hemisphere involvement in left spatial neglect from right-sided lesions: a commissurotomy study. *Brain* 107: 95-106.
- Praamstra, P., Boutsen, L., and Humphreys, G. 2005. Frontoparietal control of spatial attention and motor intention in human EEG. *Journal of Neurophysiology* 94: 764-774.
- Preilowski, B. 1990. "Intermanual transfer, interhemispheric interaction, and handedness in man and monkey" in *Brain Circuits and Functions of the Mind*, ed. C. Trevarthen. Cambridge University Press, U.S.A., 129-139.
- Prinz, J. 2007a. Mental pointing: phenomenal knowledge without phenomenal concepts. [Journal of Consciousness Studies](#), Volume 14, pp. 184-211(28).
- Prinz, J. 2007b: Accessed, accessible, and inaccessible: where to draw the phenomenal line. Commentary on Block. *Behavioral and Brain Sciences*, 30, 521-522.

- Prinz, J. 2003. A neurofunctional theory of consciousness. In A. Brook and K. Akins (Eds.) *Philosophy and Neuroscience*. Cambridge University Press.
- Prinz, J. 2000: A neurofunctional theory of visual consciousness. *Consciousness and Cognition*, 9, 243-259.
- Puccetti, R. 1989. Two brains, two minds? Wigan's theory of mental duality. *The British Journal for the Philosophy of Science* 40: 137-144.
- Puccetti, R. 1981. The case for mental duality: evidence from split-brain data and other considerations. *Behavioral and Brain Sciences* 4: 93-123.
- Putnam, H. 1975. *Mind Language and Reality: Philosophical Papers Volume 2*. Cambridge University Press.
- Quartz, S. R., and T. J. Sejnowski. 2002. *Liars, Lovers, and Heroes: What the New Brain Science Reveals about How We Become Who We Are*. Harper-Collins.
- Quine, W. 1964. *Word and Object*. MIT Press.
- Ramsey, W., Stich, S., Garon, J. 1990. Connectionism, eliminativism, and the future of folk psychology. *Philosophical Perspectives* 4, Action Theory and Philosophy of Mind: 499-533.
- Reuter-Lorenz PA, Fendrich R. 1990. Orienting attention across the vertical meridian: evidence from callosotomy patients. *Journal of Cognitive Neuroscience* 2: 232-8.
- Reuter-Lorenz, P., Jha, A., Rosenquist, J. 1996. What is inhibited in inhibition of return? *Journal of Experimental Psychology Human Perception and Performance* 22: 376-378.
- Reuter-Lorenz, P., and Miller, A. 1998. The cognitive neuroscience of human laterality: lessons from the bisected brain. *Current Directions in Psychological Science* 7: 15-20.
- Reuter-Lorenz, P., Nozawa G., Gazzaniga, M., Hughes, H. 1995. Fate of neglected targets: a chronometric analysis of redundant target effects in the bisected brain. *Journal of Experimental Psychology: Human Perception and Performance* 21: 211-230.
- Revonsuo, A. 2000. Prospects for a scientific research program on consciousness. In T. Metzinger (Ed.), *Neural Correlates of Consciousness*, pp. 57-75. MIT Press.

- Rey, G., 1997. *Contemporary Philosophy of Mind*. T.J. Press Ltd.
- Rey, G. 1994. Dennett's unrealistic psychology. *Philosophical Topics* 22: 259-289.
- Rey, G. 1983. The lack of a case for mental duality. Commentary on Puccetti, R., "The Case for Mental Duality: Evidence from Split-Brain Data and Other Consideration." *Behavioral and Brain Sciences* 6: 733-734.
- Rey, G. 1976. "Survival" in *The Identities of Persons*, ed. A. Rorty. Berkeley: University of California Press: 41-66.
- Rey, G. 1975. Persons and personalities. Presented at *Colloquium on Personal Identity* at the 72nd Annual Meeting of the American Philosophical Association, Eastern Division, December 1975.
- Rigterink, R., 1980. Puccetti and brain bisection: an attempt at mental division. *Canadian Journal of Philosophy* 10: 429-452.
- Robertson, L. 2003. Binding, Spatial Attention and Perceptual Awareness. *Nature Reviews Neuroscience* 4: 93-102.
- Roser, M., and Corballis, M. 2001. Interhemispheric neural summation in the split brain with symmetrical and asymmetrical displays. *Neuropsychologia* 40: 1300-1312.
- Roser, M., Fugelsang, J., Dunbar, K., Corballis, P., and Gazzaniga, M. 2005. Dissociating processes supporting causal perception and causal inference in the brain. *Neuropsychology* 19: 591-602.
- Ross, E. 1997. Cortical representation of the emotions. In M. Trimble and J. Cummings (Eds.), *Contemporary Behavioral Neurology*, pp. 107-216. Butterworth-Heineman.
- Savazzi, S., Fabri, M., Rubboli, G., Paggi, A., Tassinari, C., and Marzi, C. 2007. Interhemispheric transfer following callosotomy in humans: Role of the superior colliculus. *Neuropsychologia* 45: 2417-2427.
- Schacter, D. L. 1996. *Searching for Memory: The Brain, the Mind, and the Past*. Basic Books.
- Schacter, S., Singer, J. 1962. Cognitive, social, and physiological determinants of emotional state. *Psychological Review* 69: 379-99.
- Schiffer, F. 1998. *Of Two Minds*. The Free Press/Simon & Schuster.

- Schiffer, F.; Zaidel, E.; Bogen, J.; Chasan-Taber, S. 1998. Different psychological status in the two hemispheres of two split-brain patients. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology* 11: 151-156.
- Schore, A. N. 1994. *Affect Regulation and the Origin of the Self*. Hillsdale, N. J.: Lawrence Erlbaum & Associates.
- Searle, J. 1992. *The Rediscovery of the Mind*. USA: MIT.
- Searle, J. 1990a. Collective intentionality and action. In E. Lepore and R. van Gulick (eds.), *Intentions in Communications*. Cambridge, MA: MIT Press.
- Segal, G. 2000. *A Slim Book about Narrow Content*. MIT Press.
- Sergent, J. 1990. Furtive incursions into bicameral minds: integrative and coordinating role of subcortical structures. *Brain* 113: 537-568.
- Sergent, J. 1986. Subcortical coordination of hemisphere activity in commissurotomed patients. *Brain* 109: 357-369.
- Sergent, J. 1983. Unified response to bilateral hemispheric stimulation by a split-brain patient. *nature* 305: 800-802.
- Seymour, S., Reuter-Lorenz, P., Gazzaniga, M. 1994. The disconnection syndrome: basic findings reaffirmed. *Brain* 117: 105-115.
- Shallice, T. 1997. Modularity and consciousness. In N. Block, O., Flanagan, and G. Güzeldere (eds.), *The Nature of Consciousness*, MIT Press: 255-276.
- Shanahan, M., and Baars, B. 2005. Applying global workspace theory to the frame problem. *Cognition* 98: 157-176.
- Sidtis, J., Volpe, B., Holtzman, J., Wilson, D., Gazzaniga, M. 1981. Cognitive interaction after staged callosal section: evidence for transfer of semantic activation. *Science* 212: 344-346.
- Smith, A. 1969. Nondominant hemispherectomy. *Neurology* 19: 442-445.
- Smith, A. 1966. Speech and other functions after left (dominant) hemispherectomy. *Journal of Neurology, Neurosurgery, and Psychiatry* 29: 467-471.
- Smith, S.; Bulman-Fleming, M. 2004. An examination of the right-hemisphere hypothesis of the lateralization of emotion. *Brain and Cognition* 57: 210-213.
- Smythies, J. 1994. Requiem for the Identity Theory. *Inquiry* 37: 311-329.

- Sperry, R. 1990. Forebrain commissurotomy and conscious awareness. In C. Trevarthen (Ed.), *Brain Circuits and Functions of the Mind*, pp. 371-386. Cambridge University Press.
- Sperry, R. 1986. Consciousness, personal identity, and the divided brain. In F. Leporé, M. Ptito, and H. Jasper (Eds.), *Two Hemispheres—One Brain: Functions of the Corpus Callosum*, pp. 3-20. Alan R. Liss, Inc.
- Sperry, R. 1977. Reply to Professor Puccetti. *The Journal of Medicine and Philosophy* 2: 145-146.
- Sperry, R. 1975. Mental phenomena as causal determinants in brain function. *Process Studies* 5: 247-256.
- Sperry, R. 1974. Lateral specialization in the surgically separated hemispheres. In F. O. Schmitt and F. G. Worden, eds., *The Neurosciences: Third Study Program*, pp. 5-19. Cambridge: MIT Press.
- Sperry, R. 1969a. Toward a theory of mind. *Proceedings of the National Academy of Sciences* 63: 230-231 (Abstract).
- Sperry, R. 1969b. A modified concept of consciousness. *Psychological Review* 76: 532-536.
- Sperry, R. 1968. Hemisphere disconnection and unity in conscious awareness. *American Psychologist* 23: 723-733.
- Sperry, R. 1966. Brain bisection and mechanisms of consciousness. In J. Eccles (Ed.), *Brain and conscious experience*, pp.298-313. Springer-Verlag.
- Sperry, R., Gazzaniga, M., and Bogen, J. 1969a. Interhemispheric relationships: The neocortical commissures, syndromes of hemispheric disconnection. In P. Vinken and G. Bruyn (Eds.), *Handbook of Clinical Neurology, Vol. 4, Disorders of Speech, Perception and Symbolic Behavior*, pp. 145-153. Elsevier/North Holland Biomedical Press.
- Sperry, R., Stamm, J., and Miner, N. 1956. Relearning tests for interocular transfer following division of optic chiasm and corpus callosum in cats. *Journal of Comparative Physiology and Psychology* 49: 529-533.
- Sperry, R., Zaidel, E., Zaidel, D. 1979. Self recognition and social awareness in the disconnected minor hemisphere. *Neuropsychologia* 17: 153-166.
- Squire, L., and Knowlton, B. 1995. Memory, hippocampus, and brain systems. In M. Gazzaniga (Ed.), *The Cognitive Neurosciences*. MIT Press.

- Stark, R., Bleile, K., Brandt, J., Freeman, J., Vining, E. 1995. Speech-Language Outcomes of Hemispherectomies in Children and Young Adults. *Brain and Language* 51: 406-421.
- Stein, B., Price, D., and Gazzaniga, M. 1989. Pain perception in a man with total corpus callosum transection. *Pain* 38: 51-6.
- Stephan, K., Marshall, J., Friston, K., Rowe, J., Ritzl, A. Zilles, K. and Fink, F. 2003. Lateralized cognitive processes and lateralized task control in the human brain. *Science* 301: 384-386.
- Sterelny, K., and Griffiths, P. 1999. *Sex and Death*. University of Chicago Press.
- Strawson, G. 2003. The self. In R. Martin and J. Barresi (Eds.), *Personal Identity*: pp. 335-377. Blackwell Publishing Ltd.
- TenHouten, W., Hoppe, K., Bogen, J., and Walter, D. 1986. Alexithymia: an experimental study of cerebral commissurotomy patients and normal control subjects. *American Journal of Psychiatry* 143: 312-316.
- Thompson, E., and Varela, F. 2001. Radical embodiment: neural dynamics and consciousness. *Trends in Cognitive Sciences* 5: 418-425.
- Treisman, A. 2003. Consciousness and perceptual binding. In A. Cleeremans (Ed.), *The Unity of Consciousness: Binding, Integration, and Dissociation*, pp.95-113. Oxford University Press.
- Trevarthen, C. 1984. Biodynamic structures, cognitive correlates of motive sets and the development of motives in infants. In W. Prinz and A. Sanders (Eds.), *Cognition and Motor Processes*, pp. 327-350. Springer-Verlag.
- Trevarthen, C. 1978. Modes of perceiving and modes of acting. In H. Pick, Jr. and E. Saltzman (Eds.), *Modes of Perceiving and Processing Information*. pp. 99-136. Lawrence Erlbaum Associates.
- Trevarthen, C. 1976. Psychological activities after forebrain commissurotomy in man. In F. Michel and B. Schott (Eds.), *Les syndromes de disconnexion calleuse chez l'homme*. Hôpital Neurologique, Lyon.
- Trevarthen, C. 1974a. Analysis of cerebral activities that generate and regulate consciousness in commissurotomy patients. In S. Dimond (Eds.), *Hemisphere function in the human brain*, pp. 235-61. Elek Science.
- Trevarthen, C. 1974b. Functional relations of disconnected hemispheres with the brain stem, and with each other: Monkey and man. In M. Kinsbourne and W.

- Smith (Eds.), *Hemispheric disconnection and cerebral function*, pp. 187-207. Charles C. Thomas.
- Trevarthen, C. 1968. Two mechanisms of vision in primates. *Psychologische Forschung* 31: 299-337.
- Trevarthen, C., and Sperry, R. 1973. Perceptual unity of the ambient visual field in human commissurotomy patients. *Brain* 96: 547-570.
- Turk, D., Heatherton, T., Kelley, W., Funnell, M., Gazzaniga, M., and Macrae, C. 2002. Mike or me? Self-recognition in a split-brain patient. *Nature Neuroscience* 5: 841-842.
- Tye, M. 2003. *Consciousness and Persons*. MIT Press.
- Tye, M. 2000. *Consciousness, Color, and Content*. MIT Press.
- Tye, M. 1995. *Ten Problems of Consciousness*. MIT Press.
- Ułinski, P. The cerebral cortex of reptiles. 1990. In E. Jones and A. Peters (Eds.) *Cerebral Cortex Volume 8A: Comparative Structure and Evolution of Cortex Part I*: 139-216. Plenum Press.
- Van Lancker, D. 1997. Rags to riches: our increasing appreciation of cognitive and communicative abilities of the human right cerebral hemisphere. *Brain and Language* 57: 1-11.
- Varela, F., Thompson, E., and Rosch, E. 1991. *The Embodied Mind*. MIT Press.
- Wegner, D. 2002. *The Illusion of Conscious Will*. MIT Press.
- Weissman, D. and M. Banich. 2000. The cerebral hemispheres cooperate to perform complex but not simple tasks. *Neuropsychology* 14: 41-59.
- Weylman, S., Brownell, H., and Gardner, H. 1988. "It's what you mean, not what you say": Pragmatic language use in brain-damaged patients. In F. Plum (Ed.), *Language, Communication, and the Brain*. Raven.
- Wiggins, D. 1976. In A. Rorty (Ed.), *The Identities of Persons*, pp. 139-174. University of California Press.

- Williams, B. 1957, .Personal identity and individuation. *Proceedings of the Aristotelian Society* 67: 229-252. (Reprinted in B. Williams, 1973, *Problems of the Self*, Cambridge University Press.)
- Wilson, J. 1999. *Biological Individuality: The Identity and Persistence of Living Entities*. Cambridge University Press.
- Wilson, R. *Boundaries of the Mind: The Individual in the Fragile Sciences*. Cambridge University Press.
- Wolford, G., Miller, M., and Gazzaniga, M. 2004. Split decisions. In M. Gazzaniga (Ed.) *The Cognitive Neurosciences*, pp. 1189-1199. MIT Press.
- Wolford, G., Miller, M., Gazzaniga, M. 2000. The Left Hemisphere's Role in Hypothesis Formulation. *The Journal of Neuroscience* 20: 1-4.
- Zaidel, D. 1990. Long-term semantic memory in the two cerebral hemispheres. In C. Trevarthen (Ed.), *Brain Circuits and Functions of the Mind*, pp. 129-139. Cambridge University Press.
- Zaidel, E. 2000. Hemispheric contributions to pragmatics. *Brain and Cognition* 43: 438-443.
- Zaidel, E. 1998. Stereognosis in the chronic split brain: hemispheric differences, ipsilateral control, and sensory integration across the midline. *Neuropsychologia* 36: 1033-1047.
- Zaidel, E. 1991. Language functions in the two hemispheres following complete cerebral commissurotomy and hemispherectomy. In F. Boller and J. Grafman (Eds.), *Handbook of Neuropsychology*, Vol. 4., pp. 115-150. Elsevier.
- Zaidel, E. 1983. A response to Gazzaniga: language in the right hemisphere, convergent perspectives. *American Psychologist* 38: 542-546.
- Zaidel, E., Clarke, J., Suyenobu, B. 1990. Hemispheric independence: a paradigm case for cognitive neuroscience. In A. Scheibel and A. Wechsler (Eds.), *Neurobiology of higher cognitive function*, pp. 297-355. Guilford Press.
- Zaidel, E. and Siebert, K. 1997. Speech in the disconnected right hemisphere? *Brain and Language* 60: 188-192.
- Zaidel, D., and Sperry, R. 1977. Some long-term motor effects of cerebral commissurotomy in man. *Neuropsychologia* 15: 193-204.

- Zaidel D, Sperry R. 1974. Memory impairment after commissurotomy in man. *Brain* 97: 263–272.
- Zaidel, E., Iacoboni, M. 2003. Sensorimotor integration in the split brain. In E. Zaidel and M. Iacoboni (Eds.), *The Parallel Brain*, pp. 319-336. MIT Press.
- Zaidel, E., Zaidel, D., and Bogen, J. 1999 The split brain.. In G. Adelman and B. Smith (Eds.) *Encyclopedia of Neuroscience* 2nd Edition, pp. 1027-1032. Elsevier.
- Zaidel, E., Zaidel, D., Sperry, R. 1981. Left and right intelligence: Case studies of Raven's Progressive Matrices following brain bisection and hemi-decortication. *Cortex* 17: 167-186.