

# Coping with the World: Tools of Pragmatism

*Mind, Brain and the Intentional Vocabulary*

**Anders Kristian Krabberød**



Hovedoppgave

Filosofisk Institutt

UNIVERSITETET I OSLO

Høsten 2004

## Contents

<b>1. Introduction.....</b>	<b>2</b>
<b>2. Different Ways of Describing the Same Thing .....</b>	<b>7</b>
<b>3. Rorty and Vocabularies .....</b>	<b>11</b>
The Vocabulary-Vocabulary .....	13
Reduction and Ontology.....	17
<b>4. The Intentional Vocabulary and Folk Psychology: the Churchlands and Eliminative Materialism .....</b>	<b>21</b>
Eliminative Materialism.....	22
Dire Consequences.....	24
Objecting against Eliminative Materialism.....	25
1. <i>The first objection: Eliminative materialism is a non-starter</i> .....	27
2. <i>The second objection: What could possibly falsify Folk Psychology?</i> .....	29
3. <i>The third objection: Folk Psychology is used for a wide range of non-scientific purposes. Folk Psychology is not a scientific theory. In fact, it is not a theory at all.</i> .....	30
4. <i>The fourth objection: Could there be an alternative?</i> .....	31
5. <i>The fifth objection: Folk Psychology consists of functional kinds</i> .....	32
Reviewing the Objections.....	33
<i>Is Folk Psychology a Theoretical Vocabulary?</i> .....	34
<i>Mental Content</i> .....	36
<b>5. Natural Kinds and Emotions: Paul Griffiths.....</b>	<b>40</b>
Conceptual Analysis and Propositional Attitudes .....	40
Natural Kinds and Scientific Explanations.....	43
The Essentialist View.....	46
Paul Griffiths and Natural Kinds .....	47
Natural Kinds in Biology .....	48
Natural Kinds and the Special Sciences.....	51
Eliminating Emotion .....	53
<b>6. Griffiths and the Churchlands: Compare and Contrast.....</b>	<b>56</b>
The Normative Aspect of Folk Psychology .....	58
Elimination or Revision?.....	61
<b>7. Concluding Remarks.....</b>	<b>65</b>
Natural Kinds and Trans-theoretic Identity .....	68
Conclusion .....	69
<b>Bibliography .....</b>	<b>71</b>

# 1. Introduction

In this paper I will explore the tension that arises when we describe ourselves as intentional beings, on one side, and biological or physical beings on the other. That is, I will take a closer look at the age old problem of how the mind relates to the brain. More specifically, I will take a look at the relation between the scientific explanations of the mind and the commonsensical framework we rely on to describe mental phenomena, what is often called Folk Psychology.

During the past twenty years more has been learned about the brain and how it works than ever before. Armed with all this new information cognitive researchers and neurologists have been attacking questions that until recently were reserved for philosophers: the questions of what consciousness really is (or the neural correlate of consciousness as some neurologists have dubbed it), the question of the mind-body relation and the problem of self-identity. This research has produced a host of books that are written to deal with the old mind-body problem in light of the new research.<sup>1</sup> Some of these are written by neurologists with no philosophical training and few philosophical aspirations – and are quite unconvincing when they try to draw philosophical implications of the research program – and some are written with a more philosophical program in mind. Nevertheless, one thing is certain: there is an interesting and important exchange of ideas, empirical discoveries, and theoretical frameworks between philosophy and neuroscience. Just as neurologists are eager to engage in philosophical speculation on the fundamental workings of the mental, philosophers have been eager to incorporate the findings of the neurosciences in their own projects of trying to elucidate how the mind works.

A landmark in this respect is Patricia Churchland's aptly named book *Neurophilosophy* (1986). This was perhaps the first serious attempt at integrating philosophical theory with the new findings of neurology. In this paper I will take a closer look at the interplay between these disciplines and try to point out some of the salient features making up the interface between them.

At first glance it might seem that theorizing is the task of philosophy, while making new empirical data available would be up to the neuroscientist. This, although true to a certain extent, is, I believe, a much too crude division of labor. A main idea in this paper is that

---

<sup>1</sup> See for instance, Schwartz and Begley (2002) *The Mind and the Brain*; LeDoux (2002), *The Synaptic Self*; Damasio (1999), *The Feeling of What Happens; Body and Emotion in the Making of Consciousness*; Koch (2004), *The Quest for Consciousness – a Neurobiological Approach*; Bennett and Hacker (2003), *Philosophical Foundations of Neuroscience*; Quartz and Sejnowski (2002) *Liars, Lovers, and Heroes – What the New Brain Science Reveals About How We Become Who We Are*.

empirical findings and conceptual tools are interdependent, modulating each other as they go along. This interdependency of data and theory – of the observed and the framework into which it is fitted – is nothing new; it has long been the accepted view in philosophy of science. The reason for stressing the point here is that when it comes to the mind and its relation to the brain, it might seem that philosophers and neuroscientist have been surprisingly slow to catch on.<sup>2</sup> The subjective states, our pains, beliefs, desires, emotions, reasons, hopes and fears, are readily available to scrutiny by introspection and seem to be outside the grasp of the neurosciences. When it comes to our inner lives we are playing on home field and no one is to tell us what, for instance, our feelings really are. Tempers run high when the idea that mental phenomena can be given a physical description is raised. There is, the argument runs, something special about the private and subjective feel of our mental workings and consciousness that can never be captured by a third, more or less objective, party such as the neurosciences. In the lingo of the philosopher: we seem to have privileged access to our own subjective states.

However, if the neurosciences are ready to embark upon the problem of the mind, are there any reasons for philosophers to pursue the same problems? Is it not time to let “hard” science take charge? Or to put it another way: is it possible to separate the philosophical questions from the scientific? In order to get a grip on these issues I will examine the following questions: how do the methods and goals of neuroscience and of philosophy relate to one another? Is the intentional vocabulary (and Folk Psychology) compatible with that of neuroscience or will a mature neuroscience do away with the intentional vocabulary, e.g. beliefs, desires etc? If not, what are the reasons for keeping the intentional vocabulary?

There are two reasons, I think, why philosophy cannot let neuroscience do the work alone, at least not at this stage. The first reason has to do with the maturity of cognitive science and neuroscience. Although we have made tremendous progress in brain imaging and scanning techniques, in cognitive research and computational models for linking the neural with the cognitive, there are still a lot of questions that cannot be settled by empirical research alone (and for sure, there are still questions that we have not yet begun to formulate). What is needed is a place where theories can be pitched against each other and where the logical structure of those can be measured; a place where no theory is sacred and where it is acceptable to “kick the lives of every governing paradigm, examine every sacred cow, and peer behind the curtains of every magic show”.<sup>3</sup> Philosophy provides such a tumbling place. Furthermore, the rapidly

---

<sup>2</sup> See, for instance Schwartz (2002)

<sup>3</sup> P.S. Churchland (2003), p 3-4

growing amount of data gathered by neuroscientists does not come with a ready theoretical framework in which it can be placed.

The second and more substantial reason has deeper roots, and is one I suspect will not be easily disposed of. It has to do with the way the physical vocabulary (in this case the neurophysical) resists matching up with the mental (the intentional vocabulary) in a systematic, one-to-one fashion. This has to do with the interests we have in the explanations those vocabularies provide. I suspect, with Davidson, Dennett and Rorty, that there is no way to translate features under description in one vocabulary into descriptions of the other without losing something – or quite a lot – in the translation. In this paper I will look into the possibility that two vocabularies, the mental and the physical, can exist side by side without opening for ontological or metaphysical dualism.

There seems to be a metaphysical bias in the current discussion of the mind-body relation: either the commonsense conception of the mental is rendered acceptable in physicalistic terms – the reductive alternative – or the conception is to be eliminated altogether – the eliminativist position. Both these are attempts at taking physicalism seriously, and the current view is that if you reject physicalism the only alternative is some sort of unpalatable mysticism. One of these mysticisms is, of course, Cartesian dualism, where the mental realm is ontologically distinct from the physical.

The approach taken in this paper is naturalistic through and through. A minimal definition of naturalism is that mental phenomena can be explained as part of the natural order and are empirically accessible features of the world. There are different ways of interpreting what this means, and the versions of naturalism vary from the strong realistic kind, to the more relaxed pragmatic stance. The strong realist is someone who holds that every property supervenes on physical phenomena and that all events (or objects) are identical to, composed of or reducible to physical events (or objects). At the other end of the spectrum the pragmatist holds that the mode of inquiry of the physical sciences will provide theoretical understanding of the world, to the extent that this sort of understanding can be achieved, and that this mode of inquiry may also turn out fruitful in the special sciences. The only requirement is that there is some explanatory link between the special sciences and the vocabulary of physics. This explanatory link need not be reductive; examples of such non-reductive explanatory links are the link between biology and physics and between geology and physics. In both these cases there is a link between the different explanations, but there is little hope in reducing, say, biology to physics.

In order to get a better grip on the tension between mental and physical descriptions I will start the discussion, in chapter two, with a closer look at what it means to explain the same object, the brain, as two different kinds of thing.

The pragmatic kind of naturalism endorsed in this paper draws on the idea that different vocabularies instantiate different interests and commitments. In the third chapter I take a closer look at this notion. The vocabulary-vocabulary as Brandom has dubbed it allows you to talk about mental properties and physical properties without assuming the ontological primacy of the one over the other. This gives access to naturalism “without taking the route through nomological or conceptual reduction.”<sup>4</sup> Since the idea of vocabularies is significant to this paper I will devote some time to explaining the central idea behind it and lay down some of its central virtues (and perhaps search for some flaws) and spell out how it stands in relation to more traditional approaches to the mind-body problem. As the idea is elaborated it will be clear that it helps to fend off some major metaphysical and ontological problems inherent in the more orthodox approaches to the mind.

I then proceed, in the fourth chapter, to evaluate a position that is intended to remove the tension by arguing that the intentional vocabulary is widely erroneous. This is the eliminative materialism of Paul and Patricia Churchland. They claim that the intentional vocabulary, or Folk Psychology as they call it, fail to meet up with the current findings of neuroscience and consequently should be eliminated. I assess their position by first stating what they mean by eliminativism and then examine five different objections against their position. The reason for examining these objections is to determine what opposing sides in the debate see as the important elements in the commonsensical and in the scientific description of our mental lives. In essence the Churchlands claim that the commonsensical framework of the mental will have to adapt according to the findings of neuroscience.

The purpose of scientific explanation will be discussed further in the fifth chapter. For this purpose I turn to Paul Griffiths’ new and more flexible account of natural kinds. Natural kinds play an important role in scientific explanations; they are involved in the central scientific features of induction and prediction. The function of natural kinds is traditionally taken to be “to cut nature at its joints” or to capture the real essence of the members making up the kind. Griffiths thinks, on the other hand, that the traditional view on natural kinds, with universal exceptionless laws, is too rigid to be of much help to science. He has developed a new account of natural kinds that is supposed to better capture the intended function of scientific terms. He

---

<sup>4</sup> Ramberg (forthcoming) section I

wants to leave the idea of natural kinds as the most fundamental category of nature and instead view them as non-arbitrary ways of classifying the subject matter under examination. He uses his findings to argue that the vernacular concept of emotion does not form a single natural kind and concludes that it has to be eliminated in favor of a better informed theory of emotion.

In the next chapter, the sixth, I compare and contrast the different reasons the Churchlands and Griffiths hold for arriving at the eliminativist conclusion. In this chapter I will also take a look at the normative role the intentional vocabulary plays, since this is something they seem to disagree about.

The final chapter summarizes some of the lessons learned, and offers some general observations on the pragmatic approach that I am endorsing.

## 2. Different Ways of Describing the Same Thing

*Robot:* With all your modern science, are you any closer to understanding the mystery of how a robot walks or talks?

*Professor Farnsworth:* Yes you idiot! The circuit diagram is right here on the inside of your case.

*Robot:* I choose to believe what I was programmed to believe!

- "The Honking", *Futurama*<sup>5</sup>

The two last decades of the twentieth century saw an explosion of means and tools for studying the brain in more detail. For the first time it is possible to measure brain activity of a person, by non-invasive means, while the subject is still awake and responding. New techniques as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) reveal which brain areas are activated when an individual performs a particular task; electroencephalographs (EEG) and magnetoencephalographs (MEG) provide detailed information about the time course of brain processing.

Before these tools were developed researchers had to meticulously note all behavioral deficits of a subject, wait for the post mortem autopsy, and then try to correlate the observed behavioral deficits with damaged brain tissue. Still, with only these crude methods of investigation some correlations between mental functions and the brain were possible to establish. Broca's work on persons with speech impairments due to head injury in the 1850's was perhaps the first respectable evidence for cortical localization of function. He was able to conclude that a specific region of the brain is responsible for the production of fluent speech.

With the new tools at hand cognitive psychologists, cognitive scientists, and neuroscientists have launched a special epistemic project: to capture mental and subjective phenomena in a scientific language. The idea is to develop systematic and well-supported models of the operations occurring in the brain as people are experiencing different emotions and as they are carrying out a variety of different cognitive tasks including categorization, problem solving, recognition, and recalling.

At this point it is possible to discern at least three different levels of explanations involved in a scientific theory of the mental.<sup>6</sup> The top level is where the evolutionary history of cognitive development is seen most explicitly. It should include such things as ecological pressure and cognitive improvement. It probably should also include an explanation of the role

---

<sup>5</sup> Futurama, season 3, Episode 2ACV18

<sup>6</sup> See Griffiths (1997), p. 221



cognition and consciousness has in culture. This is the level of task description and involves explanation of what a given trait does for the organism. Research at this level is supported by the increasing flow of data emerging from cognitive ethology. The intermediate level is the level of computations. This should include an account of representation and how information is processed to accomplish the task described at the previous level. The lowest level is the neuronal underpinning of mental and cognitive abilities. This is where the work of neuroscientist bears most weight.<sup>7</sup> This level is where the answer to the question of how computations are physically implemented should be formulated.

Taken as scientific programs, these levels of description do not seem to be at odds with each other. But with the realization that what is being explained here are mental phenomena – our subjective experiences, emotions, beliefs, and desires – and social and moral knowledge, a sneaking suspicion that something is missing may arise. How can a theory couched in scientific terms explain the myriad of subjective states, phenomenal experiences and moral knowledge? As Paul Churchland notes: “Inevitably, these sentiments will evoke discomfort in some readers, as if, by being located in the purely physical brain, social and moral knowledge were about to be devalued in some way.”<sup>8</sup> With the suspicion of devaluation follows the conviction that there is something special about the mental and about moral knowledge. This seems to be knowledge that cannot be explained by a biologically or physically based science.

Among cognitive scientists and neurologists, on the other hand, this epistemic project is based on a firm belief that mental phenomena *can* be given a naturalistic explanation. That is, they believe that the mental can be explained without resort to super-natural entities, and they usually deny that there is a metaphysical distinction between the mind and the brain. From their point of view, cognitive and mental properties are subject to empirical studies and biological explanations.

LeDoux assures us that “[b]rain circuits and psychological experiences are not different things, but rather, different ways of describing the same thing”<sup>9</sup>. This, of course, is not a new sentiment. Ryle wrote in *The Concept of Mind*:

“He is bodily active and he is mentally active, but he is not being synchronously active in two different ‘places’ or with two different ‘engines’. There is one activity, but it is one susceptible of and requiring more than one kind of explanatory description”<sup>10</sup>

---

<sup>7</sup> Some neurologists have called the work done at this level a search for the “neural correlate of consciousness” (NCC). See, for instance, F. Crick: “Why Neuroscience May be Able to Explain Consciousness” in *Scientific American* 2002 vol. 12 no. 1, and Metzinger (ed) 2002.

<sup>8</sup> P. M. Churchland (1995), p 124

<sup>9</sup> LeDoux (2002), p. 262

<sup>10</sup> Ryle (1949), pp. 50-51

At first glance it might seem trivial to say that what we are talking about are not two different objects, but rather two descriptions of the same object, but at the same time it seems to be at odds with the natural tendency to think that ontologically speaking the mental and the physical are different things. And this way of formulating the relation between mind and brain raises a couple of important questions. What does it mean to describe the same object in different ways? What counts as a description? And, “How, to come to the point, do we know where we have two ways of talking about the same thing [...] rather than descriptions of two different things?”<sup>11</sup> In this paper I will pursue, and sketch some possible answers to these questions.

If the difference lies at the explanatory level, it is most likely an epistemic problem rather than an ontological one. A pressing challenge will be to explain how to understand the relation between a scientific approach to the mental and other ways of learning about the mental (for instance, by introspection). The main philosophical task will be to articulate what the real differences between the explanations consist of and how they affect each other. The reasons against any hope for clean-cut intertheoretic reduction between the levels are many and substantial, but this does not mean that the different descriptive levels are independent. On the contrary it is vital that findings on one level should be reflected on the other levels. (And this has in fact happened several times during the “history of the brain” as when phrenologists had to abandon their work when it was found that the lumps on the skull did not match the functional roles of the cortex below the skull.)

One important part of this paper will be to articulate what is gained by keeping both levels of description – the neurological and the psychological – and give reasons for not opting for the eliminativist strategy; that is, I will try to show that the mental vocabulary, the commonsensical framework of the mental, plays an important role in describing human behavior, and that it might be indispensable. Another significant part will be to examine when keeping different descriptions of one domain adds something to the overall explanation, and when it just leads to confusion.

The full account of a human and her faculties involves not just a physical description, but also her relation to other humans, her place in the social nexus. By eradicating the intentional vocabulary an essential part of the description of a human being would disappear. A more appealing approach would perhaps be to try to reduce the mental description to a physical

---

<sup>11</sup> Rorty (1979), p. 29. Rorty does not seem to answer this question in *Philosophy and the Mirror of Nature*; it is not compatible with the pragmatic strategy he chooses to pursue.

one, but this is only feasible if there is a one-to-one match between the terms of the mental vocabulary and that of the physical. This seems highly unlikely since the patterns picked out by the different vocabularies reflect different interests, and the one-to-one match would imply that the interest spelled out in the mental vocabulary and the physical vocabulary should be identical. This, of course, is by no means true.<sup>12</sup> The mental vocabulary captures the macro-activity of humans in a way that, it seems, could never be captured by the fine-grained vocabulary of the physical sciences.

I start off with the idea of different explanations belonging to different vocabularies, and as such reflect divergent interests. The notion of different vocabularies serving different purposes is an elaboration of the distinction between the mental and the physical Donald Davidson argued for in “Mental Events”.<sup>13</sup> His strategy did provide a useful way of perceiving the relation between the mental and the physical, but as Ramberg (2004) argues, Davidson’s characterization of the mental is not suited to elucidate the current conceptual dynamics between the natural sciences and other ways of obtaining knowledge about our subjective and moral life (philosophy, psychology, ethics, anthropology, etc.). What Davidson does provide, though, is a rewarding way of looking at the relation between the mind and body as a difference in ways of describing the same thing, by referring to different vocabularies as the means to express different interests, without assuming that one of them has ontological primacy. This idea was further developed by Rorty (1979, 1989) and Brandom (2002, pp. 156-182). The vocabulary-vocabulary, as Brandom has dubbed this idea, is the subject of the next chapter.

---

<sup>12</sup> Not true at the present, at least. However, seeing as the development of vocabularies is contingent on our interests it is important to keep an open mind when it comes to revision of those vocabularies.

<sup>13</sup> Davidson (1970): “What does it mean to say that an event is mental or physical? One natural answer is that an event is physical if it is describable in a purely physical vocabulary, mental if describable in mental terms” (p. 210)

### 3. Rorty and Vocabularies

Almost as soon as I began to study philosophy, I was impressed by the way in which philosophical problems appeared, disappeared, or changed shape, as a result of new assumptions or vocabularies.

- Richard Rorty, *Philosophy and the mirror of Nature*<sup>14</sup>

The passage above is taken from the opening of *Philosophy and the Mirror of Nature*<sup>15</sup> by Richard Rorty. In this book he began to formulate his historicist, anti-Platonic world view. The idea that vocabularies are both non-compulsory and changeable was the driving force of his attack on the representationalist epistemology in this work, and was developed further in “The Contingency of Language” from *Contingency, Irony and Solidarity*.<sup>16</sup> Here he argued that there is no way to make the notion of truth dependent on something in the world corresponding to our description of that very thing. The argument runs in three steps: i) truth is a property of sentences; what is to be assessed as true or false has to be expressed in a sentence, ii) sentences have meaning only within a corpus of other sentences; a vocabulary or a language, iii) vocabularies are products of human activity. From these three premises Rorty concludes that truth itself is a product of human activity and not something “out there” awaiting our discovery, as had been the prevailing view throughout the history of Western philosophy. Reciting the pragmatist slogan that what makes no difference to practice should make no difference to philosophy, Rorty maintains that there is no essential distinction between justification and truth. The notion of truth and correspondence has to go since “[...] assessment of truth and assessment of justification are, when the question is about what I should believe now, the same activity.”<sup>17</sup>

What matters for the pragmatist, then, is not whether our beliefs correspond to some parts of reality, i.e. whether they “are true”, by reference to the world, but whether their role in our vocabularies makes a *practical* difference. That is, if you “drop the idea that some of our schemes are distinguished by such correspondence, as Davidson has, it seems natural to say, as Dewey and Wittgenstein did, that all our idioms are tools for coping with the world.”<sup>18</sup> Rorty wants to see the “classical” notion of truth removed from the philosophical discussion.

---

<sup>14</sup> Rorty (1979), p. xiii

<sup>15</sup> Rorty (1979)

<sup>16</sup> Rorty (1989), pp. 3-23

<sup>17</sup> Rorty (1998), p. 19

<sup>18</sup> Rorty (1997), p. 53

At the heart of the notion of vocabularies, our idioms for coping with the world, is the rejection of the Kantian idea of a linguistic practice that sharply distinguishes between languages, as structures of meaning, and theories, as structures of beliefs; by appeal to the concept of vocabularies there is no longer a way to make a substantial difference between meaning and belief.<sup>19</sup> For Rorty and the pragmatist there is no way to make a distinction between the language game we engage in and the world we try to capture in that game.<sup>20</sup>

It is possible to single out several different vocabularies that serve different functions. For our purpose the important vocabularies are that of the mind – I will also refer to it as the intentional vocabulary or the mental vocabulary and later also as Folk Psychology; all seem to pick out more or less the same sort of phenomena – and the physical vocabulary, which is roughly the vocabulary of scientific explanation and the physical sciences. The physical vocabulary strives to become an “ultimate vocabulary” where strong laws operate. The generalizations of the physical sciences must be lawlike to be projectible and to be used in induction. The mental (or intentional) vocabulary on the other hand is not, on the face of it, committed to a program of projectability and induction of the same kind as the physical vocabulary. It is intended to capture the meaningful behavior of rational agents and in this respect it is ultimately a teleological vocabulary. The intentional vocabulary incorporates a complex pattern of beliefs, desires, intentions, predictions of other people’s behavior, etc. This makes it the basis of our everyday understanding of human behavior. Nevertheless, there is a discussion raging over the status of the intentional vocabulary. The Churchlands, as we will see in chapter 4, argue that the intentional vocabulary, which they (and many others with them) refer to as Folk Psychology, is a theoretical vocabulary and as such should be held to the same standards as any other scientific vocabulary.

In this metaphilosophical context the traditional mind-body problem can be viewed as a tension between vocabularies and interests, rather than as a problem of our scientific practice or our ontological commitments (although the way we perceive a vocabulary closely reflects our ontological preferences). A better way of perceiving the different vocabularies and how they are interconnected would also reduce the tension between a “greening neuroscience and a

---

<sup>19</sup> That is, there is a difference between meaning and belief, but this is a pragmatic difference, not a principled or substantial one. Sometimes it is useful to “keep one factor constant” and talk about the other factor, in order to be able to get a clear idea of one over the other. The differences are, for the pragmatist, best viewed as temporary and local.

<sup>20</sup> See also Davidson (1984), pp. 183-198, “On the Very Idea of a Conceptual Scheme”. His argument is similar; he says that there is no way to make a distinction between scheme and content. This reasoning is also similar to the Sellarsian attack on the distinction between what is “given to the mind” and what is “added by the mind” (see Sellars, *Empiricism and the Philosophy of Mind* (1957), reprint (1997))

graying philosophy.”<sup>21</sup> Of course, to say this much is only to point at what I take to be a proper solution to the mind-body problem. In order to get a better grip on the problem and the proposed (dis-)solution I will try to flesh out the details of the vocabulary-vocabulary and relate it to the more orthodox views of dualism, eliminativism and reductionism. It might also be demanded that it be defended against claims of relativism. That is, it might be that someone might claim that “since vocabularies, and thus truths, are dependent on humans, and specifically the vocabularies they use, *anything* will be possible; it is just a matter of changing vocabularies”. This postmodernist claim, I will argue, is not implied by the vocabulary-vocabulary.

### ***The Vocabulary-Vocabulary***

The idea of the vocabulary-vocabulary is intended to provide new means “for organizing our thinking about our cognitive and practical activity as knowers and agents”.<sup>22</sup> As Whorf noted:

“[...] we cut nature up, organize it into concepts, and ascribe significances as we do largely because we are parties to an agreement to organize it in this way, an agreement that holds throughout our speech community and is codified in the patterns of our language”<sup>23</sup>

But it is not just that we are parties to an agreement, it is also an important part of our biological nature:

“[...] there is a deep connection between how the mind works and how we *perceive* the world works. We are pattern seeking animals, the descendants of hominids who were especially dexterous at making causal links between events in nature. The associations were real often enough that the ability became engrained in our neural architecture”<sup>24</sup>

Concepts and categories in this way sort out the patterns in the world and they in turn make up the central parts of vocabularies. Changes in those very concepts come about largely through interest conflicts, both when we are faced with new phenomena and when we face old phenomena in a new way. This natural tendency to group salient features into concepts and categories is captured by the pragmatic tool of vocabularies. The idea is that different vocabularies will pick out different patterns, and that what ultimately is considered as a pattern

---

<sup>21</sup> P. S. Churchland (2002), p. 3

<sup>22</sup> Brandom (2000), p. 157

<sup>23</sup> Whorf (1956), cited in Scientific American, April 2004

<sup>24</sup> Michael Shermer, in Scientific American, June 2003 (my italics).

in turn is guided by our interest in the world. We do not construct patterns, they are already there, but by using a vocabulary we “choose” which patterns we take to be salient. The application of a vocabulary makes certain patterns emerge, but at the same time patterns make us select a specific vocabulary. As a result, terminological choices guide those patterns that appear. But this does not imply that the vocabularies are chosen or progress at random. There are several constraints on the development of vocabularies.

Rorty views vocabularies as different narrations about who we are. Both vocabularies that typically consist of physical descriptions and causal regularities, and the vocabulary that consists of the intentional terms is important in making the full story of who we are. He does not see that one kind of vocabulary should have any (metaphysical or logical) precedence in capturing our subjective states. (Quine on the other hand has a bias for the scientific vocabularies. They are, in his view, best suited to capture the regularities of the world and our place in it.) The question then becomes whether the needs and interest of neuroscience match up with those expressed by the intentional vocabulary. Or to put it the other way around: are the needs and interests we articulate in the intentional vocabulary best answered and met by a mature neuroscience, or are there other things at stake here? Once again the question “what counts as an explanation of the relation between mind and body” seems to be left unanswered.

The vocabulary-vocabulary provides means to spell out how those interests are linked with the categories we apply when we describe and talk about the world. Incorporating metaphors, and turning them into integral parts of how we conceive the world is the central working method of the pragmatist. “To implement a metaphorical expression [...] is to work up a complex context of theoretical and empirical use for the expression in such a way that the metaphor is sent on its ways towards literalization”<sup>25</sup> The holistic core of pragmatism implies that theoretical and empirical considerations go hand in hand; there is little hope in trying to untie the conceptual knots prior to any empirical considerations have been made.

The intellectual development of a vocabulary will never reach a point where the vocabulary is freed of all metaphors and interests. This means that there is no such thing as a neutral vocabulary; as already mentioned the formation of a vocabulary is to a large degree dependent on our own interests and it is reasonable to claim that all human activity involves interest in one way or another.

It is important to stress this close relation between the vocabularies we use to frame certain problems and the interest we have in those very problems, or what might be said to

---

<sup>25</sup> Ramberg (1999), p. 64

constitute an answer to the question posed. This reflects the anti-essentialism which Quine argued for in “Two Dogmas of Empiricism”:<sup>26</sup> there is no way to separate what a person is talking about and what they are saying about it by reference to the essence of the object being discussed. Instead, metaphors and characteristics govern what kind of questions we are inclined to ask and what we regard as answers to those questions; in this way the choice and use of a vocabulary closely reflects our interest.

There are factors involved in the process of deciding which vocabulary currently is best suited for the explanatory needs for a domain, but they might not be transferable to other contexts, and they are not provided by the world. Our interests in different aspects of a domain and the way we talk about it govern the development of the vocabulary. The interplay between the ways we ask questions, how we ask them, and what we take to be answers to them are the important guidelines in the development of a vocabulary. However, this does not open the door for rampant relativism since the world is involved, too. It does not provide criteria for choice of vocabulary, but what sort of criteria we apply will depend on how it is useful for us to engage the world. And that does depend on how the world actually is. Not every kind of vocabulary will be suitable in any given situation; objectivity is not exchanged with subjectivity. “The moral is not that objective criteria for choice of vocabulary are to be replaced with subjective criteria, reason with will or feeling. It is rather that the notion of criteria and choice (including that of “arbitrary” choice) are no longer in point when it comes to changes from one language game to another”<sup>27</sup>

First, we are not at liberty to change vocabulary; secondly, there are mechanisms at play when a vocabulary changes. The idea of changing vocabulary at all depends on the scope of the vocabulary. That is, we are at the mercy of our concepts, in some sense, but not completely, and less so at the delineated, theoretical end of the vocabulary spectrum. It depends on the scope – adequately defined highly theoretical vocabularies can be treated pretty much as tools in this respect. We can choose whether to pick out some referent by means of spatial coordinates, or by means of a qualitative description. But we cannot randomly substitute one word with another and assume that what we say when we use the new word will make sense to others. Consider this famous exchange between Alice and Humpty Dumpty:

“And only *one* for birthday presents, you know. There’s glory for you!”  
“I don’t know what you mean by ‘glory’,” Alice said.

---

<sup>26</sup> Quine (1953), pp. 20-56

<sup>27</sup> Rorty (1989), p. 6



Humpty Dumpty smiled contemptuously. “Of course you don’t – till I tell you. I meant ‘there’s a knock down argument for you!’”

“But ‘glory’ doesn’t mean ‘a nice knockdown argument’,” Alice objected.

“When *I* use a word”, Humpty Dumpty said, in rather a scornful tone, “it means just what I choose it to mean – neither more nor less.”

“The question is”, said Alice, “whether you *can* make word mean so many different things.”

“The question is,” said Humpty Dumpty, “which is to be master – that’s all.”<sup>28</sup>

The point is that no one can proclaim themselves master of the use of a word. For the pragmatist (who takes her lesson from the later Wittgenstein) to grasp the content of a concept is mastery of the use of the corresponding word in a language or in a vocabulary. Concepts are thus holistically constrained by the vocabulary they are part of. The pragmatist in this manner is doubtful that anyone can grasp the content of a single word in isolation; it is only by the framework of a vocabulary that a concept gets its meaning (i.e. it would seem impossible to entertain the concept ‘DOG’ without knowing how to make use of the word in conjunction with a lot of other words – both dog-related (as in ‘dogs bark’) and those not directly related to dogs – and to utilize it in inferences assessable as correct or incorrect). But the pragmatist takes the holistic constraint to mean that there is no way to separate the logical, or conceptual, part of a vocabulary from the empirical part. There is, in addition, some sort of normative constraint on the words and the vocabularies we employ; to engage in a language-game is to share a common backdrop of commitments. And a person not willing to participate in this game, who insists on playing by his own rules, will not be considered rational.<sup>29</sup>

At the heart of the pragmatic notion of vocabularies is a wish to reformulate and transform the way we speak about the world. New metaphors and new ways of using words are chief methods in achieving such changes. But, “the rhetorical transformation sought by the pragmatist philosopher is not, it is crucial to stress, an attempt to change the world by wishing it different, not a species of an idealism which mistakes linguistic change for substantive change in the world.”<sup>30</sup>

However, changes in the way we speak is not the same as affecting how the world *is*, but rather “[...] what is affected, when the pragmatist succeeds, are the ways in which features of the world are salient to us and so the nature of patterns of the responses which events in the world elicit from us. In a word; what is changed is a complex subset of our causal dispositions.”<sup>31</sup> A myriad of contingent, causal factors can contribute, some of them mundane –

---

<sup>28</sup> Carroll, *Through the Looking-glass*, (1939, reprinted 1988)

<sup>29</sup> See Krabberød (2002) for a further discussion of this point.

<sup>30</sup> Ramberg (1999), p. 66

<sup>31</sup> *ibid.*, p. 66

like having “gradually lost the habit of using certain words and gradually acquired the habit of using others.”<sup>32</sup> But often these contingent, causal factors are anything but mundane. They provide an important link between the objects being described and the vocabulary in which we try to describe them. We are constrained causally by the world, and this has consequences for how our vocabularies, scientific or not, develop and evolve

Changing the way we talk does in one sense change the world, but in a very limited way. The last point is because ontologically speaking we have to say with Quine: “A curious thing about the ontological problem is its simplicity. It can be put in three Anglo-Saxon monosyllables: ‘What is there?’ It can be answered, moreover, in a word – ‘Everything’ – and everyone will accept this answer as true”<sup>33</sup> In other words, the world is as it is. Theories come and go and they shape the way we perceive the world and what kind of patterns and categorizations we pay heed to, but by and large the world remains the same. What might change is what is counted as “natural”, what kind of patterns we find to be salient. It is important to stress that those patterns are not something we invent, they are already there, but new ways of talking about them might change how we view them and which ones we rely on to conceive of our practical options, our actions (thus vocabularies will effect the world, in so far as we act on it differently).

### ***Reduction and Ontology***

The pragmatic approach endorsed in this paper is intended to support naturalism without conceptual reduction. This type of post-Quinean interpretivist strategy claims to mark out a fourth possibility in addition to traditional reductionism, elimination and dualism. The main idea is that mental processes can (and ultimately will) be naturalized; i.e. the apparent metaphysical gap some see between the vocabulary that describes us as intentional beings, with beliefs and desires, on one side, and the physical vocabulary that is intended to capture the causal regularities of the physical world, on the other, will in one way or another be accounted for. The challenge here is to meet the demands of naturalism – i.e. the notion that mental phenomena can be explained as part of the natural order and are empirically accessible features of the world – and still be able to show that we can keep both the mental and the physical descriptions without resorting to dualism. This is to be done without opting for the eliminativist or the reductionist strategy. The naturalistic pragmatist does not see the gap

---

<sup>32</sup> Rorty (1989), p. 6

<sup>33</sup> Quine (1961, p. 1)

between the mental and the physical vocabulary as one that is ontological in kind. Consequently, there is not much gained from conceptual reduction, since the motivating force of reductionism is to simplify the ontological catalogue of the world.<sup>34</sup>

Since the idea of an ontological gap is left behind, the pragmatist treats the gap between the mental and the physical as a malfunction in our philosophical approach to the problem in question. By revision of the conceptual structures we rely on to maintain the impression of a metaphysical gap the pragmatist suggests that the gap will close itself. We are not free to revise the concepts away according to our own inclinations; we are at least constrained by interaction with the world. In the following chapters I will take a closer look at what revision consists of.

The common way to think of the relation between folk psychological and scientific vocabularies is to assume that on some level of analysis there is an ontological fact that needs to be resolved, and that this fact will be determined by the success or failure of the reductive or eliminative proposals. This ontological fact is supposed to validate, or invalidate, the use of folk psychological terms for scientific usage. However, as Ramberg<sup>35</sup> points out, this argument is reversed relative to the pragmatist's interests. The philosophical legitimacy reduction is thought to provide has little to do with the naturalization of the vocabulary of intentionality. The reason for this is that the success of the folk psychological categories is irrelevant to the ontological consideration at the heart of the arguments for reduction. It is possible to think methodologically regarding the different levels of explanation without burdening them with ontological preferences. And, as remarked earlier, the reductive proposal is intended to capture the causal mechanism of the macro theory as a function of the micro theory, but it is impossible to single out an ideal level where causality can be observed directly.

Folk-ontological convictions burden the traditional idea of physicalism with a dichotomous vocabulary to the extent that naturalized pragmatism has trouble spelling out its anti-dualistic commitments in relation to it. Most of the debate on the mind-body problem in contemporary philosophy has centered on the idea that there is a question to be settled and that the weight ultimately will fall down on one of two sides; either there is something ontological distinct about the mind, thus confirming some sort of dualism, or there really is only one ontological kind, the physical. To point to the mind-body problem as a puzzle that will eventually evaporate if we rid ourselves of our ontological intuitions might not be very reassuring for the eliminativist, the reductionist, or the dualist. Both the eliminativist and the

---

<sup>34</sup> For the philosopher, that is, the scientist is not moved by a desire to simplify the ontological catalogue. His main motivation is explanatory force and power. Ontological purity is simply a bonus for the scientist.

<sup>35</sup> Ramberg (forthcoming)

reductionist might claim that the naturalistic pragmatist in this way is trying to eat his cake and have it too (the dualist will probably insist that there are actually two cakes).

The naturalistic pragmatic approach, in contrast, adopts Davidson's anomal monism<sup>36</sup>, which Ramberg describes as a Bland Monism. "It is monistic, because it denies the dualist's thought that there are two ontological kinds; mental and physical. It is bland in a somewhat peculiar sense; it also denies the reductivist or eliminativist thought that there is **one** ontological kind of a sort to which our various ways of talking may stand in questionable relationship."<sup>37</sup>

This naturalized conception of the mind will not imply that all mental terms will be reduced to neurobiology. Following Davidson it might be reasonably claimed that there is no way to make that sort of reduction.<sup>38</sup> Irreducibility in this context does not constitute a metaphysical gap, or some deep ontological fact, but reflects the divergent human interests that vocabularies serve.

That means that the pragmatist view I am sketching disagrees with Searle when he claims that: "the argument for the irreducibility of consciousness is not epistemic; it is about how things are in the world. It is about ontology. There are different ways of spelling it out, but the fundamental point of each is the same: the sheer qualitative feel of pain is very different feature of the brain from the pattern of neuron firings that cause the pain."<sup>39</sup> We can agree with Searle that there is a "qualitative feel of pain", but whether this adds anything to the ontological catalog is the important question. The view I am endorsing claims that the differences between the qualitative feel of pain and the neuronal underpinnings of it are epistemological and not ontological.

In the chapters that follow I will explore ways of interpreting the difference between a description of a mental state couched in neuroscientific terms and a description of the same state in a folk psychological (or intentional) terms that does not assume that there is a ontological point to be settled, but still is sensitive to the idea that "the neuron firings cause the feeling, but they are not the same thing as the feeling."<sup>40</sup>

The idea is to keep a distinction between descriptions of the mental and descriptions of the neuronal underpinnings of these mental phenomena, without resorting to dualism. To be able to do this – both naturalize and keep a mental-physical distinction – it has to be shown how

---

<sup>36</sup> Davidson "Mental Events", (1970)

<sup>37</sup> Ramberg (forthcoming), section III

<sup>38</sup> Davidson (1970), Quine makes the same point, although the other way round: "The physical conceptual scheme simplifies our account of experience because of the way myriad scattered sense events come to be associated with single so-called objects; still there is no likelihood that each sentence about physical objects can actually be translated, however deviously and complexly, into the phenomenalist language" (1953, pp. 17-18)

<sup>39</sup> Searle (1997), p. 31

<sup>40</sup> *ibid.* p. 30

different vocabularies explain different interest and how those interests are related, and why that precludes reduction of the intentional vocabulary.

In the next chapter I will turn to Paul and Patricia Churchland's idea of eliminative materialism. They argue that the intentional vocabulary should be eliminated in favor of the framework of better informed neuroscience.

## 4. The Intentional Vocabulary and Folk Psychology: the Churchlands and Eliminative Materialism

All the key words in this explanation, by the way, are totally misleading due to the everyday quirks of language.

- Don DeLillo, *Ratner's Star*

There are many similarities between what I have called the intentional vocabulary and what Paul and Patricia Churchland<sup>41</sup> refer to as Folk Psychology; both seem to be in the same business and they go about it in the same way: explaining and predicting human behavior. I shall in this section assume that the intentional vocabulary is roughly the same as Folk Psychology (hereafter FP), the rough and ready framework that, according to the Churchlands, “denotes the prescientific, commonsense conceptual framework that all normally socialized humans deploy in order to comprehend, predict, explain, and manipulate the behavior of humans and the higher animals. This framework includes concepts such as *belief, desire, pain, pleasure, love, hate, joy, fear, suspicion, memory, recognition, anger, sympathy, intention*, and so forth. It embodies our baseline understanding of the cognitive, affective, and purposive nature of people”,<sup>42</sup> that is, the “rough-hewn set of concepts, generalizations, and rules of thumb we all standardly use in explaining and predicting human behavior.”<sup>43</sup> In a word, it is our most important tool when it comes to guiding ourselves in social situations, and when trying to make sense of other people’s behavior. This framework, they argue, is so widely misguided that it should be displaced by a better theory.

There is a resemblance between the Churchlands’ desire to reformulate FP, in light of the fact that human knowledge is speculative and provisional and the questions Rorty raised against traditional epistemology in *Philosophy and the Mirror of Nature*. But the Churchlands arrive at a different conclusion than what I depicted in the last chapter. They do not view the mind-body situation as one that is due to different descriptions of the same object; rather they see the folk psychological terms as explanatorily impotent and argue that they should be removed from a scientific account of the mental.

---

<sup>41</sup> I will in what follows assume that they both share the view on eliminative materialism and the motivating factors for it and refer to them, taken together, as the Churchlands. The exception is when a text or a quote made by one of them is being discussed.

<sup>42</sup> Churchland & Churchland (1998), p. 3

<sup>43</sup> P.S. Churchland (1986), p. 299

Now, the idea that human knowledge is speculative and provisional is something that is widespread in philosophy, but as P.M. Churchland points out, it is “sorely tested when the question at issue is the possible displacement of our familiar *self*-conception – a conception that portrays each human as a self-conscious rational economy of propositional attitudes.”<sup>44</sup> There is little room for revision in light of current neurological development in the Folk Psychological framework.

The eliminative materialist, as exemplified by the Churchlands, is someone who both is able to see that all human knowledge is speculative and provisional, and is open to a reformulation of the story we tell about ourselves and our place in the world. This means, in effect, that the eliminativist shares the view on conceptual frameworks with the pragmatist; frameworks are vulnerable to revision and overthrow. Even the conceptual scheme embodying the notion of rationality and the related conception of ourselves as folk-psychological agents, may be subject to such revision.

The Churchlands have been attacking FP for a long time and claim that its current epistemic situation resembles more and more the earlier situation of phlogiston, caloric theory, alchemy, etc. That is, although FP is seemingly capable of explaining some of the features in its domain it is widely misguided and in reality explanatorily impotent: “our commonsense conception of psychological phenomena constitutes a radically false theory, a theory so fundamentally defective that both the principles and the ontology of that theory will eventually be displaced, rather than smoothly reduced, by completed neuroscience”.<sup>45</sup> There are at least two symptoms that according to the Churchlands show the failure of FP. First, FP seems to have too great a failure rate when it comes to explaining, predicting and manipulating humans. Secondly, it has proven difficult to find even a rough disjunctive reduction from FP to current neuroscience.

In order to see why this leads to the elimination of FP I will now take a closer look at what has been the Churchlands’ main concerns for the past twenty or thirty years, and why they seem to make such a disturbance with their eliminative materialism. What is it they want to see eliminated, why, and what, if any, are the consequences of elimination?

### ***Eliminative Materialism***

According to P.S. Churchland an eliminativist is someone who holds:

---

<sup>44</sup> *ibid.* p 25

<sup>45</sup> P. M. Churchland (1981), p. 391

1. that folk psychology is a theory;
2. that it is a theory whose inadequacies entail that it must eventually be substantially revised or replaced outright (hence “eliminative”); and
3. that what will ultimately replace folk psychology will be the conceptual framework of a matured neuroscience (hence “materialism”).<sup>46</sup>

The motivating factor for the Churchlands’ thesis of eliminativism is that they see the importance of making FP stand up to empirical obligations. Folk Psychology, for the Churchlands, is a theory – “a systematic, corrigible, speculative *theory*”<sup>47</sup> – and as such must be able to answer to the empirical pressure imposed on it, but it fails to meet up with the criteria that an empirical theory should meet. It is both inaccurate and fails to explain some of the key issues one would want a theory on the intentional to elucidate, namely mental illness, sleep, creativity, memory, intelligence differences, what understanding amounts to, and the many forms of learning, to mention but a few. If FP is an empirical theory one would like it to be able to explain these. Therefore, the Churchlands have argued that there is a need to reformulate what it is to learn, understand, etc.

This means that Folk Psychology, the commonsensical framework of ordinary mental vocabulary and the common-sense understanding of the mental, will be replaced by the concepts and commitments of a mature neuroscience, at least whenever we talk seriously and literally about what we call the mind. There will be no need for the constructs of Folk Psychology once the neurosciences have reached the goal of making sense of the brain and how it functions. The Churchlands lend support to this claim by using historical analogies. By equating the posits of Folk Psychology with phlogiston, ether, witchcraft, and alchemy they try to show that we often use concepts that are useless, and explanatorily impotent, and will ultimately vanish from scientific discourse once science has caught up with what is happening in the world. In this sense the terms of Folk Psychology do not correspond to any items that will be part of a *theory* of the mental.

This gives the Churchlands reason to claim that FP is widely erroneous and explanatorily impotent. They maintain that FP is a theory and that its success should be measured by the way it stands up in comparison to other scientific theories. The intentional idioms, they claim, are isolated and the history of FP is stagnant. It is important for them to

---

<sup>46</sup> P.S. Churchland (1986), p. 396

<sup>47</sup> P. M. Churchland (1981), p. 395



insist that FP is a theory, because then they can pitch it against other scientific and often flawed theories, e.g. the theory of phlogiston, the theory of caloric fluids, and the theory of alchemy. If, on the other hand, they are wrong, i.e. that FP does not have a theory like nature, or if it can be show that the commitments of FP differ radically from those of scientific vocabularies, it would seem possible to defend FP against the attacks.

### ***Dire Consequences***

It is important to point out that the debate over the elimination of FP is not a question of the *details* of the everyday theory, but rather the complete framework. Whether or not a part of the theory can be salvaged is not up for discussion. What is at stake is the framework taken as a whole. Since some, amongst them Fodor, think that it is more than just the success or failure of a scientific program which is threatened, the stakes seem very high. After stating that there is nothing that could convince us that FP is wrong Fodor goes on to say: “Which is just as well, because if commonsense intentional psychology really were to collapse, that would be, beyond comparison, the greatest intellectual catastrophe in the history of our species; if we’re that wrong about the mind, then that’s the wrongest we’ve ever been about anything”<sup>48</sup> But he is not alone in thinking that the elimination of our everyday understanding of the mental would have dire consequences.

Baker (1987) illustrates how much some take to be in jeopardy when it comes to the question of the legitimacy of our common-sense conception of the mental. She thinks that practical reasoning, language, morality, rationality, truth, “indeed, cognitive virtue in general”<sup>49</sup> is thrown out the window if the eliminativist is right. And she provides a list of what she takes to be the consequences of the eliminative proposal:<sup>50</sup>

1. Consequences for practical reasoning:
  - a) We will lose the ability to make predictions of other peoples behavior.
  - b) Ordinary interaction among people would seem mysterious
  - c) Behavior could never go wrong; no one could ever do anything unintentionally
2. Consequences for moral and legal practice:
  - d) Explanations of behavior would be wrong
  - e) There would be no distinction between lying and honest mistake

---

<sup>48</sup> Fodor (1987), p xii

<sup>49</sup> Baker (1987), p, 129

<sup>50</sup> See Baker (1987, pp. 130-134)

- f) Moral judgments would be false or senseless
  - g) Nothing would ever matter to anyone
3. Consequences for linguistic practice:
- h) Why we say the things we do would be mysterious
  - i) Uttering of truths would be miraculous
  - j) Reports of deliberation and decision would be false
  - k) What ones does would be totally unrelated to what one reports that she thinks she is doing
4. Consequences for Psychology:
- l) Applied psychology would be bogus
  - m) Psychological explanations would evaporate

If Baker is correct then it is obvious that much hinges on whether eliminativism is true or not. However, I think that it is possible to give eliminativism an interpretation that does not imply all the points on the list. In fact, this interpretation does not mean that the skeptic will have to be “prepared to swallow all these consequences”<sup>51</sup> I maintain that the consequences presented by Bakes do not follow from a pragmatic reading of the Churchlands.

### ***Objecting against Eliminative Materialism***

The Churchlands see two kinds of objections made against their idea of eliminative materialism. One type of attack is made by those who accept the fact that human knowledge is speculative and provisional, but who, for some reason, contests the idea that this also applies to our familiar self-conception. Examples of such might be Neo-Platonists (or indeed Kantians) who argue for the inevitability of the framework features of FP, or Cartesians who holds that the epistemological capacity of introspection is one that is essentially incorrigible, and since these are the essential part of Folk Psychology no revision of the vocabulary will be needed (it might indeed be impossible). These are objections that contest the idea that humans are tentative in their epistemological speculations. This is a battle to be fought at a meta-philosophical level and does not directly affect the Churchlands’ argument for the elimination of the propositional attitudes and folk psychological account of human cognition, as it must be settled prior to the arguments the Churchlands draw on. Another kind of objection is made by those who accept our cognitive shortcomings, and the theoretical nature of Folk Psychology,

---

<sup>51</sup> Baker (1987, p. 132)

while still denying the epistemological evaluation of the framework of Folk Psychology that naturally goes with it. These are objections raised within the framework of revisionary epistemology. They directly engage the arguments brought by the Churchlands, since they are playing on the same field, so to speak. The Churchlands call objections raised against elimination within this framework for “inauthentic”. These are called inauthentic since they are, on P.M. Churchlands view, unfaithful to the contemporary epistemological perspective, that all human knowledge is speculative and subject for revision, even our knowledge about our familiar self-conception. Those who hold the “inauthentic” objections claim that conceptual frameworks are subject of revision, but not the commonsensical framework of FP. Churchland himself devotes considerably more time to these last, “inauthentic”, objections than to the objections raised at a meta-philosophical level. What do those objections amount to?

In the essay “Evaluating Our Self-Conception”<sup>52</sup>, P. M. Churchland identifies five such “inauthentic” objections. In the next few pages I will go through these objections and try to see if it is possible to discern a pattern, both in the objections and in the way Churchland responds to the objections. The hope is that if a pattern can be distinguished, by taking a step back and doing some metaphilosophical reflection it will be possible to articulate the underlying issue at stake.

The first objection is based on the notion that eliminative materialism relies on the framework of FP to refute FP; the second objection does not see what could possibly count as evidence showing that FP is radically wrong; the third objection maintains that FP is used for a vast range of different non-scientific purposes and should not be evaluated as a scientific theory; the fourth objection is based on the idea that there are no existing alternatives. The final objection is based on the claim that FP consists of *functional, or artificial, kinds* (kinds created by human choice, defined, at least in part, in terms of a function that bears on a human purpose like ‘chair’, ‘cycle’ and ‘bungalow’), not *natural kinds* (kinds that usually figure in scientific theories that are intended to capture some underlying essence or sameness, examples are H<sub>2</sub>O, gold and the like). This, allegedly, makes it “invulnerable” to empirical revolutions in neuroscience (like our conception of say iron fences does not change radically by empirical revolutions in metallurgy: what makes a fence largely remains unchanged. How to make fences, on the other hand, might change).

---

<sup>52</sup> Churchland & Churchland (1998, p. 25-38)

### 1. *The first objection: Eliminative materialism is a non-starter*

The first objection rests on purely a priori convictions and does not involve any empirical considerations. Eliminative materialism must, the argument goes, suppose the concepts it will eliminate. To put it crudely, the objection seems to rest on the idea that to eliminate FP someone must *believe* in the elimination of the concepts therein, but since ‘believe’ is a central concept of FP it cannot be eliminated (like someone saying “I believe there are no beliefs”).

This kind of objection rests on the fact that the logical conclusion from Q to not-Q is illegitimate. But this does not discourage the eliminative materialist (or the pragmatist, for that matter). She does not see that there should be anything blocking the displacement of Q based on Q coupled with further empirical evidence, and the eliminativist does insist on the empirical nature of FP (or Q in this instance). The argument runs thus: assume Q, argue from Q and some empirical evidence to not-Q and then, by *reductio ad absurdum*, conclude that not-Q.<sup>53</sup>

The non-starter objection must presuppose what it is trying to prove: that beliefs and other similar intentional states exist. Specifically, it relies on the idea that beliefs exist in the form that they take in traditional propositional attitudes philosophy. But this simply begs the question. The Churchlands are trying to provide an *alternative* to the old structure of Folk Psychology, and to assume that they have to do use the old framework when they construct the alternative is just foolish. P.S. Churchland (1986) notes that:

What the eliminativist is fumbling to say is that folk psychology is seriously inadequate as a theory. Now within the confines of that very theoretical framework we are bound to describe the eliminativist as believing there are no beliefs; however, this is not because folk psychology is *bound to be true*, but only because we are confined within the framework the eliminativist wishes to criticize and no alternative framework is available. If the eliminativist is correct in his criticisms, and if the old framework is revised and replaced, then by using the new vocabulary the eliminativist’s criticisms could be restated with greater sophistication and with no danger of pragmatic contradiction. [...] It would be foolish to suppose folk psychology must be true because at this stage of science to criticize it implies using it. All this shows is that folk psychology is the only theory available now. (pp. 397-398)

She then proceeds by showing that a similar non-starter-argument could be made in defense of vitalism; the anti-vitalist claims that there is no such thing as vital spirit, but the claim is self-refuting. If it were to be true the person making the claim would have no vital spirit, consequently she would be dead and unable to make the claim.

---

<sup>53</sup> The logical structure is  $(Q \rightarrow \neg Q)$ , which is equivalent to  $(\neg Q \vee \neg Q)$ , which then in turn reduces straightforwardly to  $(\neg Q)$ . It is, of course, the first implication which is in question.

Baker's list, mentioned above, is also based on the notion that eliminativism entails that there is no such thing as a contentful intentional state. From this she concludes that, according to the eliminativist, "all attributions of believes, desires, hopes wishes, fantasies, intentions, and so on are false."<sup>54</sup> She goes on to say: "if mental states lack content, one cannot assume anything; one cannot even suppose that one can."<sup>55</sup> However, by using the reasoning provided above it is possible to respond to Bakers criticism.

For instance, in the corollary of 1a) she says:

Suppose that I dialed your phone number and said, 'Would you join us for dinner at our house on Saturday at 7:00?' You replied, "Yes." On Saturday, I act in the way I should act if I believed that you were coming to dinner. But if neither of us had any beliefs, intentions, or other states attributed via 'that'-clauses, it would be amazing if I actually prepared dinner for you and if you actually showed up.<sup>56</sup>

Compare this with the following defense of vitalism:

Suppose that I dialed your phone number and said, 'Would you join us for dinner at our house on Saturday at 7:00?' You replied, "Yes." On Saturday, I act in the way I should act if I you were coming to dinner. But if neither of us *were alive since there is no such thing as vital spirit*, it would be amazing if I actually prepared dinner for you and if you actually showed up.

This last "defense" of vitalism just sounds silly, of course. The defender of the intentional states would say that the analogy misses the point. It is, she probably would claim, the *content* of the intentional states which is at stake. And as long as the eliminativist fails to give an explanation of what goes on in the first example, without resort to intentional terms, "his profession of skepticism about the common-sense conception is just idle play."<sup>57</sup> The vitalist would similarly say to his skeptic that as long as no one can give an adequate explanation of 'being alive' in non-vital-spirit terms, his work is idle play. We know, by now, that a better alternative to vitalism was provided by molecular biology and metabolic chemistry. In the same way the eliminativist thinks that a sufficient, non-intentional<sup>58</sup> explanation will be provided by a

---

<sup>54</sup> Baker (1987), p. 129

<sup>55</sup> *ibid.* p. 129

<sup>56</sup> *ibid.* p. 130

<sup>57</sup> *ibid.* p. 133

<sup>58</sup> A qualification might be in order here: what is meant by intentionality is the kind that the propositional attitude theorist claims to capture. There might still be room for intentionality in the new and better informed version of the mind that the eliminativist ask for, but the *description* of the cognitive and mental process and of what counts as intentionality will be radically different form the more traditional approach since the complete framework of the psychological and mentalistic terms will be revised.

matured neuroscience. Some cannot imagine what an alternative account of mental content might look like. This is partly why the eliminativist has trouble getting her message across. She might not have the complete answer ready at the present, but she has high hopes for Neural Networks, as we will see later.

If this objection were correct then it would be possible to argue that there will always be something preventing the displacement of a conceptual framework of cognition since “the same awkwardness – formulating a rejection of a framework within the framework itself – will arise *whatever* conception of cognition one happens to be using.”<sup>59</sup> In fact, it is possible to generalize this objection and say that the same argument will block any radical overthrow of a framework that enjoys the irrelevant privilege of being the one currently in use.

## *2. The second objection: What could possibly falsify Folk Psychology?*

The second objection questions what could count as evidence against the success of FP. Traditionally there are four prototypical kinds of pressures that count against the success of a theory, and the Churchlands claim all four already are exerting pressure on FP: first, there is the case of chronic poor performance (a historical example would be Ptolemaic astronomy). Folk Psychology is not a perfect theory, although it does well on average its success rate is far from 100 percent. Actually the Churchlands claim that the success rate is much closer to zero percent. FP’s success seems to rely on the fact that the subject it is applied to is a normal, language-using, adult human. It has proven to be extremely inefficient on infants and those with serious mental illnesses. If this is true it places serious pressure on FP should a better theory surface. Which leads us to the second kind: pressure is also applied from closely neighboring, well-performing theories with which the original theory does not seem compatible (a historical example would be vitalism in relation to metabolic chemistry and molecular biology). Computational neuroscience, and neural networks, provide new kinds of understanding of how the brain (or the mind) works; vector-to-vector transformation over neuronal representation does not seem to go well with the propositional attitudes, e.g. beliefs and desire, traditionally taken to be the central concepts of FP.<sup>60</sup> Thirdly, there is the case of poor (or no) extension to continuous domains. P.M. Churchland notes that FP “shows no sign of being smoothly

---

<sup>59</sup> Churchland and Churchland (1998) p. 30

<sup>60</sup> There are different ways of interpreting the status of the propositional attitudes. For the realistic physicalist, e.g. Fodor (1987, 1990, 1995), Fodor and Lepore (1992), the distributed states captured by vectors seem unable to capture the intentionality of mental states. But one can be realistic about propositional attitudes without thinking they need to turn out to be names of physical states. The Churchlands argument seems to have most bite against the strong realist.

integrable with the emerging synthesis of several physical, chemical, biological, physiological, and neurocomputational sciences. Since active coherence with the rest of what we presume to know is a central measure of credibility for any theory, FP's emerging wallflower status bodes ill for its status."<sup>61</sup>

The fourth and final kind is refuting empirical evidence designed to discriminate between competing alternatives (for instance Eddington's eclipse expedition). It is hard to see how an experiment could be constructed as to discriminate between, say FP and Neural Networks at the present time, but it could be done. For instance, in order to confirm the hypothesis that human thought is based upon inferences over propositional attitudes, it might be possible to build an artifact that operates in that way. The test then is to see if the system thus constructed can display any similarities with our cognitive capacity in real time. This would then count as evidence towards FP as a correct theory. This is typically what those in the classical artificial intelligence community<sup>62</sup> have been trying to do for years, but simulating cognitive capacities in computers, in this way, has proven exceedingly hard. Thus far the results are indecisive, but there are few indications that they will solve the problem in the near future.

*3. The third objection: Folk Psychology is used for a wide range of non-scientific purposes. Folk Psychology is not a scientific theory. In fact, it is not a theory at all.*

The third objection is raised by those who claim that the basic function of Folk Psychology is not just to explain and predict human behavior, but that it is used in a vast range of non-scientific purposes, such as to warn, promise, congratulate, etc. It serves, in other words, different purposes than scientific theories characteristically do. From this they conclude that FP should not be evaluated as a representative scientific *theory*. Stich notes that “[o]rdinary folk certainly don't take themselves to be invoking a theory when they use intentional terms to explain people's behavior.”<sup>63</sup> And he goes on to state that “Still less do they think they are using a theory when they report their own beliefs and desires.”<sup>64</sup> Typically, this conception portrays scientific theories as an abstract construction of propositions far removed from practical life. Folk Psychology on the other hand is characteristically constructed to spell out

---

<sup>61</sup> Churchland & Churchland (1998) p.8

<sup>62</sup> See Franklin (1995), and

<sup>63</sup> Stich (1993), p. 3

<sup>64</sup> *ibid.* p. 3

the practical aspects of daily life as well as the mental and cognitive processes underlying those practices.

Of course, this all depends on what a theory is, and whether or not a person has to be aware that she is using one before it can be rightfully called a theory. The Churchlands maintain that this objection rests on a too narrow and cartoonist conception of what a theory is. In the following I will take a brief look at an intuitive account of what a theory is. For the purpose of this discussion it does not matter whether it is possible to give a definition of a theory, it will suffice to outline some key points that almost everyone can agree on and then see if FP can be called a theory according to those key points.

Theory does play an intimate part of ordinary life. Think of the carpenter and his internalized theory of geometry, the jazz musician and music theory, the physician and medical theory, to mention but a few. An important idea pointed out, by amongst others Kuhn, is that when the scientist learns his trade she typically does not just learn a set of propositions and the logical connections between those. Learning a theory consists much less of memorizing a doctrine and much more of slowly grasping a host of miscellaneous skills. These skills include perception evaluation, construction, categorization etc. It also involves socializing into a social community of practices and shared goals. This conception of how theories are constructed spells out the connection between skill, practice, and the more naïve concept of theory mentioned above. After a while theories get internalized in the sense that they are automatically acted upon. This might seem to reduce their status as theories, since they become less visible, but, according to the Churchlands, theories they remain. And in this respect Folk Psychology is one of the most thoroughly internalized theories humans ever acquire. This then explains how “the claim that FP is an empirical theory is entirely consistent with – indeed is explanatory of – the intricate practical life enjoyed by its adepts”.<sup>65</sup> It is important in this context, to notice that practices as well as theories might be modified or replaced.

#### *4. The fourth objection: Could there be an alternative?*

The fourth objection draws on our ability to imagine what a good alternative to FP would be. Proponents claim that there is no existing alternative, and that there is no hope for constructing a new one in its place. Fodor seems to hold such a view: “commonsense belief/ desire psychology [roughly what I have been referring to as Folk Psychology] explains vastly more of the facts about behavior than any of the alternative theories available. It could hardly fail to do

---

<sup>65</sup> Churchland (1998, p. 34)



so: there *are* no alternative theories available.”<sup>66</sup> He to wants to see a scientific account of human psychology, but it should be constructed around the commonsensical framework of Folk Psychology. His faith in commonsense intentional psychology rests on the fact that it has been with us for as long as we have been thinking creatures, and that it thus far has proved, on his view, very successful. There are no compelling reasons to “doubt that very many commonsense belief/ desire explanations are – literally – true.”<sup>67</sup> And if they turn out to be true they are, on his account, a necessary part of the ontological catalogue of the world. It is not only that FP is the sole good vocabulary around, but it is the only one that we can imagine ever will be able to accurately articulate the different aspects of human behavior and mental life.

However, we should not be discouraged by our limited imagination. According to the Churchlands, it is possible to construct new and radically different theories that will be able to explain and predict both human behavior and mental life better than the current commonsensical framework of FP. But such a revised explanatory framework might be radically different from the existing one. The Churchlands think there are good alternatives to FP currently being investigated in computational neuroscience. One of the alternatives investigated that seems to be a particularly good substitute is Neural Networks. It differs quite radically from traditional Folk Psychology.

##### *5. The fifth objection: Folk Psychology consists of functional kinds*

The final objection takes as its point of departure the notion that FP is not constructed as a taxonomy of natural kinds, but rather of functional kinds. Since the ontological integrity of functional kinds does not depend on the possibility of being able to find a smooth theoretic reduction to a more basic science, usually understood as natural science, a failure to match the posits of FP with neuroscience does not threaten the ontological integrity of FP. The Churchlands’ main worry with this objection is not that FP is too much like ‘chair’ or ‘bungalow’, but rather that it is too much like ‘phlogiston’ and ‘caloric fluid’. That is, they are worried that FP is a false representation of the dynamic and systematic operations within us. In addition, the central terms of scientific theories have traditionally been taken to be natural kinds. If the sorts of kinds that are involved in commonsense intentional psychology are functional rather natural then there would be little hope that it will be the main part of a scientific theory of the mind.

---

<sup>66</sup> Fodor (1987), p. x

<sup>67</sup> *ibid.* p. xii)

Functionalists typically believe that any given mental state can be realized in a wide variety of physical types<sup>68</sup> (known as multiple realizability), the important thing is not the medium but the function of the mental state. As a consequence they argue that a detailed understanding of one physical system will not shed light on the fundamental nature of mind. Thus, the biological nature of the brain – neurons and neurotransmitters – will not play an important or necessary part in a complete theory of the mental. What makes something a psychological state is its role in an economy of other psychological states, not its intrinsic material properties. Laws characterizing brain processes are independent of laws in psychology.

If the defenders of FP claim that it is mainly constructed on functional kinds, then whether or not it is possible to find a rough reduction is not even a question that will ever arise. The Churchlands maintain that the most important factor about FP is that it is an *empirical* theory and should be treated as such. This is not a necessary condition if you think that it consists mostly of functional kinds. It still plays a role in the explanation, of course, but that role is greatly reduced compared to the status its defenders believe it should have. By focusing on the empirical nature of FP the Churchlands want us to take FP seriously, in the context of a larger account of nature and the place of humans in it. Focusing on functional kinds – chairs and bungalows – on the other hand, deflects attention from the empirical obligation of FP, and when the empirical consideration is eradicated then it also seems unnecessary to compare FP with competing representations of what cognitive activity amounts to. The status of the kind-terms involved in folk theoretical explanations will be the main concern in chapter 5.

### ***Reviewing the Objections***

I promised, before reciting the five objections, that I would look for a pattern in the objections, and in the response. What has emerged from the discussion? There are, I think, two key elements that run through the objections. The first is the disagreement about the status of the intentional vocabulary. Is it a theoretical vocabulary that should meet the requirements of a scientific vocabulary? Related to this question is also a question about the function of the intentional vocabulary; that is, can we separate the scientific purpose of it from all the other purposes it serves? The second element is the worry that if we eliminate FP from the account of the mental (scientific or otherwise), then a lot will be left unaccounted for. Specifically the notions of mental content and mental representation are something that the objectors feel that

---

<sup>68</sup> See, for instance, Putnam (1975), reprinted in *O'Connor & Robb* (2003, pp 210-221)

the eliminativist is unable to give a satisfying account of. In the next sections I will take a closer look at these two elements.

### *Is Folk Psychology a Theoretical Vocabulary?*

The Churchlands insist that FP is theoretical and whether it is a successful theory is an empirical question. The decisive settlement must flow from experimental research and theoretical development. But the objectors cannot agree that FP is a scientific theory; it is, they claim, involved in a range of other non-scientific circumstances and should not be evaluated as an empirical theory.

I think it will be instructive to take a closer look at the idea that our understanding of the mind is based on theory.<sup>69</sup> The idea of the theoretical nature of Folk Psychology can be traced back to Quine and Sellars.<sup>70</sup> Quine showed that the status of science does not rest on the foundation of what is immediately given to us, but is theoretical all the way down. This also applies to our everyday conception of Physics (what we could call Folk Physics). Sellars attacked what he called the Myth of the Given – the idea that some of our beliefs or claims have a privileged epistemic status because the facts that make them true are given to us by experience. One class of claims that traditionally has been awarded such privileged status was the reports of our perceptual experiences. This privileged status of perceptual experiences (or sense data, as it also was called) seemed to confirm that we could not be wrong about our own mental states; that we are infallible when it comes to our own thoughts and experiences. If we were infallible when it came to our “inner mental-states” and thus that our reports of those experiences were incorrigible they would hardly seem to be theoretic in nature. One of the characteristics of a theory is that it is fallible, that it is open for change in light of new data. Sellars rejects the idea that we are incorrigible when it comes to our own experiences of our inner mental states. Consequently, they cannot be ruled non-theoretical on the ground that they are known with certainty.

Others, besides the Churchlands, have recently defended the idea that Folk Psychology is a theory, amongst them Gopnik and Wellman<sup>71</sup>, without necessarily agreeing with the eliminativists. Gopnik and Wellman apply the method of contrast to show that the understanding of other minds relies on the ability to form a *general theory* of other minds. The

---

<sup>69</sup> The question of what constitutes a theory and what differentiates it from other kinds of frameworks is an enormous one and exceeds the scope of the present discussion, but it is possible to say something that seems to apply to theories in most cases.

<sup>70</sup> Sellars (1956), Quine “two Dogmas” (1953)

<sup>71</sup> Gopnik and Wellman (1995) “Why the Child’s Theory of Mind Really is a Theory”, printed in Davies and Stone (eds.) (1995).

two views they contrast are the idea that a “child’s early understanding of mind is an implicit theory analogous to scientific theories, and changes in that understanding may be understood as theory changes”<sup>72</sup> with the notion that what really counts is the special importance of the first-person understanding of the mind, i.e. by drawing on the first-person view “the child need not really understand the mind, in the sense of having some set of beliefs about it.”<sup>73</sup> The child does not need to form any beliefs about the other mind since she can operate on a working model of her own mind and “[t]he child’s task is to learn how to apply this model to predict and explain others’ mental states and actions.”<sup>74</sup> Gopnik and Wellman think that this dispute cannot be settled on a priori or conceptual grounds; it is a contest between two empirically testable hypotheses about the nature of Folk Psychology.

The first thing to notice is that “[t]heoretical constructs are abstract entities postulated, or recruited from elsewhere, to provide a separate causal-explanatory level of analysis that accounts for evidential phenomena.”<sup>75</sup> The abstract constructs are thought to give explanations of the empirical phenomena available and they explain the data without merely rewording the empirical data. In a sense, the theory is to reach beyond the scope of the empirical data. Another important idea is that the theoretical construct work together in coherent systems characterized by laws, structures or patterns. The explanatory force of a theory stems from both its abstractness and from its coherence.<sup>76</sup> Further, the theory should be able to predict new and unobserved phenomena and events (that reasonably fall under the domain of the theory). This is not to say that all predictions made by the theory should be correct. Since the theory goes beyond evidence, and since there is a chance the theory is wrong, some of the predictions might be wrong as well. These predictions help produce interpretations of evidence not just rewordings and generalizations about it.

What about Folk Psychology, is it a theory in this sense? Gopnik and Wellman notice that “there is a change from one mentalistic psychological theory to another somewhere between ages two and a half and around four.”<sup>77</sup> During that year and a half, or so, the child goes from one way of perceiving how other minds operate to another. In the first stage children do not rely on inferring the beliefs and desires of other agents and they seem unable to distinguish between different states the other agents might be in. They seem to perceive desires

---

<sup>72</sup> Gopnik and Wellman (1995), p. 232

<sup>73</sup> *ibid.* p. 232

<sup>74</sup> *ibid.* p. 232

<sup>75</sup> *ibid.* p. 233

<sup>76</sup> Kuhn noticed that when enough data contradicted the theory, when the coherence of the theory was challenged, it would in the end lead to the fall of the theory.

<sup>77</sup> Gopnik and Wellman (1995, p. 236)

and perceptions as fairly simply causal links between the agent and the world (without any mental representation). At the age of four (or five) the child is able to distinguish between different types of mental states and their different relations to the world of objects. When the child acquires this new understanding of minds, (others and her own) she adds to her repertoire means to make additional predictions, explanations and interpretations

This, then, gives us reason to think that the child now has a theoretical understanding of the mind, in the sense laid out above, since “the child’s understanding involves general constructs about the mind that go beyond the focal evidential phenomena.”<sup>78</sup> The child goes beyond the empirical evidence at her immediate disposal and constructs a more general account of the behavior of other persons (and perhaps animals). This new abstract and generalized account “allow[s] children to make predictions about behavior in a wide variety of circumstances, including predictions about behavior they have never actually experienced and incorrect predictions.”<sup>79</sup>

What can be concluded from this discussion, from a pragmatic point of view? The pragmatist does not deny that the intentional vocabulary might have a theoretical nature, but this does not necessarily, she will claim, lead to the conclusion that it must be eliminated in favor of another better suited *scientific* theory. When FP is used for scientific purposes then it must meet the obligation that any empirical theory should meet, but it also serves other purposes and in those circumstances it may be inappropriate to impose the requirements of scientific explanation on it. However, it would be equally wrong to conclude from the last point that FP is irreducible in principle or that it will never be replaced. This reflects the idea mentioned in chapter three that all vocabularies serve some sort of interests. And ultimately what makes a vocabulary useful or not is whether it is able to satisfy those interests and demands we impose on it at the outset.

### *Mental Content*

The second element that ran through the objections was the worry that the Churchlands’ suggested solution to the problem of mental content and representation, Neural Networks, is unfit for the task it is proposed to solve. Traditional propositional attitude theory (and belief/desire psychology) assumes that human cognition is “a dance of sentential or propositional states, with the basic unit of computation being the inference from several such states to some

---

<sup>78</sup> *ibid.* p. 239

<sup>79</sup> *ibid.* p. 239

further sentential states.”<sup>80</sup> But, the Churchlands’ argument is that “the sentential attitudes do not play a preeminent role in our cognitive economy. At best they have a limited and superficial role. There is some sentence-crunching, almost certainly, but it is not constitutive of cognitive activity.”<sup>81</sup> Another approach has to be made to the mental, one that does not rely on sentence crunching.

In Neural Networks representations are construed as partitions in multi-dimensional vector space. Neural Networks have vector-to-vector transformations which are quite different from the semantical/logical structure of the posits of Folk Psychology. The basic unit in a neural network is the neuron (or a model of it in simulations). These are highly interconnected nodes with a layer of synaptic weight between them. Vastly simplified a neural network is a parallel machine, in the sense that signals are processed in many pathways at the same time and the sentential structure of Folk Psychology is replaced with activation patterns across large groups of neurons.

Fodor and Lepore (1996)<sup>82</sup> cannot see how a neural network can give rise to mental content, or in their words “[s]ince Churchland’s attitude towards the intentional/semantic generally tends to be eliminativist, it’s unclear just what properties of contentful states his ‘state space’ representations are supposed to preserve.”<sup>83</sup>

The first thing to notice is that the way a content and mental representation is accounted for in Neural Networks is radically different from the traditional approach of propositional attitude philosophy. The traditional belief/desire philosophy, which is intended to capture the framework of Folk Psychology, is based on the assumption that the brain operates in a sentential fashion, and takes states as beliefs, perceptions, desires, and preferences as the basic units of human cognition. These states are then typically described in sentences such as “A believes that p” and “A desires that q”, where A is a person and p and q are sentences. These propositional attitudes are supposed to capture the property of the mind by which it is directed at, about, or ‘of’ objects and events in the world. Since the basic unit of cognition in a neural

---

<sup>80</sup> Churchland, P.M. (1995, p. 322)

<sup>81</sup> Churchland, P.S. (1986, p.396)

<sup>82</sup> This essay called “Paul Churchland and State Space Semantics” is more or less the same as Chapter 7 in Fodor and Lepore (1992). One of the most notable differences is that Fodor and Lepore (1992) uses “intentionality” (with a ‘t’) where Fodor and Lepore (1996) use “intensionality” (with a ‘s’). Since what is at stake here is whether the proposed alternative can accommodate for contentful states and meaning, I take it that “intentionality” (with a ‘t’) is what is meant.

<sup>83</sup> Fodor & Lepore (1996, p. 143). Another issue for Fodor & Lepore, apart from eliminativism, is whether or not content and meaning can be given a holistic account. They strongly doubt it, but that is a discussion for another place and time. However, just for the record: the approach favored by the Churchland is holistic: “What gives this vector the content ‘kitten’ is the overall role that this vector plays in the larger cognitive and motor economy of which it is an interlocking part.” Churchland & Churchland, reply to Fodor & Lepore, in McCauley (1996, p. 275)

network is the distributed representation over several thousand neurons, there seems to be no room for these propositional attitudes – “the paradigm of intentionality”<sup>84</sup> – in ‘state space semantics’, or so at least, the argument of the objectors runs. But the argument misses its target. It does so because the objectors fail to see that “they are now looking at something that does not fit any of their antecedent doctrinal categories.”<sup>85</sup> To demand that the new, revised, theory of meaning should account for the propositional attitudes of the old theory *within the framework of the old theory*, is begging the question. In other words, to object against eliminativism on the grounds that the proposed alternative to meaning, the ‘state space semantics’ of neural networks, fails to capture the intuitive approach of belief/desire psychology falls short since a “neurally grounded theory of meaning may require revision of the very intuition that now seem so secure [...]”<sup>86</sup> The eliminativist claims, on her side, that the traditional propositional attitude approach does not fit in with the evidence coming from neuroscience, and that it should be revised accordingly. The pragmatist takes the lesson from the eliminativist, but does not agree that the solution necessarily is elimination of the intentional idiom. What the pragmatist wants to emphasize is that the intentional vocabulary not only serves the scientific needs of neuroscience, but also the normative needs of everyday interaction.

It would seem that in order to justify a strong eliminativist conclusion it has to be shown that there is nothing in the scientific psychology that shares the central properties we assign to beliefs and desires, at any level of analysis. The Churchlands take FP, as mentioned, to be essentially an empirical idiom. As such it fails to meet up with the standards usually imposed on any theory that is to be part of an empirical driven explanation. The concepts making up the base of an empirical theory are typically natural kinds and it is not all too clear that the terms in FP are natural kinds or some other class of kinds. It depends on how we analyze the concepts used in the intentional vocabulary, and how we understand the concept of natural and unnatural kinds.

This might mean that eliminative materialists are right in questioning an uncritical use of FP; it might be that those very terms are in need of change. But it does not automatically follow that the terms of the intentional vocabulary are not suitable in constructing a theory of the mental that meets empirical needs. It turns on how we perceive the terms of the intentional idiom. Are they functional kinds, human- made and, as such, not suitable to “cut nature at its joint”, or are they natural kinds, concepts that point to “real” differences in nature? The

---

<sup>84</sup> Fodor & Lepore (1992), p. 193

<sup>85</sup> Churchlands second response to Fodor & Lepore in McCauley (1996), p. 278

<sup>86</sup> Churchland & Churchland (1998), p. 61

pragmatist thinks that there is room to navigate between a scientific legitimation of the propositional attitudes of Folk Psychology and the outright elimination of the Folk Psychological idiom. That is, the pragmatist wants to grant an element of the eliminativist argument – the notion that FP probably is not the best vocabulary to capture the scientific needs of neuroscience – but she does not think that the eliminative conclusion follows.<sup>87</sup> In order to see why, I will now turn to natural kinds and folk theoretical explanations.

---

<sup>87</sup> This last point is also against e.g. Fodor and Lepore, who think, like the Churchlands, that the eliminativist conclusion would follow, if the premise were true. So they do not accept the premise.



## 5. Natural Kinds and Emotions: Paul Griffiths

The proper method of philosophy consists in clearly conceiving the insoluble problems in all their insolubility and then in simply contemplating them, fixedly and tirelessly, year after year, without any hope, patiently waiting.

-Simone Weil

The worry articulated by the eliminativist is that the way we currently speak about the mental does not pick out patterns that are best suited to be a part of a scientific theory. This, in turn, means that if we are sincere about constructing theories about the mental we should expect the terms of the commonsensical framework of Folk Psychology – and, as we will see, the vernacular category of emotion – to change radically, so much in fact that have to be eliminated from a serious theory of the mental. That is what eliminativists are trying to tell us and not that we are not conscious, that we do not experience emotions, that we do not have beliefs, or anything like that. It is not so much that we do not have beliefs as it is a problem with the way we try to give an explanation of those beliefs that will fit a scientific account of the mind. But what is a good scientific explanation anyway? Is it possible to single out a set of characteristics that signals that we are dealing with a primary scientific explanation? And what should be part of a scientific theory of the mental?

### ***Conceptual Analysis and Propositional attitudes***

Even though a lot more is known about the brain today than ever before there are still deep disagreements about how to approach the connection between the mind and the brain. This might be due to methodological disagreements, but also, some think, conceptual disagreements. The main problem is that theorists of the mind (i.e. neuroscientist, cognitive scientists and philosophers) do not agree on what counts as a correct answer to the question “what is the mind?” or “what is consciousness?” This opens for a disagreement not just about whether a given theory accommodates the relevant data, but also a disagreement on what a current theory might *actually* be about. A relevant question to ask, but one that has proved exceedingly hard to answer, is whether the theory in question actually is a theory of the mental, or, to put it another way: Are we truly giving an explanation of the relation between the mind and the brain? Or, as the pragmatist would put it: have the interests that made us ask the questions in the first place been satisfied?

What some philosophers<sup>88</sup> suggest is that what is wrong with contemporary cognitive research is failure to appreciate that it faces a specifically philosophical problem that needs to be settled before cognitive scientists can do their empirical work. Advocates of this approach believe that the current mind-body problem can be solved by doing armchair philosophy.<sup>89</sup> They argue that we need conceptual clarity in order to settle the current mind-body problem and the approach they favor in order to resolve these problems is to analyze the ordinary competent speaker's use of a concept and the logical structure of the sentences in which the concepts occur. The rationale behind this approach is the conviction that if the *correct description* of ordinary speakers' use of particular concepts can be worked out – their application and the logical structure they are part of – then conceptual clarity is achieved and this will mean a long step in the direction of resolving the current problems of neuroscience and philosophy of mind.

Conceptual analysis has been the dominant method in the philosophy of mind the past thirty or forty years. At the core of this approach is the conviction that a theory of the mental can be constructed by analyzing the beliefs and desires (and related propositions) that feature in everyday psychological explanations. These beliefs and desires, the propositional attitudes, are taken to be the central theoretical entities of the model and they are often assumed to be the simplest components of thought. The mainstream propositional attitude theory assumes that our everyday understanding of mentalistic terms is sufficient for a theory of the meaning of the mental terms. This is presented as if it were a straightforward outcome of the notion that mental phenomena are occurrences of mental states that represent states of affairs in the world, and that the analysis of these states are sufficient for giving the content of that mental state.

According to Griffiths a propositional attitude theorist takes it that “Mental states are *defined* by the rules which ordinary speakers use when applying mental state terms.”<sup>90</sup> Consequently the important part for the propositional attitude theorist will be to sort out and explain how the mental terms are used by ordinary speakers. This in turn is intended to fix the content of the mentalistic concept.

---

<sup>88</sup> See, for instance, Bennett and Hacker (2003). They argue that the reason why neuroscientists thus far have failed to come up with a theory of consciousness and the mental in general that everyone can agree upon is that the question is wrongly stated at the start. In short, they claim that there cannot be an answer to the question as posed by neuroscience, not because of shortcomings in the tools available, whether it is PET-scanners or the mind itself, but because any answer to the question will be nonsensical. This, they claim, is because neuroscience has taken a conceptual question to be an empirical one.

<sup>89</sup> Conceptual analysis is often thought to be the central method of philosophy and it has a distinguished history: Plato, for example, sought an analysis of knowledge and Locke aimed to examine complex ideas in terms of simple ones, by doing analysis of the concepts involved.

<sup>90</sup> Griffiths (1997), p. 23

However, it is appropriate to ask if it is possible to make sense of such a claim. It seems to put a lot of weight on the theory of concepts in question. How is this theory supposed to account for conceptual changes due to scientific progress? Griffiths notes that: “Our best definition of water is the formula HOH, but this could not have been established by conceptual analysis! Ordinary English speakers in the past did not know that water was HOH, but this is still an important part of what ‘water’ means.”<sup>91</sup> Similarly, nothing important would be added to the scientific theory of the mental by analyzing the *current use* of the mental vocabulary. The conceptual analysis approach does not capture the fluid concept and creative analogy approach most scientific processes demonstrate.

If this line of reasoning is correct the kind of reasoning provided by conceptual analysis is erroneous because it can only tell us what our *current* beliefs within a domain are. The reason for this is that it only specifies the current conditions of application for the terms within that vocabulary. Given that the pragmatist is right when she claims that the most important changes in our theories are achieved not through arguments and analysis, but through changes in the vocabulary, conceptual analysis will not help advance the theories of the domain in question. Propositional analysis theory seems too rigid to incorporate new empirical findings into a rapidly evolving theory. It cannot accommodate the radical changes that often occur in science

Griffiths, whose main concern are emotions, observes that: “Propositional attitude theorists think conceptual analysis is the only tool they need to investigate emotions because they accept, explicitly or implicitly, a Wittgensteinian distinction between the ‘criteria’ which logically define mental state and the inessential ‘symptoms’ that can be studied empirically.”<sup>92</sup> This separation between the logical criteria that define mental states and the inessential part of them open for empirically study is something the pragmatist fundamentally disagrees with.

According to the propositional attitude theorist the analysis of concepts and conceptual relationships antecede the empirical investigation. One of the problems with this method is that conceptual revision in light of new empirical findings would, given such a division of labor between the scientist and philosopher, be, if not impossible, then at least something that would be very hard to explain. It would seem that all scientists ever could hope for was to be handed a set of neatly organized concepts by the philosophers and then try to assess their truth according to the facts they have gathered from the world. The pragmatist, who takes her lesson from Quine, denies that there is a principled distinction between conceptual

---

<sup>91</sup> *ibid.* p. 36

<sup>92</sup> Griffiths (1998), p. 23

and conceptual and empirical questions. What the pragmatist does, simply put, is to take the Quinean insight and add temporal dynamics.

This is not to say that the pragmatist conflates empirical and conceptual questions into one, she does acknowledge the important distinction between the different modes of inquiry; it is just that she sees the distinction between conceptual and empirical matters as temporary and local. She doubts that it is possible to know in advance and in isolation of the development of a theory exactly what kinds of things that are appropriate objects of the domain. That is usually something that is sorted out as we go along and the theory in question matures. And when doubt has been cast on the sort of objects that should be part of the inquiry, it is obvious that it is hard to work out the logical relation between the thus far unknown objects of the domain. This has to be done as development proceeds and it is during this process that the local and temporary differences between empirical and conceptual claims become apparent.

It would seem that the pragmatist approach is orthogonal to that of the propositional attitude theorist. Whereas the latter tries to free neuroscience and philosophy from its present mind-body problem analyzing the logical relations among the psychological terms and suggesting the correct application of those terms, the pragmatist insists that we have to be open to the idea that our vocabularies are changing and that we at the present time cannot say what is the correct use of that vocabulary. Our interest in a given domain might change as well. Since there is a close interplay between the vocabularies we use and the interest we have in those vocabularies we should expect this to influence the use of a vocabulary.

The main tool for the pragmatist is re-description and metaphors; by exploring new connections between modes of thinking she tries to shake our intuition and make us accept new ways of approaching things that already were taken to be literal truths. This makes the enterprise of giving a stable and exact description of what mastery of a given concept is very difficult. In fact, the pragmatist philosopher doubts that such a description could be made available. If it were to be given it would only give a temporary definition of the term.

### ***Natural kinds and Scientific Explanations***

A consequence of this line of reasoning is that we will have to rethink our relation to some of the most common philosophical tools. One of the things that should be reconsidered is how we perceive the notion of natural kinds. As we have seen, some have been criticizing the eliminativist position by claiming that the framework of folk psychological terms, what I have been calling the intentional vocabulary, is made up of a taxonomy of functional kinds rather

than natural kinds. They proceed by arguing that the intentional terms will never be reduced to neuroscience since the ontological status of Folk Psychology does not depend on it being possible to make a rough reduction of the relevant terms. It is ironic that this also is the reason why eliminativists want to see the elimination of the Folk Psychological framework and it might point to a problem with the current conception of how kinds are to be constructed and interpreted.

As mentioned earlier I am endorsing a view that claims to offer ways of thinking about the mental and the physical that implies neither reduction nor eliminativism. I think this can be given better support if we reconsider the relation between natural and functional kinds. Natural kinds tend to be central in scientific descriptions and explanation. Traditionally conceived, science has been involved in giving explanations of the causal processes in nature. Natural kind terms are central to such a project since they are usually taken to denote categories that allow reliable extrapolation from samples of the category to the whole category. An ultimate scientific explanation, some held, required that these categories be subject of universal, exceptionless “laws of nature”. This, as it turns out, demands too much of even the most robust natural kinds. But luckily it is possible to recast this requirement and generalize from the idea of “law of nature” to the somewhat weaker idea that statements are to a varying degree lawlike; that is, that they do support counterfactuals, but that this support is always to some degree conditional.

There are at least two good reasons to believe that a proper account of natural kinds would be useful to science. First, the theory of natural kinds captures an important aspect of human concept formation and how concepts are used. It is therefore not only a very good scientific tool: since the formation of concepts needs to be informed by what we know it also provides a glimpse of the psychology of concepts. Secondly, it can be modeled in such a fashion that it allows us to mediate between competing strands of empiricism and realism. That is, in Griffiths’s words, “If the theory of natural kinds is a central part of the best scientific account of concept formation and use, then an ability to make sense of this becomes an adequacy condition on any account of how thought and language relate to the world.”<sup>93</sup> It is only the strong metaphysical realist who would claim that the link is one-directional: the real structure of the world dictates which concepts that best fit in a theory describing it. For Griffiths there are no theory-independent kinds that will anchor such a theoretical view. Categories and kinds are dependent on the theoretical concepts that represent them and theories

---

<sup>93</sup> Griffiths (1997), p. 175

and concepts change as we interact with the world. Natural kinds can provide an explanatory link between the world and the concepts we use to categorize it.

It is also desirable that an account of natural kinds is in accordance with what we now know about cognitive development, and how we form concepts. Experiments show<sup>94</sup> that as children grow older they tend to regard natural kinds as grouped together by some sort of inner essence, distinct from their superficial properties, while kindergarten kids put more weight on physical appearance. For the older kids both genealogy and function are more important than the overall physical appearance. But in both cases experiments suggest that children in general privilege the intended function of objects when it comes to human made artifacts while they believed that biological species have some unseen properties that survive even if their superficial characteristics have been transformed.

A study done by Santos, Hauser and Spelke<sup>95</sup> concludes that the ability to classify objects in a domain-specific manner is something humans, in a large degree, share with other non-human primates. They set out to answer the question “How do organisms decide which features to attend to in order to build effective strategies for classifying the complicated assortment of objects in their world?”<sup>96</sup> The study was carried out on two human children and two non-human primate species – captive cotton-top tamarins and free-ranging rhesus monkeys. They used foods and artifacts as the two different domains the subjects were to be confronted with. As already mentioned, human children tend to choose artifact according to their functionality (e.g. shape, rigidity, and size) over physical appearance. Both the tamarins and the rhesus monkeys showed the same behavior. They reliably chose tools according to their functionality, rather than physical appearance, when presented with novel tools with varying colors and shapes (these tools were typically used for retrieving rewards). The experiment then turned to another domain, food, where the features of shape and orientation play little role.

Children from around the age of three or four start to predict that objects of the same color will have the same smell and taste. In a word-labeling test Santos et al. was able to show that “children used the feature of color, not shape, when generalizing labels to new food objects.”<sup>97</sup> In the same manner they were able to show that the monkeys generalize their learning about edible objects (watching others eat) according to color, but not shape. This means that there is a propensity both in humans and in monkeys to generalize in a domain-

---

<sup>94</sup> See Griffiths (1997, p. 182-186) and Santos Hauser and Spelke (2002)

<sup>95</sup> See Santos, Hauser and Spelke (2002), for the details of the study.

<sup>96</sup> Santos, Hauser and Spelke (2002), p. 205

<sup>97</sup> *ibid.* p. 210

specific manner.<sup>98</sup> This suggests that objects are not clustered on a neutral basis from which understanding is achieved, but they are rather clustered on the basis of background theories. And this is what we should expect. Different features are likely to be part of different domains of inquiry. Sometimes it is relevant to ask of the age and sex of an animal sometimes it is not, it all depends on the question you want to be answered. And this should be reflected in a well informed theory of natural kinds.

We should be careful before we draw the conclusion that the propensity to group objects in kinds according to domain-specific properties is genetically programmed, though. This is might not be a good way of capturing how a trait that has evolved is realized in a new generation. It disregards the idea that traits depend on a whole of range developmental resource in order to be realized.<sup>99</sup>

The traditional view of natural kinds, with exceptionless laws, would seem too rigid to allow anything to count as a scientific biological explanation. For instance, how are we to account for speciation in light of exceptionless laws? Individuals within species usually differ radically in wide range of properties – morphological, physiological, and behavioral, to mention a few. Genes and phenotypic characters vary too much within the groups over which biologists need to generalize. Since species are one of the prototype examples of natural kinds, the theory should be able to accommodate for such local variation within species. Griffiths' idea, therefore, is to rework and soften the notion of natural kinds, in light of the needs posed by biology. By abstracting from the specific need of biology we will see a theoretical tool better suited to explain the categorization of different concepts involved in scientific explanation.

### ***The Essentialist View***

Speciation in biology is intended to show a deeper sameness or essence in different members of the same species and help describe the differences with other species, and “one way to express this deeper commonality is to say that members of a kind share an essence – a property common to all the members of a kind and responsible for each member being the kind of thing that it is.”<sup>100</sup> The meta-task for biologist is to find a method that allows for a way to capture and

---

<sup>98</sup> “Like humans, monkeys represent and reason about objects differently in different domains” (Santos et al. 2002), p. 212

<sup>99</sup> For a good discussion on the interplay between genetic programming and the environment see *Nature vs. Nurture* by Matt Ridley (2003). He shows in a very informative way that it is usually impossible to draw sharp line between the influence of the genes and the environment when it comes to determining our phenotypic development.

<sup>100</sup> Griffiths (1999), p. 209

explain this underlying “nature”. The underlying nature is then taken to constitute a natural kind.

But, it has proven very hard to agree on what exactly should count as an essential property and how to describe it. What should count as the underlying “nature” of two members of the same species, the “nature” that biological explanation is intended to capture? Different approaches are possible, but it seems to be agreed upon that “the shared nature of a biological kind cannot be either a set of phenotypic characters or a set of genes. Both vary too much within the groups, such as species, about which needs to generalize.”<sup>101</sup>

### ***Paul Griffiths and Natural Kinds***

I will now turn to a view of natural kinds that is sensitive to both evolutionary reasoning and the revisionist perspective central to the pragmatist’s position. Paul Griffiths has recently given such an account of natural kinds.<sup>102</sup> His notion of kinds is one that not only “cuts nature up at its joints” by analyzing the referents of the terms involved; it is also historical in essence. On Griffiths’ liberal view, natural kinds are not the most fundamental categories of nature as was the common idea in the early days of natural kind-talk. Instead they are viewed as non-arbitrary “ways of classifying the world that correspond to some structure inherent in the subject matter being classified”.<sup>103</sup> The order of the correspondence here is important to emphasize. For Griffiths it is not the case that we have to make our concepts correspond to some real structure in the world in order to achieve scientific progress, rather it is “empirically adequate background theories which allows us to construct projectable concepts.”<sup>104</sup>

Griffiths view does not *only* rely on the idea that kinds should be natural and useful in scientific explanation. What makes his account particularly interesting is that it explicitly states that what is taken to be a natural kind is relative to the domain of which it is a part. In Griffiths’ own words: “The use of different concepts promotes different agendas.”<sup>105</sup> This makes it very interesting in the context of the pragmatic stance I have been sketching, with different vocabularies as promoters of different interests. Although his approach was mainly constructed

---

<sup>101</sup> *ibid.* p. 213

<sup>102</sup> Griffiths (1997) and Griffiths (1999). He, in turn, bases his account on natural kinds on the work of R. Boyd. He has also suggested that to rid his account on natural kinds of the historical baggage attached to the term it should be substituted with ‘investigative kind’ (see Griffiths (2003)). I will use the old term in this paper since the motivational force behind it – scientific explanation, how induction and prediction is possible – is more or less the same as the old one.

<sup>103</sup> Griffiths (1997), p. 6

<sup>104</sup> *ibid.* p. 175

<sup>105</sup> *ibid.* p. 198



to accommodate emotions and their role both in psychological and neurological theories, and the need for biological explanations generally, I think there is a lot we can learn by extrapolating his idea of kinds and see whether it is applicable to the mental domain in general.

Another reason that makes Griffiths approach interesting in our context is that he regards himself as an eliminativist.<sup>106</sup> This is a conviction he has in common with the Churchlands, although the subject of elimination differs somewhat. While the Churchlands argue that the folk psychological framework is up for revision and elimination, Griffith attacks what he calls the vernacular categories of emotion. This way of referring to emotions shows that it has some common features with Folk Psychology: both are taken to be commonsensical frameworks and to be what people (lay persons) refer to as emotions or mental phenomena. When the Churchlands claim that FP is widely erroneous and will be eliminated in favor of a matured neuroscience they base this on the fact that it fails as an empirical theory. Griffith claims, on the other hand, that the vernacular category of emotion is not suited to the scientific theory of emotion since the concept does not form a natural kind. This is also an empirical matter, but it would seem that Griffith's position is stronger, in the sense that he does not have to rely, as the Churchlands do, on historical analogies. At the same time Griffith is less categorical when it comes to the elimination of the framework. He is not opposed to the idea that the commonsensical framework serves other, non-scientific, purposes. It is the scientific use of the vernacular category of emotion he attacks, and in this sense it is up for elimination. It would at this stage be instructive to see why Griffiths opts for the elimination of emotion and then compare it with the Churchlands' idea of elimination with respect to the vernacular mental categories. I think it is possible to point at some instructive similarities and differences. Let us start by taking a closer look at Griffiths' argument for the elimination of the vernacular concept of emotion.

### ***Natural Kinds in Biology***

Griffiths' account of natural kinds is by and large based on examples from biology. According to Griffiths, biology is facing the following problem: if taxa are genealogically or historically defined, they cannot be natural kinds. Natural kinds are traditionally thought of as subject of universal laws (spatiotemporal unrestricted laws of nature), they must have universal applicability and cannot be historically or genealogically defined. If natural kinds are thought of

---

<sup>106</sup> "The refusal to identify emotion in general with affect program phenomena leads to a form of elimination of the general concept of emotion", Griffiths (1997), p. 242, (see also pp. 14-17)

in this way the status of biology as a science is threatened, since the constructions of biological laws are contingent on history and thus not candidates for universal laws (in fact, with regard to evolutionary explanation, it is not clear that we are dealing with explanation by way of law at all). Without laws enabling biologist to make reliable predictions and to extrapolate from observed to unobserved instances, biology is reduced to the study of how things happen to be here and now. Biologists should be able to extrapolate from observed to unobserved instances, that is, the kinds used in biology should be projectable.

In evolutionary biology the notion of natural kinds is intended to group certain species based on clusters of similar properties and give us reliable categories.<sup>107</sup> But there is a problem with the notion of similarity. There are numerous properties an object can have and there are indefinitely many ways of grouping them together. This has posed a problem for the taxonomist who hoped to group together different animals on the account of their overall similarity. But since there is no such thing as overall similarity, only similarity relative to a set of properties, finding algorithms that choose the suitable similarities is basic to the scientific explanation of the taxonomy of species.

There are two fundamental sets of categories brought about by evolution: The first types are homologies: this is when the same structure is providing the same function in different animals. A homologue structure is similar among related organisms because those organisms have all descended from a common ancestor that had an equivalent trait. According to Darwin, bats and whales possess similar bone structures because they inherited them from a common ancestor, not because they were constructed according to the same archetype. The second types are analogies; two unrelated structures are selected by natural selection for the same ecological and functional role. A good example would be wings in bats and birds. The wings of bats and pigeons are functionally alike, but genealogically different. Since speciation is intimately linked to genealogy biologists cannot rely on functionality to show what kinds of species are closely related.

Categorizations made by analogies do not have the explanatory power to help form natural kinds since “similarities due to analogy (shared adaptive function) are ‘shallow’”.<sup>108</sup> This shallowness assures that the deeper you dig the more dissimilarities are likely to become apparent, and the more the mechanisms under scrutiny will diverge. Bird wings and bat wings, for example, are very different in internal structure, although they have similar aerodynamic

---

<sup>107</sup> That is: “the species category is supposed to reliably collect morphological, physiological, and behavioral properties”, Griffiths (1999), p. 215

<sup>108</sup> Griffiths (forthcoming), p. 5

properties. Homological relationships (shared ancestry) on the other hand, are “notoriously ‘deep’: even when the function has been transformed, the deeper you dig the more similarity there is in the underlying mechanisms.”<sup>109</sup> The homologies are reliable enough to serve as a basis of scientific explanation; that is, they form reliable categories with correlated properties. This clustering of properties is explained by a mechanism referred to as causal homeostasis: “A category brings together a set of objects with correlated properties. The category has causal homeostasis if this set of correlations has some underlying explanation that makes it projectable. A successful category captures [...] a causal homeostatic mechanism – something which means that the correlations can be relied on to hold up in unobserved instances.”<sup>110</sup>

In evolutionary biology the search for causal homeostatic mechanisms was solved by relying on evolution to guide the clustering of properties. Algorithms that in this way selectively pick out patterns of phylogenetic relationships and evolutionary history are called cladistic. Cladistically determined taxa have an historical essence, as opposed to intrinsic essence, since “nothing that does not share the historical origin of the kind can be a member of the kind”<sup>111</sup>

This way of grouping organisms together gives more explanatory power than relying on, say, resemblance in intrinsic properties. By using cladistic algorithms no degree of resemblance will put an organism in family with another if they descend from different ancestors. The success of this strategy can be explained by the fact that organisms have evolved by natural selection, and the cladistic algorithm traces this process. Evolutionary theory helps the taxonomists to search for helpful patterns in the biological data, patterns that do not rely on superficial resemblance, but on properties that are more useful to biological explanation.

This approach to natural kinds differs from the traditional essentialist approach in at least two ways. First, there is no requirement of an occurrence of a universal and unique property occurrence in all the members of a kind. Species can be natural kinds even though the members do not share a qualitative essence. Second, Griffiths’ account of natural kinds allows for species to evolve without excluding them from the natural kind category. Species can vary both over time, and at a time. However, that is not to say that there is no limit to the variability allowed in a species. Species must be sufficiently stable so it is possible to make better than chance predictions about some of the properties of a species.

---

<sup>109</sup> *ibid.*, p. 5

<sup>110</sup> Griffiths (1997), p. 188

<sup>111</sup> *ibid* p. 219

The homeostatic causal mechanism is present in other sciences as well, apart from biology. In chemistry, for example, the underlying mechanism that provides causal homeostasis is the microstructure of the molecules.

Now we can see the emergence of a theory of natural kinds, one that holds up to the demand of projectability without placing a too demanding constraint on the clustering mechanisms. A kind, on this view, is *minimally natural* if it is possible to make better than chance predictions about the properties of its instance. This might seem to be a very loose definition of a natural kind, but it is not without force. First, it does not open for an arbitrary classification of the nature; some causal process in nature must be present to warrant categorization of objects or properties influenced by that process. Second, it assures that the objects being classified are being categorized on the basis of the ‘deeper’ similarity they share (whether it is homological traits in evolutionary biology or microstructure in chemistry). Third, even though the generalizations of the special sciences fail to live up to the ideal of universal exceptionless laws of nature they still provide some counterfactual force, a key part of the conception of natural kinds. And finally, “the minimal account of naturalness lends itself to successive restriction that allow us to distinguish between kinds of greater or lesser naturalness and hence of greater or lesser theoretical value.”<sup>112</sup>

### ***Natural Kinds and the Special Sciences***

Parallel to the idea that natural kinds are supposed to capture the universal laws of nature runs the idea that the natural sciences are the basis from which the other sciences ultimately will derive their legitimacy. This, of course, is the motivational force behind the reductive proposal. Reductionism is taken to be an empirical doctrine and it plays a regulative role in scientific practice. For such a program the search for bridge laws connecting the properties of the special sciences with the physical sciences becomes imperative.

The philosophical account of natural kinds is mainly a contribution to the philosophy of science; it is intended to show how the natural scientist can make predictions and inductions. But the model of natural kinds presented by Griffiths is applicable not only to an ideal – or ultimate – physical science, but to the special sciences as they stand. Griffiths notes that traditionally there has been a difference between the way philosophers use the term natural kind and the way social scientists (linguists and psychologists, in particular) use the term. For the philosopher it is intended to show a real distinction in nature (his own examples are chordate

---

<sup>112</sup> Griffiths (1999), p. 217

and chlorine), while social scientists use ‘natural kind’ to designate categories that all human societies use “such as tree or bird”.<sup>113</sup>

Since the idea of intrinsic properties, or essence, shared by all the members of the kind is abandoned, processes and concepts previously thought to be outside the grasp of natural kinds can now be incorporated into it. Socially constructed kinds are among those who might be natural kinds on this view. Griffiths uses money as an example.<sup>114</sup> The lawlike generalizations of economy do not hold because of the shared microstructure of the monetary units. It is the social conventions of the economy that make the generalizations possible, but this is enough to guarantee that money does form a natural kind. It is easy to see that the requirements of a minimally natural kind are satisfied: it is possible to do better than chance predictions about the economy and there are some underlying causal mechanisms that warrant the predictions. For instance, a definition of inflation might read: “A persistent increase in the level of consumer prices or a persistent decline in the purchasing power of money, caused by an increase in available currency and credit beyond the proportion of available goods and services.”<sup>115</sup> The fluctuation of the economy market can be reliably predicted since there are causal factors that regulate it. Economists can predict when prices will fall and when they will increase and these predictions will be better than chance, based on the causal, albeit human-made, mechanisms that underlie it. This is not to say, of course, that they will always be right, but the economist can assert that the decrease in the money supply typically leads to a tightening of the economy, all things being equal. The economist’s statement carries counterfactual force, a key element of a scientific law, and it establishes the fact that the monetary fluctuations are regulated by causal mechanisms.

It is possible to generalize this line of reasoning into insight into the nature of the special sciences. The special sciences derive their legitimacy, not from the search of bridge laws that will ultimately explain the generalizations made by the special sciences in terms of basic physical science, but from the possibility of finding homeostatic causal mechanisms that explain the lawlike generalizations of that science. This breaks down the traditional distinction between natural kinds and kinds made by humans. The upshot of this way of viewing natural kinds is that the “categorizations from any special science that enter into the generalizations of that science are now commonly regarded as natural kinds. Inflation and schizophrenia take their

---

<sup>113</sup> Griffiths (1997), p. 4

<sup>114</sup> Griffiths (1999), p. 218

<sup>115</sup> From dictionary.com

place alongside electrons and stars.”<sup>116</sup> The bias for basic physical science as the science that best captures the way the world really is, is abandoned.

### ***Eliminating Emotion***

How does Griffiths use this conception of natural kinds to argue for the elimination of emotion? The main reason Griffiths opts for elimination is that the psychological states and processes that fall under the vernacular (or folk theoretical) category of emotion do not allow for a unified scientific explanation. They fail to do so because emotion, on his view, does not form a natural kind. Ultimately, natural kinds are supposed to be “categories which admit of reliable extrapolations from a sample of the category to the whole category”<sup>117</sup> In this way the proper use of natural kinds is to allow for scientific unity and explanation.

But emotion terms are more like the “partial reference” terms, like jade (or lily). Jade is both jadeite and nephrite and as such cannot be treated as a single kind by geologist. When it comes to explaining the occurrence of jade in geological deposits, their geological origins, and when predicting the abundance of jade in unknown geological deposits, two different explanations must be invoked for the different kinds of jade. This, Griffiths maintains, is also true of emotions; different explanations must be invoked in order to explain different occurrences of, for instance, anger. What is referred to in the vernacular as anger might differ substantially in the biological and neurological realization in different circumstances. One instance of anger might be an instantiation of a pure affect program – a short lived highly automated response –in one situation. In another situation it might be a more complex cognitive response to more highly analyzed information. Since these, though all being called anger in the vernacular, have different neurological and biological realizations they are in need of different sorts of explanations. These different aspects of anger are only superficial similar and as such is much like the “folk category of lilies [which] is of extremely limited use for explanation and induction. It collects no cluster of properties beyond a certain vague visual gestalt.”<sup>118</sup> If the sciences of mind are to give scientific explanations of emotion, then they have to develop different theories of different emotions and even different theories for what we now take to be the same emotion. By analyzing the status of emotion in relation to kind-talk, Griffiths wants to see whether the way emotions today are individuated stand up to the needs of

---

<sup>116</sup> Griffiths (1999), p. 216

<sup>117</sup> Griffiths, (forthcoming).

<sup>118</sup> Griffiths (1997), p. 191

a serious science of emotions. To use the terms developed in the previous section: there are no underlying homeostatic causal mechanisms in the vernacular category of emotion.

Griffiths' suspicion is that any psychobiological theory of emotion would have to rely on analogies and as such would not be good scientific theory. Such a theory, although it might be a useful ecological generalization, would not be useful for psychological or neuroscientific generalizations. Empirical discoveries would probably show that the domain of affective and motivational phenomena is large and diverse. The domain does not fall under a natural kind, but it would be possible to discern different kinds of emotions (i.e. primary emotions, secondary emotions, background emotions, etc) within it. In chemistry, for comparison, the basis is provided by atoms and their bonds, "but lumping together mixtures, compounds, alloys, pure elements and pure isotopes on the grounds that they are all 'chemical substance' is not very helpful."<sup>119</sup>

The conceptual analysis theorist may object to Griffiths and claim that he conflates questions about natural kinds with questions of whether the concept of emotion has been given a univocal analysis. But it is easy to see that these two questions have little to do with each other. Superlunary objects, for instance can be given a univocal analysis, something is superlunary if and only if it is outside the orbit of the moon. But to say that the objects outside the orbit of the moon form a natural kind is highly dubious (once again there is an absence of homeostatic causal mechanisms). Especially when we keep in mind that natural kinds, in Griffiths' sense, are basic parts of scientific explanations, and should help form scientific theories. The category of superlunary objects ceased to be a candidate for natural kinds with Galileo's astronomical discoveries and has not played a part in any serious scientific theory since. Griffiths' idea is that the same should be the case with emotions; they do not form a distinct kind suitable for scientific enquiry. Griffiths conclude that the "question about emotion, therefore, is not whether we can give a single 'account' of the category in the sense of a philosophical analysis of the emotion concepts, but whether the category thus singled out is a productive object of scientific enquiry."<sup>120</sup> Thus he stresses the scientific part of the question. In the next chapter I will take a closer look at the idea that it might be that the vernacular categories of emotion and of the mental survive elimination if we argue that they are mainly involved in non-scientific purposes. That is, I will take a closer look at the third objection raised against the Churchlands in the previous chapter, but this time the normative aspect of Folk Psychology will be the crucial point, not whether it is mainly constructed of functional kinds or

---

<sup>119</sup> Griffiths (forthcoming,) p 10

<sup>120</sup> Griffiths (2003), p. 10

not. I start off by comparing the motivation the Churchlands have for the elimination of Folk Psychology with the account presented by Griffiths.



## 6. Griffiths and the Churchlands: Compare and Contrast

Marco Polo describes a bridge, stone by stone.  
“But which is the stone that supports the bridge?” Kublai Kahn asks.  
“The Bridge is not supported by one stone or another,” Marco answers, “but by the line of the arch that they form.”  
Kublai Kahn remains silent, reflecting. Then he adds: “Why do you speak to me of the stones? It is only the arch that matters to me.”  
Polo answers: “Without the stones there is no arch.”

-Italo Calvino, *Invisible Cities*

We have just seen two different kinds of eliminative strategies and natural questions that arise at this point are: what do they have in common and, in what way do they differ? Does the softer pragmatic view on natural kinds lend support to the idea that the intentional vocabulary should be eliminated on the grounds that it is unfit for scientific purposes? Or is it the case that it is irreducible on the grounds of its mainly normative aspects? Natural kinds are not given to us by the world as ahistorical types or objects of lawlike generalizations, but incorporate historical and contingent factors. In the field of biology this means that historically defined taxa are natural kinds with the corollary that evolutionary homologies are natural kinds as well. What happens when this lesson is carried over to the story of our mental lives?

As a first approximation we might say that what differentiates the eliminativist position of the Churchlands with that of Griffiths' is that the Churchlands compare FP with phlogiston and caloric fluids, while Griffiths equates the vernacular category of emotion with superlunary objects. The difference lies in the ontological status of the phenomena the commonsensical framework is being compared to. As already mentioned, the category of superlunary objects refer to something that exists (all objects above the moon), but that particular way of grouping objects has no theoretical value, it will not be useful to science. Phlogiston and caloric fluids, on the other hand, are no longer believed to exist in any ontologically committing way. This means that they are not just unfit for scientific use; there is no way to make the concepts intelligible by reference to nature. Accordingly it seems that Griffiths allows that there is such a thing as emotion (as it is being referred to in the vernacular), while the Churchlands, by contrast, apparently claim that what the framework of FP refers to does not really exist. However, they are not dogmatic in their claim. Echoing the question raised by Rorty – quoted earlier, in the second chapter – Patricia Churchland notes that: “The treacherous difficulty here is that we cannot be sure when we are asking the right question. We do not know whether such categories

as ‘emotion’, ‘declarative memory’, ‘procedural memory’, ‘learning’, and ‘consciousness’ pick out a unified, single *natural* kind, or whether these categories herd together quite motley collections of disparate phenomena – whether searching for the neural substrate for ‘memory’ is like looking for the ‘principle’ that unites jewels, such as amethysts, diamonds, amber, and pearls.”<sup>121</sup>

The Churchlands argue quite forcefully for the elimination of common-sense psychology and although they probably are open for a non-scientific use of FP, in a limited sense, they do not seem to be open for the idea that it might radically change in light of the discoveries made by neuroscience. Griffiths, on the other hand, has a more relaxed attitude towards elimination. He makes the point that common-sense psychology may well develop and change its understanding and classification of emotion in light of scientific discoveries. And even though they are in different businesses that does not preclude the idea that the common-sense framework should not and cannot learn from science.

All this being said, it is worthwhile to notice that the motivation for the elimination of the commonsensical framework is something the Churchlands and Griffiths seem to share. They both see the need for a vocabulary that better serves scientific needs. The classification of objects into superlunary and sublunary objects does not pick out categories that are useful in any serious scientific enterprise; in a word, they do not help astronomers in making inductions and explanations about heavenly objects.

Griffiths argues that the vernacular category of emotion groups together properties with different biological and neurological realizations. There is no shared homeostatic causal mechanism supporting the vernacular category of emotion and as such they are not useful to the scientific study of emotion. Correspondingly, in the Churchlands’ case there is little, or no, hope of finding neurological correlations to the terms of Folk Psychology (or the terms of propositional attitude theory) and therefore FP is not suited to the scientific study of the mental, or cognitive activity in general.

In this sense, categorizations that are not of interest to science do not necessarily categorize unreal objects; it is just that the categorization itself does not serve the needs of science. This raises, of course, the question of whether it is possible to make a viable distinction between scientific vocabularies and other kinds of vocabularies, vocabularies that serve different, non-scientific purposes.

---

<sup>121</sup> Churchland, P. S. (1986), p.152. Emphasis is hers.

One such purpose for the intentional vocabulary is to articulate the different normative constraints and commitments of everyday interaction amongst humans. At this point it might seem that we are only reiterating the third objection raised against Eliminative Materialism in chapter 5: Folk Psychology serves a vast range of non-scientific purposes, and therefore it is not even a subject of reduction or elimination. The point I want to make, however, is that even if at the present there is little or no hope of getting a reductive explanation of the mentalistic terms this has little to do with the idea that Folk Psychology is non-theoretic or non-scientific. There is, in light of the pragmatic view I have been sketching, no longer possible to operate with a distinction between the kinds of vocabularies that get to settle what there really is, and other kinds of vocabularies. This means that just because a concept finds no home in a project of systematic natural scientific explanation, we should not assume that it does not classify real things in a useful way. This is not to deny that there are some vocabularies better suited for scientific practices, but that does not make a particular vocabulary better at capturing the “real” world. It is just that the vocabulary is more useful to the scientist at the present.

### ***The Normative Aspect of Folk Psychology***

One way to object to the elimination of the vernacular psychological terms is to claim that they are primarily normative kinds and whether or not they constitute a natural kind is beside the point. Some claim that the irreducibility of Folk Psychology stems from this normative aspect. Scientific theories, they claim, are only descriptive and the idea of reducing a largely normative theory to a descriptive one would be equivalent to try to derive an ‘ought’ from an ‘is’. The intentional vocabulary incorporates patterns that single out the normative aspect of the interaction between people, in a way that could never be captured by a neuroscientific theory.

Paul Churchland (1981) identifies an objection made against eliminative materialism by those who think that the normative aspect of Folk Psychology guarantees its irreducibility. They build their defense of FP on the logical structure of the proposition at the core of FP, that is, they typically believe in the propositional attitude theory. The structure of FP is, on this theory, intended to be a “characterization of an ideal or at least a praiseworthy mode of internal activity.”<sup>122</sup> It outlines what it is for someone to have beliefs and desires, and consequently what it is to be rational. We might not live up to the ideal rationality laid down by FP, but this does not diminish the normative dimension of FP. They conclude that “though neuroscience may usefully augment it, FP has no pressing need to be displaced, even as a descriptive theory;

---

<sup>122</sup> Churchland, P. M. (1981), p. 399

nor could it be replaced, qua normative characterization, by any descriptive theory of neural mechanism, since rationality is defined over propositional attitudes like beliefs and desires.”<sup>123</sup> We saw in the previous chapter that Griffiths turns down a propositional attitude theory approach to emotion on the basis that it fails to add anything important to a scientific account of emotion (that is, it can only give a description of the current use of the terms involved).

Paul Churchland, in his attack on FP, does not entirely dismiss the idea that there is something normative about the folk psychological terms, but he thinks this is just a coincidence that we “just happen to value”<sup>124</sup> the intentional terms. His argument is that the normative aspect of Folk Psychology does not warrant its survival in light of recent empirical discoveries.

He notes that the rationality that FP ascribes to us is not an ideal rationality as some suggest. Even if our current conception of rationality is largely constituted within the framework of FP, it is not an insurance against the possibility that this framework might be inadequate to capture a deeper a more accurate account of cognition, which he sees necessary.

We have already seen that there is very little a propositional attitude can add to a scientific theory, but it may be that it captures the normative aspect of human interaction in a way that the scientific account cannot. According to Paul Churchland, “the fact that the regularities ascribed by the intentional core of FP are predicated on certain logical relations among propositions is not by itself grounds for claiming anything essentially normative about FP.”<sup>125</sup> He maintains that the logical structure of the regularities predicated by FP is on par with, for instance, the classical gas law, and then goes on to say: “A normative dimension enters only because we happen to *value* most of the patterns ascribed by FP.”<sup>126</sup> And what is more: those evaluations, those normative convictions, are often changed by fresh insight, something the propositional attitude theory seems unlikely to capture. The pragmatist acknowledges the fact that the normative dimension is contingent, but this does not seem to diminish its practical aspect.

There is another path open for Griffiths which turns on a consequence of the historicist approach to kinds that he favors: it is doubtful that it is possible to separate the normative from the descriptive use of a concept. Griffiths does not deny that there is a normative aspect of the concept of emotion, but this does not relieve concepts of emotion from all pressure put on them by empirical discoveries, as it is commonly thought to do. Normative concepts, no less than natural kinds, are equally part of an open-ended, since both normative as well as epistemic

---

<sup>123</sup> *ibid.* p. 399. Churchland attributes this view to Daniel Dennett.

<sup>124</sup> *Ibid.* p. 404

<sup>125</sup> *ibid.* p. 404

<sup>126</sup> *ibid.* p. 404

projects offer motive for changing extension and intension of the concept. Griffiths cannot see that there is a way “that description and prescription can be cleanly separated”.<sup>127</sup> On his view the very same kind (human made or natural) “are used to predict and explain human behavior and to prescribe and condemn that behavior.”<sup>128</sup>

The naturalistic pragmatist has more sympathy with this idea that there is no principled way to separate the normative from the descriptive use of concepts. She also thinks that any full account of human behavior has to include both, especially when those terms figure in essentially normative practices. One such practice is the ability to interpret another speaker. What makes the task of interpretation “practicable at all is the structure the normative character of thought, desire, speech, and action imposes on correct attribution of attitudes to others, and hence on interpretation of their speech and explanation of their action.”<sup>129</sup>

The pragmatic approach I described in chapter three is exactly a stab at explaining the link between language, thought and world. The big difference is that thus far I have been trying to cash out the relation as one that has to do with our vocabularies and how we use them without saying anything very precise about how vocabularies are constructed (or how they come to be constructed; it is indeed less a matter of conscious construction than of development through use), and how they change over time in light of new empirical discoveries and new norms. The pragmatic approach would be greatly enriched, I think, by adopting an account of natural kinds that is more versatile than the old conception of natural kinds. Another important factor is the notion, advanced by Griffiths, that the different categories are natural only relative to the domain in which they are being applied, and that they are connected to this domain by background theories. The pragmatist says that what counts as a correct description of a domain depends on the vocabulary, which in turn is relative to the interest that guides it in the first place. Different normative and theoretical commitments spell out the commitments and constraints at the heart of that vocabulary and the normative and descriptive part of the vocabularies will interact so that the realization that some of the concepts of FP do not constitute a single kind will have impact on its role as a normative kind.

---

<sup>127</sup> Griffiths (2003), p. 19

<sup>128</sup> *ibid.* p. 19

<sup>129</sup> Davidson (1990), p. 325. See also Krabberød (2002), here I argue, with Rorty and Davidson, that there is no way to stop prescribing and just describe. This normative constraint is one important factor in the process of interpretation we cannot escape without abandoning our rationality.

## ***Elimination or Revision?***

Now that we have seen two different eliminativist positions – the Churchlands claim that FP is radically false and Griffiths claim that the vernacular categories do not single out natural kinds – it is worth while to point out some common nontrivial confusions and misunderstandings often contributed to the eliminativist position. Some have taken eliminativism to mean that the *mental as such* should be eliminated. This is, of course, not what is meant by elimination. Just changing the way we talk does not necessary remove the revised referents of the terms from the world. Ontologically speaking what there is is there, no matter how we talk about it.

Changing from a theory that says that stars move in crystalline spheres to a better informed theory that makes no references to crystalline spheres or superlunary objects, does not remove the shine, or the movement of the stars. What is changed is how we perceive the movement of the stars, and what we think makes them shine.

But, there is no way to be sure that the current mode of talking about those items is the best way to capture and describe the reality we are trying so hard to cope with, indeed there is no way to guarantee that there is anything that would count as a best description. This is the vital lessons to be learned from pragmatism and eliminativism: we should always be willing to revise our vocabularies in order to better talk about the world; that is, we should not be afraid of trying to refine our tools for coping with the world, in order to make them better answer our needs and interests.<sup>130</sup> Consciousness and emotions will most likely always play an integral part in understanding and predicting humans, but how we describe the mental and how it relates to our bodies will always be open for revision.

The position of eliminative materialism has often been accused of being a too materialistic theory not sensitive to the “real” issues, humans and their place in the world. The fear is that if the framework of Folk Psychology is eliminated or displaced, a scanty picture of the humans is left. This is the heart of the “de-humanization” objection, but I think it is a motivating factor for all objectors to Eliminative Materialism. What is feared is that it is not only the folk psychological vocabulary that is the object of elimination, but humanity as such.

This, I think, is a common misunderstanding among those who write and criticize the eliminativist position. But a proper, charitable, reading of the eliminativist position reveals that it does not imply anything similar to such dehumanization. On the contrary, at the heart of the

---

<sup>130</sup> This does not preclude the idea that some of our interests are best served by a vocabulary that is highly theoretical and have explanatory force. But the theoretical vocabularies should not be viewed as “the best descriptions” of the world.

eliminativist position lies a wish for a better understanding of how human behavior and their social interaction, that is the “the hope of a comparably superior social practice rooted in a comparably superior account of human cognition and mental activity.”<sup>131</sup> As viewed from the eliminativist position this superior account of human cognition is best achieved by radically reformulating the rough framework of Folk Psychology, since this framework does not seem to meet up with the empirical standards anyone who takes it to be a serious idiom imposes on it. And whether or not we take it seriously does not depend on whether it is a scientific theory. In fact, it does not rely on the theoretical nature of FP at all. The important factor here is that it plays an important role in the explanation of our cognitive enterprises, and as such should be held accountable to the data made available by cognitive science and neuroscience.

Despite what critics might think, then, there is an optimistic side to eliminativism. The eliminativist is confident that if the defective framework of FP is displaced by another and better one it would lead to enhanced understanding of our place in the world.

The main difference between the eliminativist and the pragmatist is that the pragmatist adheres to the idea – originating from Davidson – that an explanation of any causal relation in the world falls under some particular kind of description, and that kinds of description may be distinguished by reference to interest. This is what the vocabulary-vocabulary is meant to capture. Different interests are instantiated by different vocabularies and different interest makes us choose different vocabularies. The upshot of this reasoning is that what looks like a perfectly good explanation in one domain – that is, a description made by one particular vocabulary – might look highly dubious in another. This is especially true if the interests in the second vocabulary are orthogonal to the original one.

What the eliminativist urges is that we have to rethink how we perceive the mental. She wants to see a better informed theory of the mind, since the current theory fails, according to the eliminativist, to accommodate all the new evidence neuroscience has provided the last thirty or so years.

The pragmatist does not think there is anything called a “perfect description”, different kinds of explanations require different descriptions and the descriptions of different causal relations are useful for different reasons. And as such it is in the interest of the pragmatist to articulate an alternative version of kinds that is sensitive to these pragmatic concerns.

It might be wrong to conclude, as the Churchlands do, that Folk Psychology – or the intentional vocabulary – is vacuous by making an analogy with other scientific developments

---

<sup>131</sup> Churchland (1998), p. 35

made during the history of science. Newtonian physics was indeed wrong in many areas (especially when it comes to objects traveling at extremely high velocities), but it still has great heuristic value, and is accurate within a very large range. It still works on acceleration of “normal” sized objects, and objects traveling at “normal” velocity, although it is not *literally* true – “literally true” in this context cashes out the knowledge that there are other descriptions, other vocabularies – in particular Einsteinian physics – that have greater explanatory scope and force.<sup>132</sup> Further, there are cases where the old theory was not entirely replaced by a better developed theory, but rather was incorporated in it. And finally, there are cases when we have a better informed theory, while still retaining the old “commonsensical” framework and explanations. Water is defined as H<sub>2</sub>O, but what is being referred to as ‘water’ in an everyday situation is hardly ever pure H<sub>2</sub>O. What matters for the layman is that the substance in question retains some of the important functions that we usually contribute to water: it must be liquid (well, ice is in some sense solid water, but this is not a complicating factor since it is referred to as ‘ice’), is potable, can be used to extinguish fires, gets you wet, etc.<sup>133</sup>

It might be that what we now mean when we use one of the mentalistic terms is not what the term will refer to when neuroscience has matured, but this is an empirical question that should not worry us very much. The legitimacy of the intentional vocabulary is not secured by the success of a reductive program, whether it is of the “translation” form or of the “disappearance” form.<sup>134</sup> Its legitimacy derives from its contribution to our everyday effort at trying to explain and predict other humans.

There are different situations that would call for the elimination of a vocabulary. One situation, the first situation is when we come to realize that some of the functions of the vocabulary are misguided; that is, sometimes through vocabulary changes we change our view of what needs to be explained. This is what the Churchlands are arguing for: the commonsensical framework of the mental does not fulfill the needs of a scientific account of the mental. But this overlooks the fact that the FP is not constructed to fulfill scientific functions. The other situation that would call for elimination would be the case where the new

---

<sup>132</sup> Vocabularies with greater explanatory scope and force tell us something more of the kinds operative in the physical domain. “*Force* is a measure of the reliability of the predictions made using that generalization. *Scope* is a measure of the size of the domain over which the generalizations are applicable.” (Griffiths, 1999, p. 217) Newtonian kinds are more *restricted* than Einsteinian, and that, in the context of theoretical physics, is an important drawback. Newtonian physics has less scope and force than Einsteinian physics.

<sup>133</sup> It seems that whether water is made of XYZ or HOH does not matter to the lay person when deciding whether or not a substance qualifies as “water”, but it does matter to the scientist trying to make some kind of chemical substance involving water (or HOH).

<sup>134</sup> See Rorty (1965). He uses the term “translation” for the kind of reductive strategy that subsumes and explains the old theory, by “disappearance” he means, roughly, what later have been known as eliminative materialism.



replacing classificatory scheme would guarantee to fulfill all the explanatory, inductive and other functions of the theory. However, even though the explanatory needs might be fulfilled there are still other things we do with the intentional vocabulary that run parallel to the needs of the neuroscientist. We use the intentional vocabulary to manipulate and affect other people. It is unlikely that this practical function of the intentional vocabulary is to be reduced to a purely physical vocabulary. As Griffiths says: “When it can be shown that the explanatory and inductive potential of two schemes overlap, so that neither can subsume the other, then both must be retained and they are not in genuine competition.”<sup>135</sup>

This is not to say that the intentional vocabulary is irreducible in principle, it is just that the needs and interests, which features we take to be salient, of neuroscience differ quite substantially from the needs of everyday interaction with other beings in our environment.

Empirical discoveries, cognitive research and neuroscience, may reveal that there is no single object or single description in neuroscientific language that fully covers a term in Folk Psychology, but this does not imply that the term has to be eliminated from the vocabulary of Folk Psychology. It means only that it has no role to play in a neuro-scientific account of how the mind works.<sup>136</sup> In time it might be that Folk Psychology will revise its vocabulary accordingly, but this is not necessary, especially when the heuristic value of Folk Psychology is considered.

Consider two extremes of a heuristic spectrum: Newtonian physics and witches. Newtonian physics though not strictly literally true still can have immense practical and theoretical value: “Newton’s simple mechanics (the three laws of motion plus the gravitational laws) is strictly false as judged from the perspective of Einstein’s more penetrating mechanics. But it would be entirely too precious and puritanical to insist that the former be eliminated in favor of the latter in all practical contexts”.<sup>137</sup>

At the other end of the heuristic specter are concepts like witches. Although witches no longer play a part in a (scientific or otherwise serious) ontological catalogue the term is still in use, as a negative remark, as a part of a fantasy movie or book, or as something you can dress up as on Halloween, and most importantly it plays an important part in explaining why several thousand women were killed between the fourteenth century and late eighteenth century.<sup>138</sup>

---

<sup>135</sup> Griffiths (1997), p. 199

<sup>136</sup> The pragmatist does not want to preclude that there may be other kinds of scientific explanations of how the mind works.

<sup>137</sup> McCauley (1997), p. 251

<sup>138</sup> It is estimated that about 100.000 were accused of being witches, of these approximately 40.000 were sentenced to death (see <http://www.ub.uit.no/fag/historie/hekser.html>).

## 7. Concluding remarks

Let us return to the question asked in chapter two. What does it mean to say that we are giving different descriptions of the same thing?

In the current popular literature on the relation between the mind and the brain – see, for instance, Schwartz (2002) LeDoux (2002) Damasio (1999), Koch (2004), Quartz and Sejnowski (2002) – there is a tendency to jump between the different modes of explanation. In one breath they use mentalistic terms – they use words like thinking, feeling, emoting etc – and then in the next breath they use the terms of the neuroscientist – neurons, synapses, dopamine, etc. They all make use of the intentional vocabulary alongside with the neuroscientific vocabulary, but they do not give a detailed description of what the relationship between the two consists of. The conceptual analysis theorist might say that this way of trying to explain the relation between mind and brain is logically off the mark and that the product is nonsense.

The confusion here lies in the fact that the vocabulary of neuroscience and the intentional vocabulary mimic each other. Neuroscientists use terms from the vernacular and give them new content, and as people learn more about the new findings of neuroscience, they too incorporate some of the new meanings into their commonsensical framework. This, in turn, makes it hard to evaluate the categories of the different vocabularies. And it makes it more challenging to evaluate whether the reasons for elimination, offered by Griffiths and the Churchlands, are sound. Griffiths notes that “categories that need to be eliminated are more usually ones whose apparent causal homeostasis results from their approximating some other category with real causal homeostasis. Replacing the crude approximation with the category that is doing the actual work is pure epistemic gain. Retaining the old category would not be pluralistic, but merely foolish.”<sup>139</sup> In the present context this observation brings new force to the Churchlandish question of the whether folk psychological terms should be eliminated or not. That is, is it the case that FP just piggybacks on another and better theory, neuroscience, (which is, in essence, what Churchland claims), or is it a way of clustering together properties that are not relevant to the questions asked by neuroscientists, but which deserve answers in their own right?

For the pragmatist there is no way to articulate a principled or *universal method* that gives means to choose one vocabulary over another when the two are in direct conflict. The

---

<sup>139</sup> Griffiths (1997), p. 191

solution to the conflict will vary from case to case. One problem with attempts to employ such a genuine method is that it is always possible to articulate more than a single such method. This, of course, is a consequence of the idea that different vocabularies instantiate different interests and there is no way to tell which of our *interests* in the current domain of inquiry is correct. As a consequence there is little hope, from the pragmatist point of view, in deciding at the present which of the two kinds of descriptions, the mental or the physical, is the one the neuroscientist or the lay persons should use.

However, this is not something that should raise a philosophical worry. On the contrary, as Ramberg points out<sup>140</sup>, it seems to be a consequence of a systematic evasion of having to choose one over the other. It helps bringing the narrative vocabulary and the neurobiological vocabularies closer. The correct stories of our cognitive and subjective capacities are not based on either the intentional, narrative vocabulary **or** the physical, biological vocabulary. They are in this respect equally suited to capture the true story of our mental lives, and none of them are currently disposable, since they both play important and different parts in our cognitive endeavors.

Ramberg suggests that the vocabulary currently used by neuroscientists and cognitive scientists, is best viewed as an interface-vocabulary trying to match up two different domains of inquiry. It is sensitive both to the scientific need of the physically based sciences, such as neuroscience, and other needs and constraints, such as the normative commitments of the intentional vocabulary.

By following Griffiths in his understanding of natural kinds it is possible to argue that the physical, neuroscientific vocabulary and the intentional vocabulary can and should exist side by side since his approach to natural kinds “embodies the realization that there is more structure in the world than can be captured by a single taxonomy of nature.”<sup>141</sup> Currently the intentional and the physical vocabulary seem to have equal right to exist since they explain quite different aspects of human undertakings. However, this does not preclude the idea that they are in some way continuous and dependent on each other.

Griffiths notes that “The current received view is that the dynamics of physical systems can only be adequately captured using a hierarchy of theoretical vocabularies, each irreducible to the vocabularies below it. Irreducibility is guaranteed by the fact that descriptions in on vocabulary can be made true by indefinitely many arrangements of the structures described in

---

<sup>140</sup> Ramberg (2004)

<sup>141</sup> Griffiths (1999), p. 217

lower-level vocabularies.”<sup>142</sup> If this is true of physical systems, then it must be equally true when it comes to the dynamics between vocabularies intended to capture purely physical phenomena and those that display the normative obligations we, as language users, share with those in our community.

If the argument provided above is correct then there is no way to make an isomorphic map of the intentional vocabulary upon the physical (or the other way around). There will always be some residue left that the reduction cannot account for. This again reflects the different interests that guide the different vocabularies of the different levels of inquiry. Some of the questions posed at the top level will be meaningless at the lower and vice versa. This is not to say that some concepts cannot change levels. When we learn more about the brain it is likely that some of the questions that we face at one level of analysis will have answers on another level. But when this happens, this will also, to some degree, represent a reinterpretation of the original question.

The idea that it is possible to make a distinction between what counts as the philosophical and what counts as the scientific part of a problem, seems highly implausible to the pragmatist. The pragmatist thinks that philosophical activity is parasitic on, and a response to, changes elsewhere in culture and society and cannot, as such, be singled out as an autonomous type of inquiry. The idea, therefore, that philosophical questions – in this context, conceptual questions – are prior to empirical and scientific ones is doubtful.

For the pragmatist there is no way to make sense of the notion that some modes of inquiry are typically scientific and others are non-scientific. That is, there is no principled way to distinguish between science as demarcated by ontological responsibility and science as search for systematicity and generality in explanation.

There is, for the pragmatist, nothing “out there” that makes a particular kind of description more “correct” than any of the alternative true descriptions. The world does not by itself provide criteria one can turn to that tell us whether one description of the world is more suitable or closer to the objective, non-human-dependent nature of reality than another. What matters is whether or not the description is useful, and whether it yields a grasp of the world that advances our current interests in the domain of inquiry. However, this should not be taken to mean that the pragmatist endorses the idea that “anything goes”. We are causally constrained by the world, and we seek to articulate real patterns.

---

<sup>142</sup> *ibid.* pp. 215-216

The pragmatist insists that, when it comes to revealing our cognitive interests, the level of analysis should not be separate sentences and their relation, or the logical relation between words in that sentence. Vocabularies as a whole are where such interests are embedded, and where changes in interests come to expression. This makes us realize, for example, that “the fact that Newton’s vocabulary lets us predict the world more easily than Aristotle’s does not mean that the world speaks Newtonian.”<sup>143</sup> It is not the world as such that changes, but what we take to be salient features of the world. And these salient features are what we try to capture, both in the scientific vocabularies and the intentional vocabulary.

### ***Natural Kinds and Trans-theoretic Identity***

Another important implication of this liberal notion of natural kind and conceptual change is that it captures the idea of trans-theoretic identity. It is possible to make intelligible the claim that when we refer to stars they are the same objects as the ancients referred to even though they talked about fixed objects on the crystal spheres of heaven, while we talk about massive incandescent gaseous objects traveling through empty space at very high velocities. The underlying causal mechanism assures that a vital component of the meaning of a term is retained across radical shifts in theory. There is no incommensurability barring the identity of our stars with those of the ancient peoples. This idea is in opposition to the idea that radically different scientific paradigms are incommensurable, and that the basic terms of conflicting paradigms do not refer to the same things.

If natural kinds are thought of to provide some way of classifying similar instances and objects to the same class, without being thought of to provide and support universal laws, but retain some sensitivity to historical changes and essence, it is possible to have radical changes in the scientific changes that do not imply incommensurability. The account of natural kinds Griffiths offers is able to accommodate for such changes in what we take to be the salient features of the world.

If we allow kind-talk to be flexible and to take note of – indeed react to – empirical discoveries and changes in the way we think of the world, scientific progress and intertheoretic reduction would become a much simpler and more commonplace process. When we see that the ontic commitments of our scientific theories are not absolute it will make us more at ease with the idea that theories come and go and that what is perceived as the best explanation in a domain today may radically change tomorrow. The search for natural kinds is an open-ended,

---

<sup>143</sup> Rorty (1989), p. 6

ongoing and empirical epistemological project: “a natural kind *concept* is a concept that it makes sense to seek to clarify through empirical inquiry. Such concepts are ongoing projects of inquiry in which extension and intension are altered to preserve inductive and explanatory power.”<sup>144</sup>

This realization also makes it easier to accept the way theoretical concepts in science are revised as theory change. This, of course, has been a major problem for the philosophy of science for decades. By denying the conceptual analysis approach that takes the current understanding and ordinary use of a term to exhaust the meaning of the given term, reconceptualization becomes one of the chief methods to cash out epistemological gains from theory change, instead of being a headache for philosophers and a way for strong relativists to claim that changes in the way we speak about the world alter it in a somewhat radical and ontological speaking fundamental way. There are changes taking place, for sure, but those are not radically changing *what there is* (understood in an ontological committing way), they reflect our changing interests in the patterns that are salient to us.

## **Conclusion**

The consequence I would like to draw from this discussion is that the current situation does not call for the elimination of Folk Psychology. It might not be the best suited vocabulary for neuroscience, but as the situation is currently there would seem to be little hope of reducing it entirely (reducing it away or what Rorty calls the “disappearance” version of the Identity Theory<sup>145</sup>). The interests that guide the two opposites in the tug-of-war, the scientific use of the intentional vocabulary and the rough and ready use of it we rely on to make useful predictions about our fellow humans, are not the same. The scientist needs a vocabulary that contains terms that are projectable and are useful in making predictions from observed instances to unobserved instances. We non-scientists do not have that need for the intentional vocabulary. In this respect the position I have lined up here is analogous to the “Ryle-Dennett view that when we use a mentalistic terminology we are simply using an efficient vocabulary – the vocabulary characteristic of what Dennett calls the “intentional stance” – to predict what an organism is likely to do or say under various sets of circumstances.”<sup>146</sup> The intentional vocabulary provides an immensely powerful tool for predicting human and other animal (and sometimes other

---

<sup>144</sup> Griffiths (2003), p. 16

<sup>145</sup> See Rorty (1965), p. 176, Reprinted in Hackett (2000), pages refer to the reprinted version

<sup>146</sup> Rorty (1989), p15

organisms') behavior. Churchland is right when he claims that the current way we use Folk Psychology might not be the best for scientific purposes, but it seems odd, from the pragmatist perspective, to articulate this point by claiming that the terms of FP do not refer to "anything real", like caloric fluid or phlogiston.

The upshot, then, is that the pragmatist does not see the current problems in neuroscience as consisting of just conceptual confusion. There are other issues at stake; not only bad application of more or less good concepts – an ailment that might be solved by armchair philosophy and conceptual analysis. Rather, it would be more to the point to characterize the situation as a "mismatch between important human interests and some of the tools with which we are pursuing them"<sup>147</sup> And this is where the philosopher might do some important work, since: "[I]ike the engineer and the lawyer, the philosopher is useful in solving particular problems that arise in particular situations – situations in which the language of the past is in conflict with the needs of the future."<sup>148</sup> The philosopher's job, contrary to the conceptual analysis theorist's claim, is not to sort out the logical tangle (or conceptual confusion) of conflicting concepts, and then subsequently leave the empirical confusions to the scientist, but rather to be able to see the conflict raging between needs expressed in an old vocabulary and the needs expressed in a new vocabulary that try to explain and describe the same relations in the world.

---

<sup>147</sup> Ramberg (1999), p 65 "Dennett's Pragmatism"

<sup>148</sup> Rorty (1995), p. 199

## Bibliography

- Baker, L. Rudder**, 1987, *Saving Belief*, Princeton University Press, New Jersey
- Bechtel, W. / Mandik, P. / Mundale, J. / Stufflebeam, R. (eds.)**, 2001, *Philosophy and the Neurosciences – A Reader*, Blackwell Publisher Ltd., Cambridge Massachusetts
- Bekoff, M. / Allen, C. / Burghardt, G. (eds.)**, 2002, *The Cognitive Animal – Empirical and Theoretical Perspectives on Animal Cognition*, MIT Press, Cambridge, Massachusetts
- Bennett, M.R. and Hacker, P.M.S.**, 2003, *Philosophical Foundations of Neuroscience*, Blackwell Publisher Ltd., Cambridge Massachusetts
- Brandt, Robert (ed.)**, 2000, *Rorty and his Critics*, Blackwell Publisher Ltd., Cambridge Massachusetts
- Carroll, Lewis**, 1939, *Through the Looking-glass*, Reprinted in *The Complete Works of Lewis Carroll*, (1988), Penguin Books, London
- Churchland, Patricia S.**, 1986, *Neurophilosophy – Toward a Unified Science of the Mind/Brain*, MIT Press, Cambridge, Massachusetts
- Churchland, Patricia S.**, 2002, *Brain-Wise – Studies in Neurophilosophy*, MIT Press, Cambridge, Massachusetts
- Churchland, Paul M.**, 1981, “Eliminative Materialism and the Propositional Attitudes”, reprinted in O’Connor & Robb, 2003
- Churchland, Paul M.**, 1984, *Matter and Consciousness (revised edition, 1988)*, MIT Press, Cambridge, Massachusetts
- Churchland, Paul M.**, 1995, *The Engine of Reason, the Seat of the Soul*, MIT Press, Cambridge, Massachusetts
- Churchland, Paul M. / Churchland, Patricia S.**, 1998, *On the Contrary, Critical Essays, 1987 -1997*, MIT Press, Cambridge, Massachusetts
- Dahlbom, Bo (ed.)**, 1993, *Dennett and His Critics*, Blackwell Publisher Ltd., Cambridge Massachusetts
- Damasio, Antonio**, 1999, *The Feeling of What Happens – Body and Emotion in the Making of Consciousness*, Harcourt, New York
- Damasio, Antonio**, 2003, *Looking for Spinoza – Joy, Sorrow, and the Feeling Brain*, Harcourt, New York
- Davidson, Donald**, 1970, “Mental Events”, reprinted in Davidson 1980
- Davidson, Donald**, 1980, *Essays on Actions and Events*, Oxford University Press, Oxford
- Davidson, Donald**, 1984, *Inquiries into Truth & Interpretation*, Oxford University Press,



Oxford

- Davidson, Donald** , 1990, “The Structure and Content of Truth”, *Journal of Philosophy* 87 (1990): 279-328.
- Davidson, Donald**, 2001, *Subjective, Intersubjective, Objective*, Oxford University Press, Oxford
- Davies, Martin and Stone, Tony (eds.)**, 1995, *Folk Psychology*, Blackwell Publisher Ltd., Cambridge Massachusetts
- Dennett, Daniel C.**, 1969, *Content and Consciousness*, reprint 1996, Routledge, New York
- Dennett, Daniel C.**, 1978, *Brainstorms*, Bradford Books, Cambridge, Massachusetts / MIT Press, Cambridge, Massachusetts
- Dennett, Daniel C.**, 1982, “How to Study Human Consciousness Empirically, or Nothing comes to Mind”, *Synthese*, 59, p 159-180
- Dennett, Daniel C.**, 1987, *The Intentional Stance*, Bradford Books, Cambridge, Massachusetts / MIT Press, Cambridge, Massachusetts
- Dennett, Daniel C.**, 1991, *Consciousness explained*, Penguin Books, London
- Dennett, Daniel C.**, 1995, *Darwin’s Dangerous Idea: evolution and the Meanings of Life*, Penguin Books, London
- Dennett, Daniel C.**, 1996, *Kinds of Minds – Towards an Understanding of Consciousness*, Basic Books, New York
- Dretske, Fred**, 1995, *Naturalizing the Mind*, MIT Press, Cambridge, Massachusetts
- Flanagan, Owen**, 1992, *Consciousness Reconsidered*, MIT Press, Cambridge, Massachusetts
- Fodor, Jerry A.**, 1987, *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*, Bradford Books, Cambridge, Massachusetts / MIT Press, Cambridge, Massachusetts
- Fodor, Jerry A.**, 1990, *A Theory of Content and other Essays*, Bradford Books, Cambridge, Massachusetts / MIT Press, Cambridge, Massachusetts
- Fodor, Jerry A.**, 1995, *The Elm and the Expert: Mentalese and Its Semantics (Jean Nicod Lectures)*, Bradford Books, Cambridge, Massachusetts / MIT Press, Cambridge, Massachusetts
- Fodor, Jerry A. / Lepore, Ernest**, 1992, *Holism, A Shopper’s Guide*, Blackwell Publisher Ltd., Cambridge Massachusetts
- Fodor, Jerry A. / Lepore, Ernest**, 1996, “Paul Churchland and State Space Semantics” in McCauley (1996)
- Franklin, Stan**, 1995, *Artificial Minds*, MIT Press, Cambridge, Massachusetts

- Griffiths, Paul E.**, 1997, *What Emotions Really Are*, The University of Chicago Press, Chicago
- Griffiths, Paul E.**, 1999, "Squaring the Circle: Natural Kinds with Historical Essence", in Wilson, Robert A. (ed), 1999, *Species – New interdisciplinary Essays*, MIT Press, Cambridge, Massachusetts
- Griffiths, Paul E.**, (2003), "Emotions as Natural and Normative Kinds" *Philosophy of Science Assoc. 18th Biennial Mtg - PSA 2002: PSA 2002 Symposia*, downloaded from <http://philsci-archive.pitt.edu/archive/00000998/>, pages refer to this version.
- Griffiths, Paul E.** (forthcoming), "Is emotion a Natural Kind?" forthcoming in *Philosophers on Emotion*, Robert Solmon (ed.) Oxford University Press. Downloaded from <http://philsci-archive.pitt.edu/archive/00000566/>
- Hahn, Lewis Edwin**, 1999, *The Philosophy of Donald Davidson*, The Library of Living Philosophers, Illinois
- Krabberød, Anders K.**, 2002, *Truth, Meaning and Triangulation - Reflections on Donald Davidson and Richard Rorty*, Semesteroppgave i filosofi, hovedfag, Universitet i Oslo
- LeDoux, Joseph**, 2002, *Synaptic Self – How Our Brains Becomes Who We Are*, Viking Penguin, New York
- Loewer, Barry / Rey, Georges**, 1991, *Meaning in Mind, Fodor and His Critics*, Blackwell Publisher Ltd., Cambridge Massachusetts
- McCauley, Robert (ed.)**, 1996, *The Churchlands and Their Critics*, Blackwell Publisher Ltd., Cambridge Massachusetts
- Metzinger, Thomas (ed.)**, 2000, *Neural Correlates of Consciousness*, MIT Press, Cambridge, Massachusetts
- O'Connor, T. / Robb, D.**, 2003, *Philosophy of Mind – Contemporary Readings*, Routledge New York
- Quine, W. V. Orman**, 1953, *From a Logical Point of View*, reprint 1980, Harvard University Press, Cambridge, Massachusetts
- Quartz, Steven R. / Sejnowski, Terrence J.**, 2002, *Liars, Lovers, and Heroes – What the New Brain Science Reveals About How We Become Who We Are*, Harper Collins Publishers, New York
- Ramberg, Bjørn**, 1989, *Donald Davidson's Philosophy of Language: An Introduction*, Basil Blackwell, Oxford
- Ramberg, Bjørn**, 1999, "Dennett's Pragmatism", in *Revue Internationale de Philosophie*, 1/1999 – no. 2007 – pp.61 - 86

- Ramberg, Bjørn**, 2004, "Davidson og sjelsvitenskapene", in *Norks filosofisk tidsskrift*, vol 39, Nr. 1-2, s 91-105, Universitetsforlaget, Oslo
- Ramberg, Bjørn**, (forthcoming), "Naturalizing Idealizations: Pragmatism and the Interpretivist Strategy", to appear in *Pragmatism Now*
- Ridley, Matt**, 2003, *Nature via Nurture – Genes, Experience, and What Makes Us Human*, Harper Collins Publishers, New York
- Rorty, Richard**, 1965, "Mind-Body Identity, Privacy, and Categories", *Review of Metaphysics*, XIX, 1 (September 1965), 24-54, Reprinted in Rosenthal (2000)
- Rorty, Richard**, 1979, *Philosophy and the Mirror of Nature*, Princeton University Press, New Jersey
- Rorty, Richard**, 1989, *Contingency, irony, and solidarity*, Cambridge University Press, New York
- Rorty, Richard**, 1991, *Objectivity, Relativism and Truth, Philosophical Papers vol. 1*, Cambridge University Press, New York
- Rorty, Richard**, 1995, "Philosophy and the Future" In Herman Saatkamp (ed.). *Rorty and Pragmatism*. Vanderbilt University Press; Nashville, TN
- Rorty, Richard**, 1997, "Davidson between Wittgenstein and Tarski", *Critica*, Vol. XXX, No. 88 (april 1997)
- Rorty, Richard**, 1998, *Truth and Progress, Philosophical Papers vol. 3*, Cambridge University Press, New York
- Rosenthal, David (ed.)**, 2000, *Materialism and the Mind-Body Problem, (second edition)*, Hackett Publishing Company Inc., Indianapolis.
- Santos, Laurie R. / Hauser, Marc D. / Spelke, Elizabeth S.**, 2002, "Domain-Specific Knowledge in Human Children and Nonhuman Primates: Artifacts and Foods", Printed in Bekoff, Allen and Burghardt (2002)
- Schwartz, Jeffery / Begley, Sharon**, 2002, *The Mind and The Brain – Neuroplasticity and the Power of Mental Force*, Regan Books, NY
- Searle, John R.**, 1997, *The Mystery of Consciousness*, New York Review of Books Press, New York
- Sellars, Wilfrid**, 1956, *Empiricism and the Philosophy of mind*, Reprint, 1997, with an Introduction by Richard Rorty and a Study Guide by Robert Brandom, Harvard University Press, Cambridge, Massachusetts
- Stich, Stephen P.**, 1996, *Deconstruction the mind*, Oxford University Press, Oxford
- Szasz, Thomas**, 1996, *The Meaning of Mind*, Syracuse University Press Edition, NY