PhD dissertation

The cognitive origins of soul belief: Empathy, responsibility and purity

By Gretchen Koch

The Graduate School of Theology and the Study of Religion
Department of the Study of Religion
Faculty of Theology
University of Aarhus
October 2008

Contents

Introduction

As much variety as there is in the religions of the world, there appears to be enough in common for it not to seem absurd on its face to adopt "the study of religion" as a meaningful course of education. And yet the discussion on what constitutes the essential nature of religion continues, whether the question is being posed from a psychological, anthropological, or philosophical standpoint. To approach the question from a cognitive standpoint entails certain suppositions. The most important is that there is a kind of thinking which can be characterized as religious, and therefore learning about the way humans think in general will tell us about their religious thinking as well, in the way that learning about cooking generally will shed light on how to prepare a specific dish. Another is that asking someone directly what they believe will not necessarily give the whole story—they may respond that they believe in one way but behave as if they believe differently, or give different answers when the questions are phrased differently. A third supposition is that human minds might be structured in such a way as to constrain the kind of religious concepts that humans tend to find meaningful. If this is so, then we should adopt a viewpoint of generalization rather than differentiation and look for which religious concepts have wide appeal, then ask what the underlying cause might be. We should also ask whether a particular religious belief is causal in the sense of disposing people toward certain behavior, or whether it is more likely the case that it is parasitic on that behavior and acts more as a post hoc explanation or justification.

Because these are concerns a cognitivist needs to bear in mind, she needs to be prepared to adopt a cross-disciplinary approach. Experimental psychology can give valuable information about how people respond to and use religious concepts consciously and unconsciously, whereas anthropological data will give insight on which concepts are most compelling and widespread, and how they

are used "on the ground." And always in the background linger the philosophical questions that are so often the center of religious concerns: Who am I? Why are we here? What is our purpose? Why should we be good, and what does it mean to be good, anyway? Being able to draw from these disparate fields is a source of both weakness and strength—weakness because it stands in the way of becoming a specialist, attaining as deep a knowledge as possible of a specific area, but strength in that it allows one to look at a religious category from multiple perspectives, drawing on the wisdom of sources from each. For the purposes of this dissertation, that religious category is the soul. It is a loaded term, certainly, one for which there are multiple definitions, and I do not promise to come out with a perfect one by the end. But that is not the problem that constitutes the primary concern for this discussion. Rather, that problem is the question: if we intuitively, perhaps unshakably, perceive humans as having some kind of immaterial essence to them, one which survives death and is the source of their true nature, what effect would that have on our day to day thoughts about each other, especially how to treat each other? Could this effect be seen by examining our common sense, or "folk," ideas about what it means to be human? And, perhaps more disturbingly: if empathy requires us to accurately understand what goes on in other peoples' minds, is it possible that belief in a soul could in some cases act to forestall empathy, rather than aid it?

No doubt, these are monumental questions to ask. It is fortunate for me, then, that the dialog concerning what constitutes religious agency among scholars of the cognitive science of religion is ongoing, and has been for quite some time now. I will use that as my starting point, in hopes that the methodological approach of this dissertation will add something to the conversation. As cognitive inquiries into a specific area often do, this one involves thinkers borrowing from each other's ideas and building on them as well as pulling in thoughts from previously unconsidered sources, perhaps sources who themselves never anticipated that their research would be put to such a purpose, or answer these particular questions. But the questions are important, because asking them helps to dissolve the fog of mystery surrounding concepts central to the very basic elements of faith. Part of the methodological toolkit a cognitivist needs when examining particular religious beliefs is a good foundation on how we think about *non*-religious concepts. If a god or spirit is a religious concept and a human being is not, then what are the important

differences? Our understanding of god and spirit psychology is continually advanced by what we learn about human psychology. And if it turns out that we intuitively treat other humans as if they are walking vessels of concealed divinity, this might reveal a good deal about the give and take between the realms of sacred and profane.

In the course of this dissertation, I am going to argue from a cognitive perspective that some of our intuitive inferences about human agency which we apply to other humans, every day, could be classified as "supernatural" in that they involve that locus of human agency, the soul, being a separate thing from the body. That is, it is a thing which exerts control over the body from the top down. The goal in doing this is to push at the boundaries of current methodology in the cognitive science of religion in terms of how we conceptualize what being supernatural really means and thereby better describe the nature of religious agents generally. I will argue that religious thinking is continuous with folk notions of how the world works, but also challenge the idea that folk notions about what it means to be a person are really that different from what it means to be a spirit.

In Chapter 1, I will provide a sketch of some prevailing theories of religious agency in the cognitive science of religion, and point out overlaps and areas of contention, as well as discussing some of my own criticisms. In Chapter 2, I will examine research on cognitive empathy in order to build a case that humans intuitively attribute human agency, human goals and beliefs, to invisible causes. In Chapter 3, I will argue that this invisible agency is tacitly viewed as an essential entity, divorced from body and society in order to justify beliefs about moral responsibility and free will. In Chapter 4, I will make a case that this agency—the soul—is viewing as something pure and unique, defining the boundary between person and material reality including the boundary that exists between ourselves and all other members of the animal kingdom. By the end I hope to have made a convincing argument that by sneaking in the back door, so to speak, our intuitions about what it means to be human call into question whether perhaps we need a somewhat different concept of supernatural agency.

In his book *Darwin's Dangerous Idea*, philosopher Daniel Dennett recruits the metaphor of skyhooks versus cranes. A skyhook supports from above, hanging on mystery, essentially miraculous. A crane, meanwhile, supports from the bottom-up, building upon lower and simpler layers. I want to suggest that our folk psychology treats the soul as a kind of skyhook that exists in humans, upon which

appear to hang the precious gifts of empathy, responsibility, and purity. And what's more, that this kind of top-down agency is what we ought to think of when contemplating what counts as supernatural.

Chapter 1

What is a supernatural agent– and why does it matter?

One of the central questions for scholars of religion is the matter of religious agency—what makes an agent religious, and what makes an agent the object of religious belief? In this chapter I am going to be paying most, but not all, of my attention to the latter question. I will cover the different concepts of a supernatural agent in cognitive theories of religion, and compare them against each other with concern for theoretical workability. In order to do this, we will need to examine some of the typical candidates offered up for this category, and determine what critical aspects they have in common which would merit them being lumped into a group called "supernatural." That will also require, of course, an examination of what the term "supernatural" means itself.

For a person coming from a background of monotheistic religion, the first supernatural agent who comes to mind will most likely be "god." Or, more likely, they will say "God," or "Allah," or "Jehovah." Asked to think a little further, this person will note that members of other religions may worship more than one god, or may even have a pantheon of gods they worship or fear. Or the term "supernatural" might send them off on another track and they will mention ghosts, spirits, and other characters associated with the paranormal. The questions one can ask about supernatural agents are virtually boundless: Where do they exist—here on earth, in some other plane, or both simultaneously? Do they have personalities like humans? Have they ever *been* humans? Have they always existed, or were they born, and do they have the capacity to die? Are they good or evil or some mix of the two, and do they have opinions about the morality of humans? Do they have access to information we don't? And if so (or even if not), why do they care?

A central idea that I will be arguing throughout this thesis is that we think of supernatural agents very much in the same way that

we think of human agents—that our intuitive, everyday thoughts about how human beings think and behave are actually very much like what we tend to think of as "supernatural," and vice versa. But in order to do that, I will need to explain how cognitive theorists differ in their view of supernatural agency from how other scholars of religion might see it.

A cognitive theorist is interested in how the architecture of the human mind makes some concepts appealing, and others boring or difficult to grasp. In the context of religion and agency, a cognitive approach would require looking at how concepts of agency spread or stagnate (their epidemiology), what aspects of our thinking about agency make certain representations seem intuitive or counter-intuitive, and why some inferences about the way supernatural agents think and behave are easier than others. From examining these questions, we hope to reach some illumination about those already mentioned. If it is discovered, for example, that people have a hard time imagining spirits who lack a beginning but no trouble imagining them ending, then we will be skeptical about the likelihood of finding a society in which all spirits come into existence as babies but never die. In order to conduct this investigation, however, it will be necessary to clarify some terms.

1. What is an agent?

It seems like a basic question, but the matter of what constitutes an agent is trickier than it might first seem. In filming the *Lord of the Rings* trilogy, the directors faced the problem of how to portray thousands upon thousands of people fighting furiously on a battlefield without having to hire multiple thousands of actors. In order to do this, they asked the animators to design a program which would produce a kind of agent—an "actor" with artificial intelligence whom they could order to walk in a specific way, fight in a specific way, and behave in specific ways when confronted with obstacles, other AI agents, or open spaces. In doing this, they were able to portray a battle in which movements were choreographed without being choreographed. Nobody directly tells *this* orc to attack *that* human, but the orc is programmed to know that when he detects that a human is near, he should attack—and *this* human happens to be nearest. In this way, the process is deterministic without being guided; a rule-based heuristic. To make things even more complex, they may have also factored in rules related to what an actual soldier in combat might

find important. If my opponent is much larger and stronger looking than I am, then I might choose to fight someone else. The animators reported that to their astonishment, upon reviewing the battle sequences they sometimes discovered "deserters"—individual orcs or humans dashing off away from the battle and toward the trees. It turned out that the rules they had given the soldiers to follow required that if there were no opponents in the immediate vicinity, they should run forward. The result of this was that soldiers who had begun the battle facing away from their opponents would simply run without stopping, the action fading behind them. This kind of reaction might seem entirely rational to human beings with a strong desire to survive, but the animators didn't want cowardly orcs on their hands, and had to amend their behavioral algorithm.

I refer to these soldiers as agents, but are they really? Generally we think of an agent as having much more awareness, much more control over his or her actions than an angry animated orc. An agent is an entity to whom we ascribe beliefs, goals, and emotions. When eventually computer scientists create a form of artificial intelligence which passes the Turing test, the debates will begin in earnest on whether to consider that entity a person, equivalent to a human agent in terms of freedoms and responsibility. Already some advocates are arguing that the level of agency exhibited by a chimpanzee entitles it to be viewed as a person.[1] As will be discussed later in this paper, we have certain expectations about what conspecifics—fellow human beings—are capable of doing, and what they are prone to do. These expectations are based on our own, first-hand experience of what it's like to be a human, and are projected onto people who look and behave similarly to ourselves. We know what it's like to have dispositions toward things, desires for things, and beliefs about things, and so we expect other humans to do the same. Therefore, our idea of what it means to be a human agent has certain intuitive parameters. A person with no beliefs is hard to imagine. Equally difficult to imagine is a person with no goals or desires. Even agents with artificial intelligence have goals—because they are programmed to have them. But human agents have goals of their own, often very elaborate ones. Our ability to form these goals, backed up by equally elaborate systems of belief, is something most humans don't spend much time pondering. It is also something we

[1] Most notably the Great Ape Project. For more information about this movement, see their web site at http://www.greatapeproject.org/

apply effortlessly to supernatural agents as well, which leads us to ask...

2. What is a supernatural agent?

There simply isn't enough room here for me to provide a complete history on cognitive treatments of supernatural agency, so I will not attempt to do so. Rather, I will present a selective sketch of the progression of cognitivist thinking on the subject, in order to provide a background for the discussion on agency and mind-body dualism which will comprise the remainder of the dissertation. Hopefully, the dissertation will itself become part of that ongoing discussion.

The year 1980 saw the publishing, in the journal *Current Anthropology*, of anthropologist Stewart Guthrie's theory of religion as systematized anthropomorphism using *Webster's Third New International Dictionary's* definition of the term, "the interpretation of what is not human or personal in terms of human or personal characteristics." In his article Guthrie acknowledged the long history of preceding thinkers who have regarded animistic thinking as central to religion, but set his own theory apart by emphasizing his view of religion in intellectual terms as a means of recruiting anthropomorphic notions of the universe in order to explain that universe. Humans, he said, are often confronted with ambiguous phenomena. In order to interpret those phenomena, they explain the world around them as if otherwise mysterious entities possessed properties that make them human-like. Humans are chosen as a model for this interpretation because they are typically the most important feature in an individual human's atmosphere in terms of establishing what is most important for that individual to know and do. They are complex and multi-faceted and therefore generate a variety of phenomena, and given the highly social nature of human life, are very likely to be present at any given time. For Guthrie, religion is a method of attempting to explain and control the world, like common sense or science, but is distinguished by the fact that it invokes anthropomorphic thinking for that explanation. Religion explains the world in terms of the behavior of human-like but non-human agents, such as gods and spirits. Every religion is populated by these beings, as much as theologians might try and de-personalize our thinking about gods in order to elevate them and make them more perfect. Guthrie spends a good bit of time in the article describing how theologians such as Paul Tillich have striven to drive anthropomorphism out of the concept of God entirely, but says

that this effort leaves religions anemic: "My assertion . . .is that human-like but non-human beings both are universal in and characterize religious belief and that they not only are made plausible by experience, but are made (contingently) persuasive by it."[2]

There are a few other important things to note about Guthrie's theory besides its emphasis on anthropomorphism. For one, it is cognitive in the sense that Guthrie describes religion as a rational way to interpret existence, rather than a form of neurosis (Freud) or wish fulfillment and relief of anxiety (Malinowski). Guthrie portrays religion as logical, but also quite natural. It is not a unique form of thinking—it recruits our already existing tendency to see human-like forms and behavior everywhere we look in order to make clear what would otherwise confuse. This anthropomorphism is, as Guthrie puts it, a kind of "reasonable illusion."[3] Religion, then, is motivated by practical and cognitive needs rather than "cultural" or emotional ones. He rejects the functionalism of Durkheim on the grounds that it makes religion strictly about human social relations because the gods do not exist, but Guthrie's theory does not require that they exist either—simply that they are perceived as existing, and that this existence is viewed as explaining something important about the universe.

Guthrie notes that E.B. Tylor defined religion substantively as "the belief in Spiritual Beings"—that is, a way of conceiving of the world. Another notable aspect of the theory as presented in this article is that in it Guthrie discards the label "supernatural" as useful in characterizing agents central to religious belief. The term, he says, is a Western folk category and not a cross-cultural one. He quotes anthropologist Benson Saler approvingly: "It is misleading to refer to beings or powers in non-Western world views as 'supernatural' when supernatural is not salient in native thought" and Robin Horton as suggesting that "supernatural" is "no more descriptive of African traditional thought than of Western lay understanding of nuclear explosions."[4] So while non-human and formerly human agents are central to Guthrie's theory of religion, specifically identifying them as "supernatural" or "superhuman" is not. In a later paper, Guthrie would again join Tylor in referring to these entities as "spiritual beings."[5]

---

[2] Guthrie 1980: 185

[3] Ibid., 181

[4] Ibid., 185

[5] Guthrie 2008

Guthrie's book *Faces in the Clouds*, published in 1993, put forth an elaboration of his theory of religion as anthropomorphism. By this, as Guthrie puts it, he means that religion "makes nature humanlike by seeing gods there."[6] Guthrie maintains that all religions have a god or gods; by the term "god" he is using the Oxford Universal Dictionary definition of the term: a "superhuman person . . . worshipped as having power over nature and the fortunes of mankind."[7] He notes that gods may also be reviled, and that "gods may be only marginally superior to humans, and may grade into spirits, ghosts and demons, which may be inferior to humans."[8] The anthropomorphism to which Guthrie refers is a form of *pareidolia*—a psychological phenomenon involving the perception of patterns in random visual or auditory stimuli. Pareidolia is what causes white noise to sometimes sound like music, but it is also what makes faces appear in the clouds, as Guthrie's chosen title notes.[9]

Of course, part of the definition of pareidolia is that it entails that the perception of said patterns is erroneous, which Guthrie notes as the reason that theologians refuse to countenance anthropomorphism as the essential nature of religion. Guthrie doesn't simply assert that humans humanize their deities, but that religion is *itself* an act of humanizing—a form of seeing agency in nature which is taken very seriously, cloud faces that have personalities, interests, and influence, and who care about human behavior. These "culturally postulated superhuman agents," as Lawson and McCauley labeled them in 1990,[10] are identified by the human capacity to perceive their images and behavior in the natural world. Guthrie's portrayal of these agents rests on a tendency in humans to have a hair-trigger perception of agency. Agency is the attribute which Guthrie defines as human, which humans then project onto the world around them. His definition, however, is hostile to religious belief because it relies on the assumption of their falsity. "Anthropomorphism," as Guthrie puts it, "by definition is an overestimate of likeness. It is not simply an assumption of likeness since, in fact, many things *are* like us in

---

[6] Guthrie 1993: 177

[7] Quoted in Guthrie 1993: 241

[8] Ibid., 241

[9] The earliest known use of this term appears to have been in a *Skeptical Inquirer* article published in 1994, so Guthrie does not avail himself of it in *Faces in the Clouds*. Nevertheless, it is a fitting term for what he describes. Source: http://www.wordspy.com/words/pareidolia.asp

[10] Lawson and McCauley 1990

various ways. It is a mistake about likeness. We can label it as anthropomorphism only after seeing it as error."[11] Guthrie acknowledges philosophers who attempt to remove anthropomorphism from religious dialog by using terms like "ground of being" and "the Unconditional," but notes that such attempts simply fail to reach the vast majority of religious believers. This is because, he says, a god who has been stripped of all human characteristics simply doesn't resonate with lay believers. They want something more "substantial"—that is, something more like a person. As for the question of why we anthropomorphize in the first place, Guthrie's answer is that it is functional: "I believe we anthropomorphize because we perceive the world in terms of our interests, which usually involve humans."[12]

The key points to draw from Guthrie's theory of religion are that a) humans anthropomorphize in order to try comprehend the world around them, b) religion is a systemized form of anthropomorphism c) symbolic communication with humanlike beings perceived in nature (gods, spirits, and so on) is a common feature of all religions, which means that d) religious thinking is continuous with other sorts of thinking, rather than *sui generis*. This latter point is especially important to keep in mind when considering theories of supernatural agency which I will be addressing next.

Anthropologist Dan Sperber has done significant work on the epidemiology of religious beliefs—that is, he pays attention to those aspects of a belief which make it likely to spread quickly and easily amongst members of a population. A key factor for beliefs concerning supernatural agency, Sperber says, is their *relevance*. Sperber and Diedre Wilson developed relevance theory in 1986 to describe the context in which we interpret utterances—the correct interpretation of an utterance is that one which permits the most new implications to be drawn from it, using the least amount of information. In relation to supernatural agents, relevance suggests that those concepts of such agents which portray them as having maximal implications for human behavior should be the ones which are most often transmitted from one person to the next. In order to be interesting, a god or spirit must be interest*ed*—it must care about what people are doing. If a god or spirit is omniscient, but is not in any way concerned with human life, is not very compelling. A god or

[11] Guthrie 1993: 183

[12] Ibid., 185

spirit who "knows when you've been bad or good," on the other hand, is one with whom we might be able to formulate some kind of exchange. Even if the spirit of this particular forest doesn't know anything about what is going on in the forest ten miles away, the spirit is still important if it knows that your neighbor stole food from you, or when your tribe might win in battle against another, or how to win the affections of someone you desire.

In his book *Why Would Anyone Believe in God?,* psychologist Justin Barrett strives to convey to a general audience exactly how natural religious belief and behavior is. It is not surprising, then, that the first line in his first chapter on belief in gods is "Belief in gods requires no special part of the brain."[13] Barrett goes on to explain that we come to believe in gods in very much the same ways we come to believe in other people—that we use the same "agent detection device" that has come in handy over millions of years of human evolution. Being able to detect the presence of other agents, even if we can only partially see them or can't see them at all, is of strong importance for human survival since often other agents may wish us harm. Barrett harnesses this concept to advocate for what he calls a hypersensitive agent detection device, or HADD. Humans are not the only animals who have this sort of device—when a dog jumps and growls at the sound of crashing thunder outside, it is exercising its own form of agency detection. The device is "hypersensitive" because we often may need to guess at the presence of another agent before we can have its presence fully demonstrated to us. It's better to suspect that someone is following you in the forest when you hear leaves crunching behind you, rather than your mind waiting until it has full visual confirmation. This approach does result in false positives, but as anthropologist Stewart Guthrie has noted, it's better in terms of survival to mistake a tree stump for a bear than the other way around. While we do often see faces in the clouds (and in flowers, water, tortillas,[14] doors,[15] and the occasional grilled cheese sandwich[16]), behavior alone is often sufficient to convince us that there is some*one* around, rather than merely some*thing*. Over time, people learn to discount a lot of false positives in their perception of agency around them—after all, leaping up in fear every time a leaf

[13] Barrett 2004: 21

[14] http://www.jesustortilla.com/

[15] http://www.jesusinthedoor.com/id2.html

[16] http://news.bbc.co.uk/2/hi/americas/4034787.stm

falls to the ground can become quite counter-productive. But certain invisible agents—ghosts, spirits, and gods—seem to stick around in the human consciousness, and get transmitted to other members of the group. What happens between the initial perception that a crunch in the forest might be an agent, and the establishment in a community of the idea that spirits inhabit the forest and want us to do things for them?

As Barrett describes it, HADD kicks in when we perceive non-inertial, goal-directed movement on the part of an object and seek to explain said movement. In other words, if we can't see that an object has been pushed or otherwise acted upon by something else in order to move, and it appears to be moving for some specific reason (a goal), then we tend to attribute agency to that object. As Barrett says, "If a person or other known agent clearly accounts for the action, the object that moves need not be identified as an agent. If no such known agent is responsible for the movement, the object itself becomes a prime candidate for agency. Thus, we treat remote-control toys, cars, and computers as agents only when they 'act' in a way that challenges our own agency (or the agency of another person)."[17] To put it simply, the appearance of goals is what activates HADD. For a vivid illustration of how HADD works in this way, we can contemplate the fact that in a famous study from 1944, both adults and children attributed motivation and even gender to animated triangles and circles on a computer screen.[18]

3. Counterintuitive agents

Pascal Boyer is Henry Luce Professor of Individual and Collective Memory at Washington University. As might be guessed from his title, a primary concern for Boyer in developing a theory of religion is, as with Sperber, determining what makes religious concepts "catchy"—in a world where cultural concepts are swimming through and between minds in huge numbers, what makes some more likely than others to be conceived, maintained in memory, and spread to others in one's community?

In his 2002 book *Religion Explained*, Boyer articulates his own theory of how religious belief and behavior is "natural," and how it stems from the way the mind works. He starts by acknowledging

[17] Barrett 2001: 32-33

[18] Heider & Simmel 1944

long-standing previous theories of what religion is "about"—explaining the world, allaying anxiety, supporting morality, gullibility—and stresses that while there are elements of truth to them, they don't explain why religion is produced by the kinds of minds that humans have. Boyer's approach is from the inside out, attempting to convey how religious beliefs could arise from internal cognitive processes. With a background in both psychology and anthropology, he attempts to formulate a functional theory of religion, one which is not satisfied with documenting the transmission of concepts between minds, but also their production within the mind.

In order to do this, Boyer describes a method of identifying supernatural concepts, and what makes them memorable. He argues that popular and compelling religious concepts can actually be predicted by determining whether they conform to a specific formula which he has developed—they "centre around a small catalogue with recurrent features."[19] This catalogue is established by ontological categories in the brain which are present from very early on in life. In order to qualify as supernatural, Boyer explains, a concept must fulfill the following requirements: 1) it must activate a set of ontological categories, 2) it must specify information which violates the inferences of one or more of these categories, but otherwise 3) it must retain activation of other inferences of the ontological category which are not violated. I will go through these one at a time.

What are ontological categories? They are categories into which we intuitively place things we encounter in the world. Boyer lists them as PERSON, ANIMAL, ARTEFACT, PLANT, and NATURAL OBJECT. Each category brings with it a set of automatic expectations. The category ANIMAL, for example, brings with it the expectation that the young of the animal will be of the same species that it is. We expect a cat to give birth to kittens, not puppies—this is an intuitive biological expectation. In addition to intuitions about biology, Boyer explains, we also have an intuitive physics and an intuitive psychology. Not all of these necessarily apply, of course. A mountain, for example, would fall into the category of NATURAL OBJECT, but we don't have an intuitive biology of mountains.

Let us imagine, however, a volcano which must have its anger soothed by consuming human sacrifices. This involves a transfer of biological and psychological traits to the category of NATURAL OBJECT, making it special. A ghost activates the ontological

[19] Boyer 2000

category of PERSON, but the fact that it can float through walls violates our intuitions concerning physics. A violation may entail either a breach or a transfer of traits that go with an ontological category. For example, a chair which disappears is a breach of our intuitive physics regarding ARTEFACTS. A transfer, on the other hand, entails applying the inferences regarding one ontological category to a member of a different category. For example, if the chair knew what you did last summer, that would be transferring to the chair a notion from our intuitive psychology, which is associated with the PERSON category. Using this formula, one can create a template for possible counterintuitive concepts by combining different ontological categories with either breaches or transfers of different inference systems. A statue which weeps, for example, is an ARTEFACT [ontological category] which violates our intuitive biology [inference system] by transferring a trait from the PERSON category [type of violation]. There are fifteen possible combinations of these three characteristics, which makes for a rather short list of kinds of counterintuitive concepts.

Lastly, there is the fact that the concept still activates all other inferences which would normally be made about its ontological category. The statue may weep, but we still expect it to stay where it is rather than flying up into the air. Except for the biological violation, we still mentally place it in the category of ARTEFACT and draw the normal set of inferences which would go with that. We don't have to stop and think "Now wait, does the statue still weigh the same?" Our inference systems do that work on their own. "People tacitly represent spirits as having minds," notes Boyer. "That is, spirits are presumed to perceive events, to have beliefs and form intentions on the basis of their beliefs, and so on."[20]

Not all supernatural concepts become successful reproducers, and manage to get themselves into the long-standing beliefs of a religious community. There are a variety of reasons for this, and one is that a concept which is counterintuitive in too many ways is hard to remember. A table which laughs is interesting and may catch in the memory. A table which laughs, flies, exists only on Tuesdays, and grew out of the ground is also interesting, but less likely to be memorable because there are too many violations involved. For this reason Boyer stresses that only *minimally* counterintuitive concepts are likely to be preserved in the memory and passed on. As for why

[20] Ibid.,198

other concepts are not often recruited from the catalogue of the supernatural even though they may also be minimally counterintuitive, Boyer's answer is that the ones which are most relevant to social relationships are most likely to be successfully transmitted. As it turns out, the ones most relevant to social relationships are those which start from the PERSON template, such as gods, spirits, witches, ghosts, and ancestors—they are the ones which activate our "theory of mind" inferences, our folk psychology, even though they may have breaches of physical and/or biological expectations. Also popular are supernatural concepts of ARTEFACTS with psychological properties—that is, objects which are man-made but yet also trigger theory of mind. A statue which speaks would be an example of this.

Mickey Mouse is an example of a counterintuitive concept, according to Boyer's theory (ANIMAL + transfer of psychological expectations), and Mickey is certainly a memorable character. But the reason why there is no religion centered around him, Boyer explains, is because of our evolved mental mechanisms for processing information and engaging in cooperation.[21] "Religious concepts," he notes, "are those supernatural concepts that *matter.*"[22] And what has mattered most to humans throughout their evolutionary history has been functioning in a social group.

The thing that makes concepts of gods and spirits so compelling, Boyer argues, is that they are "full access strategic agents." This term needs some unpacking. First, he establishes the difference between strategic information and non-strategic information. Strategic information is not necessarily important in the sense of being life-saving (although it certainly could be), but it is information that is of high interest in human social relationships. Boyer's definition is "Strategic information is the subset of all information currently available (to a particular agent, about a particular situation) that activates mental capacities that regulate social interaction."[23] The contents of my closet are strategic information if it means the closet contains a) something I stole from a friend, or b) a secret lover who is hiding, or c) the corpse of my enemy, etc. Because gods and spirits have advantages such as being immaterial or omniscient, they may have special access to important information about our social world which we do not. Boyer is careful to note that

---

[21] Ibid., 203

[22] Boyer 2002: 155

[23] Ibid., 204

this does not necessarily mean that they are wiser than us—there are spirits which are not very bright and can be tricked—but rather that they seem to *know* more.

As human-like thinking beings, gods and spirits also have desires. These desires, and the human ability to satisfy them, can dramatically shape the human social world. The proper sacrifices and rituals must be made in order to ensure that hazardous activities in life may be conducted safely—that a woman will not die in childbirth, that a hunt will be successful. Supernatural beings are also capable of punishing those who do not cooperate with others in their community, those who have stolen or practiced witchcraft to harm their neighbor. If a god knows everything, then it follows that he knows that the woman next door is dying of some disease we don't understand, and what's more he knows why. It may be because he was displeased with the goat we sacrificed to him last week, or it may be because the woman disobeyed her parents, which displeased him and caused him to curse her. A god or spirits who not only is aware of when humans are acting immorally but has the power to reward or punish them is very important concept for them to know about, indeed. Boyer uses the example of anthropologist Evans-Pritchard studying the Zande people of Sudan. When roof of a mud house collapsed, the Zande explained the event in terms of witchcraft. Evans-Pritchard's suggestion that the roof's collapsing was caused by termites was brushed aside—the people understood termites, and that houses fall apart eventually. What they wanted to know was why the roof collapsed at that specific *time*, with certain people sitting underneath, rather than earlier or later. They invoked supernatural agency as a means to explain, but not simply for general explanation—for explanation of a strategically important event.[24]

As a human, then, I perceive that the most control I have over the world is at my fingertips. If something is within the realm of my ability, and I desire it, then I can behave in such a way as to bring about the desired result. If it is not within the realm of my ability, my second choice would be to ask for the assistance of another person—another person is likely to comprehend my request, to be capable of fulfilling it, and is susceptible to my attempts to convince him or her to do what I want. If what I want is beyond the realm of human ability, or there aren't any humans around who can comprehend my request or have the desire to fulfill it, then I can appeal to the gods.

[24] Ibid., 14-15

Since gods are human-like, they may be disposed to respond to my requests in a similar manner to humans, but also able to fulfill larger requests—making rain in a time of drought, for example, or saving the life of a very sick relative. Given this supposition, one should not expect to find that concepts of gods and spirits who know no more than humans, and do not care about human behavior, will be commonly found in religious belief.

Communities all over the world are faced with the challenge of getting along. When members rely on each other, they must be able to determine who is trustworthy and who is not, who they should help and who will help them. The question of how altruism evolved has been an ongoing area of concern for biologists and anthropologists over decades, from trying to figure out why worker ants toil in service of the queen to determining why a human soldier would sacrifice himself on the battlefield. From an evolutionary standpoint, it stands to reason that animals would be willing to help out those who are closely related to them, as this makes it more likely that at least some of their genes will be passed on to the next generation. Those con-specifics who are most closely related will be favored, as this presents the greatest chance for the individual's genes to be perpetuated. This model, called "kin selection," was formulated by evolutionary biologist William Hamilton, and the idea behind it can be summed up by JBS Haldane, who when asked if a man should lay down his life for a brother replied jokingly "for two brothers, or four cousins."[25] Robert Trivers, also a biologist, would later develop a theory of reciprocal altruism to explain sharing and cooperation even among even among animals who are not closely related, but who have the capacity to remember which con-specifics have helped them in the past as opposed to who has not. Under such circumstances the animal can behave kindly to another with the expectation that such kindness will be repaid in the future. The theory of reciprocal altruism became famous in large part because in a evolutionary world where individuals are expected to behave strictly to selfishly further their own interests, the presence of unselfish behavior seemed mysterious, in need of an explanation. Trivers argued that over our evolutionary history the practice of reciprocal altruism became part of human nature.

But even combined, kin selection and reciprocal altruism do not give a full explanation for why humans should practice altruism

[25] http://www.guardian.co.uk/books/2005/aug/27/featuresreviews.guardianreview9

and cooperation. In his book *Passions Within Reason*, economist Robert Frank theorized that the passions play a major role as well—shame, guilt, indignation, spite, compassion, and other such emotions serve to regulate human helping behavior by providing strong incentives and disincentives. These hard to fake responses signal a willingness to punish freeloaders, and can inspire trust in others when the emotions of condemnation are directed at freeloaders who are harming them rather than oneself—that is, sympathetic condemnation. If you are wronged and upon learning of it, I become angry on your behalf, this can act as a signal of trustworthiness on my part. Not only am I upset when someone betrays my trust (which possibility itself is a deterrent to prospective betrayers), but also I am indigent on your behalf, which shows that I have an interest in your not being betrayed as well. Frank's argument is that emotional responses, when viewed in this light, are actually very rational in terms of establishing trustful relations. Here is Boyer's formulation of the idea:

> First, people should be prepared to punish cheats even if it is costly. So they should have powerful emotions that help them disregard the cost. This is clear in many everyday situations. The anger triggered by queue-jumpers and parking space thieves is quite disproportionate to the actual damage they inflict. Second, we should be outraged when cheating is not punished even if we did not incur any cost because the existence of 'suckers,' by making cheating a viable strategy, is a threat to our own safety. So people should punish cheats, not just because they feel outraged but also because others will feel outraged if they do not. This passion too is a constant in human interaction. The queue-jumper makes you feel angry even if you are standing in another line.[26]

What does all of this have to do with gods and spirits? According to Boyer, the answer is that they are connected to the moral emotions because a full access strategic agent has access to all dimensions of a given behavior, *including* its rightness or wrongness. If I believe that it was wrong for my neighbor to steal my car, and God knows everything about the theft of the car, then it stands to reason that God also believes that stealing the car was wrong. Now I not only have my own moral outrage, but also God is on my side and is thus outraged

[26] Boyer 2002: 213

too. If I am the one who has done something wrong, and God has full access and therefore knows all about it, then this will compound my guilt. Boyer's point is that we have already evolved the tendency to behave morally, but that gods and spirits are parasitic on this tendency because they support our moral conclusions—they appear to give them an origin which is understandable, whereas our mental machinery which actually produced the conclusions is beyond our access.

Another theory of how supernatural agents function to bolster group cohesion is costly signaling. A "signal" in this case is an expression or behavior which makes a statement about the character of the person in question. Signals become costly if making them involves heavy demands on the performer, whether those demands are a matter of health, finances, time, opportunities, and other things which are of value to the signaler. Examples of costly signaling can include tattoos and other painful and/or elaborate body ornamentation or engaging in a physically taxing ritual such as the Sun Dance or ceremonial fasting. A materially costly signal could entail a generous donation to one's church, or the sacrifice of cherished food items or favorite possessions. A hard-to-fake emotional display can be considered a costly signal, especially if it is perceived as being hard on a person emotionally or psychologically. Ritual behaviors which involve such displays indicate something particular about the people who participate in them, which is that their allegiance to the gods or spirits being honored in that ritual is genuine. A non-believer could participate in the rituals, but his/her emotions may not be as convincing as those of someone who is a passionate and fervent believer. Performance of rituals, as well as living an obviously devout life, gives religious adherents the opportunity to send a message to others in their community that "I believe in the same things you do. I worship the same gods you do. I can be trusted." In this way, such participation becomes like a secret password for respected membership in the group. Knowing the doctrines of the religion of your community becomes a way to prove devotion to both the group and the otherworldly beings to whom it is devoted. A person can make costly signals which communicate dedication to other groups, such as a military unit or a fraternity, but religious costly signaling may be especially effective because it requires members to signify devotion to entities whose existence will be non-obvious to a non-believer. Scholar of religion Joseph Bulbulia theorizes that this factor creates a tendency to make doctrines more particular and less

believable, so that the costliness of the signals required is increased and therefore their status as evidence of devotion is likewise enhanced. If the doctrine of the religion in question is exclusivist, then making costly signals indicating adherence to one religion will preclude displaying allegiance to another, thus enhancing their appearance of veracity.

4. Theological (in)correctness

There are elements of Boyer's theory of supernatural agency which I think do not quite work, but I am not going to address them just yet. First, I want to discuss another take on how to interpret beliefs about religious agency. "Theological Correctness: Cognitive Constraint and the Study of Religion" is the title of a 1999 article in which Barrett addresses a problem with attempting to learn about people's religious beliefs simply by asking them straight out: they do not always act as though the things they *say* they believe are true. Depending on how we are asked about our religious beliefs and how long we are given to ponder the question, we may give contradictory answers. Why is this so? Barrett argues that our cognitive processing acts to limit our "online" representations. By "online," he is referring to a kind of thinking which is implicit, non-reflective, and rapid. If we imagine a continuum of complexity in cognitive processing, then online thinking is at the end of simplicity. At the opposite end are explicit, reflective, more slowly developed models of thought. This continuum exists in other domains besides religion—science and philosophy are at the complex end, together with theology. Barrett uses the term "theological correctness," or TC, to illustrate how more reflective thinking causes people's answers about religious beliefs to more closely match the dominant theology of their religious, whereas when they are *not* given the chance to reflect their reports tend to be more simple, "gut-level" and basic.

In support of this argument, Barrett cites evidence such as an experiment in which he was involved, entailing that people of several different faiths in the United States and India were asked to use their concepts of divinity to process and remember short stories about one of a number of gods, depending on their particular denomination. One such story was:

> A boy was swimming alone in a swift and rocky river. The boy got his left leg caught between two large, grey rocks and

> couldn't get out. Branches of trees kept bumping into him as they hurried past. He thought he was going to drown and so he began to struggle and pray. Though God was answering another prayer in another part of the world when the boy started praying, before long God responded by pushing one of the rocks so the boy could get his leg out. The boy struggled to the river bank and fell over exhausted.

"Subjects whose basic god concept is anthropomorphic," notes Barrett, "might infer that God finished answering one prayer before attending to the boy. If so, they would tend to misreport that this is actually what the story said. Subjects using a god concept in which God performs many tasks simultaneously would be unlikely to report that the story said God finished answering one prayer before saving the boy It is just as likely that God rescued the boy *while* answering the other prayer."[27] As it turned out most subjects did in fact anthropomorphize gods when confronted with these scenarios. They tended to apply a concept of god with a limited attention span, fallible access to knowledge, and a physical location. When asked to reflect upon the properties of the gods using a questionnaire, however, they would say that gods have limitless attention, access to all possible knowledge, and no location in time and space—that is, a more theologically correct depiction of gods. When considering god concepts indirectly, instead of being asked straight out to list the properties of gods, they followed a shortcut to a less cognitively demanding representation. Barrett compares this to scientific knowledge, which can also be quite complex but which we can intuitively avoid elaborating by continuing to use expressions such as the sun "setting" when we know quite well that it actually is in one position, and the planet revolves around it.

Barrett's argument is that theology involves deduction, starting with general principles and then applying them to specific religious concepts. But when people are using "online" thinking, which is non-reflective and tacit, they portray those concepts differently by using abductive thinking. A theological concept, for example, would be that if God is omniscient and knows everything there is to know, then it naturally follows that he knows the titles of all of the books on my bookshelf, since those book titles are a sub-set of the category of "everything." But when not in "online" mode, our

---

[27] Barrett 1999: 329

abductive reasoning makes us more prone to trying to derive a general principle from specific circumstances, so that the situation makes sense in light of the general principle which has been constructed. The story of the Zande mud hut's roof collapsing would exemplify this—in light of the need for the Zande to have a *moral* reason for the collapse, an agent (in this case, a witch) was inserted into the explanation in order to make it satisfying.

So Barrett argues that religious groups maintain two levels of conceptualization—the theological and the basic—even though these may at some points be contradictory. In order to explain why this happens, he refers to the same ontological categories that Boyer does (artefacts, animals, persons, etc.) and the intuitions we draw from them, such as naïve biology and psychology. Even though humans can process the idea of elaborate, complex religious concepts when asked to do so, he says, we revert to our "folk" concepts, which are produced by these categories and inferences, the rest of the time. This intuitive knowledge constrains online conceptions of supernatural agents, says Barrett: "since gods are intentional agents, willful initiators of action, the default assumptions used to conceptualize them is intuitive knowledge about agents, or naïve psychology. So it is no surprise that the basic concepts of participants in the experiments sketched above were found to reflect many of the assumptions of human agents—the prototypical intentional agent."[28]

From a scientific viewpoint, then, the notion of theological correctness and incorrectness is highly important in terms of getting an accurate perception of the religious views in a culture, as well as which elements of those views they most likely have in common with other cultures. For example, if Barrett's experimentation holds out as accurately representing online beliefs about the nature of gods and spirits due to cognitive constraints, then we should expect to see the same or similar inferences made by devotees of other religions across the globe. The notion of cognition shaping religious concepts in this way means that we should look for two parallel levels of religious thinking in every community that can be characterized as religious. Viewing religious systems in this way may also serve to explain why the theological beliefs of a community seem inconsistent or contradictory—if the answers people give to questions and their religious behavior which is observed appear to be in conflict with other stated beliefs, it may simply be that the observer is inadvertently

[28] Ibid, 330

comparing theology to online thinking, as opposed to comparing theological beliefs to each other. "Since they are tied to intuitive knowledge," Barrett says, "basic level concepts are more stable across groups than are theological concepts."[29] He also warns against comparing the theological level of one group to the basic level of another, as this will lead to mistaken conclusions about the complexity of religious thinking of some groups compared to others. Lastly, there is the implication of TC that we must be careful to articulate which level is being studied, in order to avoid taking the stated beliefs of clergy as necessarily representing those of all believers in a particular group, or hastily concluding that believers apply the same level of theological reasoning (or lack thereof) in all contexts.

5. Perceiving agency in design

In this section I am going to examine a kind of online/folk/intuitive way of attributing agency based on perception of the workings of the universe, in order to consider how our inferences regarding theory of mind work in practice. For that purpose, I will use a real world ongoing struggle over existential meaning as illustration.[30]

At the basic level, simply seeing agency in the universe to begin with seems to create a sense of relevance for human beings. A skeptic might say that the teleological argument for the existence of God, the argument from design, is essentially an argument from pareidolia—if the universe looks designed, then it is designed. There is a reason for why something exists rather than nothing, and that is because some*one* wanted there to be. Aquinas argued that elements of the universe, such as astral bodies, act in pursuance of ends though they have no intelligence, and that this therefore points to the existence of an original intelligent agent who put them into motion—this being is called God.[31] The theologian William Paley would later compare the universe to a watch found on the beach. Upon picking up this watch and examining it, a person would naturally infer that it had been designed and created intentionally.

Today, advocates of Intelligent Design argue that the "irreducible complexity" of some organisms indicates that they could

---

[29] Ibid., 334

[30] Journalist Gordy Slack chose to title his book on the Kitzmiller v. Dover trial *The Battle Over the Meaning of Everything*, which from a certain perspective is no exaggeration.

[31] *Summa Theologica*, fifth proof

not have evolved, but had to have been the special creation of some intelligent agent—officially this agent might not have been a god, but as the battle rages over the question of whether Intelligent Design ought to be taught in American public schools, advocates "on the ground" inevitably reveal their hand by proudly noting that the designer they have in mind is none other than Christianity's god. At the recent and highly publicized trial Kitzmiller v. Dover in Dover, Pennsylvania, biochemist and expert witness for the defense Michael Behe declared in his testimony that intelligent design can be recognized in an organism from its "purposeful arrangement of parts." But in spite of having written elsewhere that "intelligent design theory focuses exclusively on proposed mechanisms of how complex biological structures arose," he could not, when pressed, articulate exactly what those proposed mechanisms are. Lauri Lebo, a local reporter covering the trial, would later write that "Behe could not name the mechanism because there was none. Behe's sole argument for intelligent design was, if it looks designed, it is designed."[32]

Members of the school board who had initially attempted to have a statement promoting intelligent design read by biology teachers in class, and the defendants in the case, would later admit that funds for the intelligent design textbook *Of Pandas and People* had been solicited from school board president Bill Buckingham's church congregation, and that Buckingham had earlier appealed to the board to have the book taught in biology class by saying, "We're just looking for a textbook that balances the teaching of evolution with something else, like creationism."[33] School board members who opposed such measures reported that they later received harassment from those who disagreed with them on the issue, both board members and other Dover residents, in the form of being branded as "atheists"—a label particularly hurtful to teachers Bryan and Christy Rehm, who taught vacation bible school. Such slurs make transparent the belief of these people that if the universe was designed, it was a designed by a god, and anyone who didn't want such a thing taught in public schools must not believe in a god. Biologist Kenneth Miller, however, an expert witness for the prosecution, would later testify that he reconciled his Catholic faith quite comfortably with his belief in evolution, and that intelligent design simply doesn't work as a

[32] Lebo 2008: 154-156

[33] Ibid., 162-163

scientific theory and therefore cannot be taught in a science classroom.

The revealing thing about the continued battle between proponents of ID and those of evolution in the public schools—there is no actual controversy in the scientific community, much as some ID advocates try to portray—is that Intelligent Design supporters portray the issue as a matter of meaning and morals, rather than scientific validity. Philip Johnson, understood by many as the "father" of the modern ID movement and co-founder of the Discovery Institute's Center for the Renewal of Science and Culture, an organization that (among other things) provides legal support to efforts to get ID into public schools, authored the now-famous "Wedge Document" in 1999.[34] This document outlines plans to create a Wedge movement, "aimed at the court of public opinion--which seeks to undermine public support for teaching evolution while cultivating support for 'intelligent design theory'," [35] according to Barbara Forrest, who serves on the board of directors for the National Institute for Science and Education and is author of *Intelligent Design and its Critics*, a history of the ID movement in America and its ideological underpinnings. She points to Johnson's candor in saying elsewhere that "If we understand our own times, we will know that we should affirm the reality of God by challenging the domination of materialism and naturalism in the world of the mind. With the assistance of many friends I have developed a strategy for doing this...We call our strategy the 'wedge.'"[36] The Wedge Document itself makes ample testimony to this goal, stating that:

> Debunking the traditional conceptions of both God and Man, thinkers such as Charles Darwin, Karl Marx, and Sigmund Freud portrayed humans not as moral and spiritual beings, but as animals or machines who inhabited a universe ruled by purely impersonal forces and whose behavior and very thoughts were dictated by the unbending forces of biology, chemistry, and environment. This materialistic conception of reality eventually infected virtually every area of our culture, from politics and economics to literature and art.[37]

---

[34] This document can be viewed at http://www.antievolution.org/features/wedge.pdf

[35] http://www.infidels.org/library/modern/barbara_forrest/wedge.html

[36] Johnson 1997: 91-92

[37] http://www.antievolution.org/features/wedge.pdf

To the members of the Center for the Renewal of Science and Culture, then, a universe ruled by impersonal forces is not one in which humans can be moral or spiritual, where free will and responsibility cannot exist. In order to escape this amoral existence, we must maintain that the universe was created by a personal force—some intelligent, deliberate entity. This is not a scientific argument. It is an argument from how things ought to be, rather than how things are. Not only is the universe designed because it looks designed, it *must* be designed—our morality and dignity depend on it.

I spend so much time discussing Intelligent Design here because its existence, and the fact that so many find it convincing in spite of the fact that it is not actually a scientific theory, points to a human need to see purpose in the universe—a level of agency above our own which makes it possible for meaning and morality to pre-exist our own creation of it. If the universe is designed, then we are not the authors of its significance—that privilege belongs to the designer. This is why appeals to irreducible complexity are not satisfying on their own. Life *is* complex—any biologist, theist or otherwise, will happily admit this—but that complexity on its own is not enough to demand an author. Rather, it is all of the things intuitively associated with an agent—purpose, goals, intelligence, responsibility—that make it seem like a universe which was *not* authored by an agent therefore cannot be populated by real agents. If our world was not created by someone with purpose, meaning, and morality, then we cannot really have such things. The agency must come from above.

5. Non-gods: other forms of supernatural (?) agency

"Our species' defining ability to differentiate unobservable minds from observable bodies"[38] makes it possible for us to think of human agents as supernatural as well, asserts psychologist Jesse Bering. His work is devoted to exploring how we conceive of this separation, focusing on our ability to thinking of people—indeed, our inability to *refrain* from thinking of people—as continuing to exist after their death. He has discovered so far that even if people profess to be materialists, they often can't help attributing certain mental states to

[38] Bering 2006: 453

others after they have died. This was determined first with children by portraying a story in which a mouse was eaten by an alligator:

> The mouse strolled across a small puppet theater made up with a plastic tree and artificial grass, where the alligator hid. The mouse explained his woes: He was lost, sick, sleepy, hungry, and thirsty. After the mouse's brief soliloquy, Mr. Alligator emerged perfunctorily and ate the mouse. The end.

The script called for a quick and painless murder: "Uh-oh! Mr. Alligator sees Brown Mouse and is coming to get him!" the experimenter would say as his alligator hand puppet devoured the hapless rodent. "Well, it looks like Brown Mouse got eaten by Mr. Alligator. Brown Mouse is not alive anymore. Will the mouse ever need to eat food again? Will he ever grow up to be an old mouse? Will he ever need to drink water again? Is he still hungry? Is he still sleepy? Can he still hear the birds singing? Can he still smell the flowers? Does he still want to go home? Is he still scared of the alligator?"[39] He found that while kindergarten-aged children were in high agreement (88 percent) that the dead mouse's brain no longer worked, only a small number (24 percent) didn't think he would still want to go home. Older children (10 to 12 years old) agreed on the first question (81 percent), but a larger number (46 percent) were skeptical about whether the mouse still desired to go home. The same pattern repeated throughout the experiment and over different groups of children — younger kids understood the physical reality of death but were more apt to believe that sensory perceptions and feelings existed after life ceased.

This tendency, however, is not limited to small children. Bering also tested undergraduate students with a mean age of 20 years, telling them stories about a man or a woman who experienced various psychobiological, perceptual, epistemic, emotional, and desirous states, who then were killed off mercilessly. Then he quizzed the students to see which aspects they would attribute to the person once they were dead. "Is he still *hungry*?" "Can he *see* the paramedics trying to resuscitate him?" "Is he *thinking* about his wife?" "Is he *angry* at his wife?" "Does he *want* to be alive?"[40] As you might expect, the more likely we are to put something in the

---

[39] http://www.newtimesbpb.com/Issues/2006-03-09/news/feature_4.html

[40] Bering 2002

category of "mental" rather than "physical," the more likely the test subject was to attribute such ability to the dead person. Even those who had classified themselves as "extinctivists" when asked their position on life after death sometimes gave strange answers. When asked if the character in question knew that he was dead, one man even said "Yeah, he'd know, because I don't believe in the afterlife. It is non-existent; he sees that now."[41] Bering contends that this, paradoxical as it is, is an entirely natural response for us because it is so very difficult for us to *stop* empathizing with others—to put our theory of mind to rest, so to speak, along with them. While it's not very difficult to imagine not being hungry or sleepy, since we experience those states often, we can't imagine what it must be like to just stop existing as a conscious being, so we can't imagine what it must be like for someone else to cease to exist either. This is what Bering refers to as a simulation constraint. We go right along attributing thoughts and emotions to them as if they were still around, because it is so difficult to imagine otherwise.

"Folk psychology" is a term that refers to our common sense, intuitive beliefs about how other people think. The folk psychology of the soul,[42] then, as Bering puts it, refers to our intuitive beliefs about the element of human identity which is separable from the body. As illustrated above, people are very likely to attribute mental states to dead or invisible agents, even if they are not nearly as inclined to attribute physical states. Tying the individual into the concept of Intelligent Design, Bering notes that "if the soul is not the product of intelligent design, then there is no teleological function that it is designed to fulfill, no *raison d'être* to explain its existence beyond human attributions of purpose."[43] Belief in the soul, and in its immortality, Bering notes, is directly supported by belief in gods:

> Those who believe that one's life is owned by God are more likely to view suicide—as well as abortion, capital punishment, and medical euthanasia—as being morally wrong. It is not suicide per se that sheds light on this teleo-functional bias, but the moral repugnance for the act. Religious rules against suicide reveal a more complex cognitive stance in which the

---

[41] http://www.edge.org/q2005/q05_7.html

[42] Bering 2006: 453-498

[43] Ibid., 458

> self's will is seen as imposing itself over the will of the creator of the self's will. According to the premise that a person's life belongs to God, an individual does not have the right to purposefully causes his or her death, because this right is seen as being God's alone. This conception suggests that suicide is viewed as a moral transgression in which an individual "cheats" God by stealing the latter's power of intentionality in causing the self's death. Suicide therefore becomes a form of intellectual theft; the self redesigns its end in an act of mutiny against its creator.[44]

The significance of our identity as souls, as such portrayed, is at least in part dependent on the significance of the over-arching purposefulness of living in a universe which was created by an intelligent entity.

5. Again—*what* is a supernatural agent?

This chapter suffers from an abundance of overlapping terminology. Rather than being departures from each other in explaining the origins of supernatural agent concepts (though Guthrie does not call them supernatural), the theories presented in this chapter build off of each other, a process which involves the production of new labels in order to convey exactly what each theorist means. But this leaves us with a word salad, unable to give a concise yet cohesive explanation of what a supernatural agent is—or, for that matter, what exactly makes it supernatural. So I am going to go back through some elements of the theories presented and try to iron out a clearer concept. All of the theories presented are cognitive in the sense that they attempt to explain belief in supernatural agents by examining what makes our minds naturally disposed toward such belief—but they differ in terms of identifying the reasoning for such dispositions, as well as the nature of the agent concepts which are produced by them and the implications of those concepts. Rather than defining "supernatural" in the standard sense of being over or beyond nature, unexplained or inexplicable by science, Boyer classifies as supernatural those concepts which are minimally counterintuitive in that they constitute minor violations of our mental ontological categories. When those

[44] Ibid., 459

supernatural concepts are taken seriously, they can be classified as religious concepts. Being taken seriously, in this case, means that they "have a direct and important effect on [peoples'] lives."[45]

The most succinct criticism that one could make of Boyer's portrayal of religious concepts, from Guthrie's point of view, is that it is no coincidence that those "counterintuitive" concepts which are most catchy, and most likely to be taken seriously, are either a) non-PERSON concepts which are ascribed agent-like properties, or b) PERSON concepts which are ascribed non-PERSON-like properties. Not only are agent concepts central to religion because gods and spirits can function as full-access strategic agents, but also because the pull to attribute agency not only to gods and spirits but also to everything else—artefacts, plants, animals, and natural objects—is irresistible. It is implicit, automatic, and non-reflective—which is to say, it is intuitive. In his 2000 article, Boyer anticipates this criticism. When considering the question of "why socially significant supernatural concepts are largely about agents," he notes that one explanation is "that we generally tend to project human features onto unknown entities, because human features are most familiar. This is misleading on two counts. First, what is projected is not specifically human. Intuitive theory of mind is spontaneously activated by the presence not just of humans, but also of most animals. Second, not *all* human characteristics are projected, only intuitive theory of mind."[46] The first defense misses the point. It is possible to learn enough about how an animal thinks that you develop a theory of mind for it which does not anthropomorphize the animal. Anthropomorphism, however, is not the projection of dog agency onto dogs; it is the projection of *human* agency onto dogs—hence the "anthro" prefix. Humans may develop a complex understanding of how that animal thinks, but that doesn't qualify as "spontaneously activated theory of mind," and it doesn't mean they are treating the animal as a supernatural agent, by Boyer's definition. Humans who do not know much about animal minds, and/or those who are in online, non-reflective mode, tend to treat most animals as though the animals think like humans. Gods and spirits who have animal forms are also ascribed human agency—a raccoon god, for example, would not think like a raccoon (or at least, not strictly like one) but like a human. Boyer's second objection, that not all human characteristics are projected when anthropomorphizing

[45] Boyer 2002: 104

[46] Boyer 2000: 203

but simply intuitive theory of mind, is beside the point as a theory of religion as anthropomorphism does not require otherwise.

There is also the matter of what properties Boyer maintains constitute violations of intuitive inferences from the PERSON ontological category. The notion of theological correctness is instructive in establishing this, since theologically *in*correct concepts are what we are looking for when trying to establish what counts as intuitive. Barrett's theory is that TC concepts will be complex and require reflection, whereas non-TC concepts come rapidly and without much thought—they are intuitive. So intuitive notions about humans include the idea that they have a physical and temporal location as opposed to being distributed amongst many different places and throughout time, the idea that they do not know everything and do not have access to all information there is to know (they are only partial access strategic agents), and they perform tasks in sequence rather than being able to do as many things simultaneously as they choose, since they cannot pay attention to everything at once. This is not a summary of all intuitive beliefs about humans, however, and that is the problem. The question of what falls into the category of our naïve psychology, our intuitive theory of mind, has still not nearly been answered completely. And so when we talk about PERSON concepts with a violation or breach in terms of inferences concerning intuitive psychology and biology, we do not have access to the full extent of what that may mean—or what it rules out. In *Religion Explained*, Boyer does note that

> intuitive psychology references are applied to intentional agents in general, not just persons. So it is quite likely that concepts of gods and spirits are mostly organized by our intuitive notions of agency in general (the abstract quality that is present in animals, persons, and anything that appears to move of its own accord in pursuance of its own goals) rather than just human agency.[47]

But though humans might have a general concept of agency, does that really mean that they intuitively apply *that* to gods and spirits, rather than a naïve psychology of the human mind more specifically? As mentioned above, the idea that our intuitions about gods and spirits are based in anthropomorphism seems much more plausible.

---

[47] Boyer 2002: 164

One other problem with Boyer's theory of counterintuitive supernatural concepts, I believe, is that it tacitly assumes that people's intuitive inferences about the world around them are accurate. That is why the notion of religion as systematized anthropomorphism does not work with it, because anthropomorphism is a particular way of mistaking the way things are; seeing human agency where it is not present. Since even the most distinguished psychologists and biologists do not know everything there is to know about their respective fields, it is fair to say that the vast majority of humans, over the vast majority of time, have made intuitive but incorrect inferences about how human psychology and biology "works." The body of inferences that can be labeled as "folk psychology" and "folk biology" is therefore bound to include some inferences which are simply erroneous. As Bering's work indicates, human cognitive constraints make it impossible to know what it is like to be dead. Even if we fervently believe that there is no afterlife, that consciousness stops along with the body, we are simply unable to process non-consciousness. Non-consciousness is therefore radically counterintuitive.

Chapter 2

Empathy: Connections and Disconnections in Theory of Mind

Empathy is a term used in many different ways by those who study it—it can mean anything from emotional contagion, such as when one baby's crying is picked up by others and becomes a chorus, to literal telepathy.[48] While shopping online for a book by psychologist William Ickes entitled *Everyday Mind-Reading*, I was startled to discover negative reviews of the book from people who had expected it to teach them how to practice actual extra-sensory perception. What they got instead was a book detailing a number of experiments in which people were asked to describe what partners of varying degrees of familiarity and similarity were thinking at different points while watching videos of previous interactions between the two. To the extent that the book addressed what exactly makes people better at understanding what others are thinking, its title was certainly accurate, though the "mind-reading" in question was in no way supernatural.

The fact that this confusion is possible, however, hints at something about empathy in folk understanding—we're not exactly sure how we do it. The content of another person's mental states can be communicated in myriad ways, oftentimes mostly unintentional. So, too, can our understanding of them be unintentional. Most people will have no trouble interpreting a person with a slumped posture, down-turned mouth, and tears running down his cheeks as being sad—indeed, they are unable to avoid it. They may also be unable to avoid feeling sorry for him, especially if they know him and/or can identify with his predicament. Identification is the essence of most

[48] Portions of this chapter have been taken from Koch (forthcoming)

definitions of empathy: what does "stepping into the shoes" of another person tell me about what they are thinking, and what does it make me feel toward them? These two aspects of empathy are often called cognitive and affective empathy, respectively, though there is of course considerable overlap between them. To understand empathy we need to understand the connection between knowing what a person is thinking or feeling and feeling *for* them, and what happens when this connection is not made. In this chapter I will discuss the development of theory of mind both in the individual and over human evolutionary time, with special attention to our capacity to imagine and consider what is going on in the minds of others.

1. Ontogeny: empathy in development

Humans are born empathizers. We've had to be—our ancestors have been social since before the species who would become chimpanzees and moderns humans diverged. As the brain size of humans increased, so did our vulnerability at birth and young childhood, requiring children to stay with their parents for much longer than other primates do. In order to maintain social arrangements which would make this extended child-rearing possible, the same big brains which necessitated it also made it possible for us to practice cognitive empathy—to understand other peoples' beliefs, including their false beliefs. The ability to practice cognitive empathy entails the ability to deceive. A classic test developed by Heinz Wimmer and Josef Perner to study development of theory of mind is the Sally-Ann task, in which a doll named Sally is shown placing a ball in a box, and then going outside. A doll named Ann, who watched Sally do this, is depicted removing the ball from the box and placing it somewhere else. Then Sally returns to the room, and the child is asked where Sally will look for the item. A four year old child whose understanding of false beliefs is developing normally will reply that Sally will look for the ball where she left it, not knowing that it had been moved. Autistic children of the same age, however, typically will not make this inference and instead say that Sally will look in the place where Ann has hidden the ball.[49] False belief tasks require the ability to know what is true, but also know that another

[49] In her book *Autism: Explaining the Engima*, Uta Frith shows how this experiment can be repeated in many different ways, using pictures instead of dolls.

person's beliefs are not the same as yours, and do not correspond to the truth. The cognitive ability this requires is more complex than it might seem, especially when we are expected to connect chains of belief such as that Molly is cheating on her husband Bob, who does not know it but who is also cheating on Molly with Jenny, who knows that Molly is cheating but won't tell Bob, and Molly's husband Richard who knows about Molly and Bob but not Jenny. Being able to process such chains of belief is of course a vital prerequisite for watching televised soap operas, but it also came in very handy in our evolutionary history. Anthropologist Robin Dunbar has illustrated how this ability to understand varying levels of intentionality enables us to keep track of complex ideas about the knowledge and purposes of many different people at once, which makes gossip a highly popular activity and, much later in the evolutionary timeline, suggests an explanation for the enormous appeal of television soap. Stories about the complex levels of relationships between characters may be appealing because we cannot "turn off" our intense interest in social relations, or because thinking about fictional occurrences of duplicity and manipulation can help us to have a grasp of how to handle such instances even if we haven't had much practice in our own lives.

Teaching is made possible by the comprehension that other people have different beliefs and do not share your knowledge. If I can tell which things you do not know, then I know what you need to learn. Young children have been described as having the "curse of knowledge," in which the barriers between what I know and what everybody else knows have not been fully constructed, making small children very poor deceivers. Paul Bloom relates a story of a friend of his whose family planned to surprised him with a pie on his birthday. All was apparently going according to plan, until his three year old niece, who had been instructed repeatedly to keep the secret, walked up to him and shouted "There is no pie!"[50]

The origin of understanding the beliefs and goals of others appears to lie in joint attention. Normally developing children begin to point declaratively—that is, pointing in order to draw another person's attention to a third object, as opposed to imperative pointing which is more like reaching—at thirteen to fifteen month.[51] A chimpanzee can point imperatively, but evidence of declarative

[50] Bloom 2004: 20-21

[51] Tomasello 2001: 65

pointing even among human-raised chimps is scant.[52] Psychologist Michael Tomasello argues that this disparity exists because human children are capable of joint attention, the ability and desire to share with another person the experience of an outside entity, at nine to twelve months whereas no other primates are at any point in their development. Small children delight in pointing things out to their parents, and are gratified when the parent acknowledges and comments. A child who points out a train, for example, saying "Train!" and looks back and forth from the train to her mother, is attempting to participate in joint attention. If her mother does not look at the train, or looks but otherwise shows no reaction, the child may become upset.

The ability to learn language in the first place appears heavily dependent on joint attention, which may explain why autistic children, who do not commonly point declaratively or do so much later in their development, also experience a delay in language acquisition or may not begin to speak at all.[53] The mutual focus on a train which allowed the mother to teach her daughter its name to begin with required joint attention, in order to convey that the word "train" in fact refers to that large, long carriage moving down a track, rather than a meaningless exclamation on the mother's part. Later, if the child and her mother want to talk about trains, they will have to have an invisible object as their outside referent. There is nothing to point to physically, but they have a shared verbal pointer to a concept. Parents and children can also repeat this process by using picture books from a very young age, so that the child can learn words for things through shared focus with the parent.

Tomasello argues that development of these skills in infants also has another very important connection. When they begin to engage in joint attentional sequences, they likewise begin to recognize others as intentional agents like themselves. Intentional agents have goals and take deliberate action to implement those goals, which includes directing their attention to specific objects to be used in that implementation. Though the broader term for this set of abilities is theory of mind, there is considerable debate about whether "theory" is in fact the correct term for it. Calling it a theory implies certain things about its origin and how it is implemented—a theory is something explicit, reflective, conscious, and effortful. This would suggest that

[52] Tomasello & Call (1999)

[53] Frith 1989: 102

children learn to view others as intentional agents by being specifically instructed to do so, or at least by consciously applying trial and error reasoning in their minds in order to work out how different kinds of behaviors affect different people. This version of understanding intentionality is referred to as "theory theory" for that reason. While such instruction certainly does take place, it cannot account completely for the rapidity at which children come to grasp the idea that others have different thoughts and feelings than they do.

The alternate possibility is that we come to this understanding by simulation—that is, we come to identify with others by creating for ourselves a mindset which is as close as possible to that of the person we are thinking about by accessing memories from our own experience, which allows us to share in the person's experience to some extent. A parent whose child has just hurt someone else may prompt the child to identify with that person as a means of recognizing their own wrong-doing: "You wouldn't like it if Bobby hit *you*, would you?" But it would seem that there must be some pre-existing capacity to identify with the emotions of others which makes this type of instruction effective. The reason that this prompting to identify with Bobby's feelings works because the child most likely has been hit at some point in his or her life previously, and can remember what that is like in order to realize that it is not something appropriate for him to do to Bobby.

Tomasello favors a simulationist explanation, maintaining that we begin to understand that others have goals and motivations "like me," though it need not be a matter of having explicit conscious access to one's own goals and motivations before gaining the ability to attribute them to others.[54] The difference between this view and "theory theory" is that simulationists tend to view our understanding of the goals of others as more or less unconscious and automatic, whereas TT proponents portray it more as a matter of explicit, proto-scientific theorizing. This is of course a grossly simplified version of the conversation, and there are scholars of developmental psychology at many different points on the spectrum between the two.

In his book *The Work of the Imagination*, developmental psychologist Paul Harris draws a connection between imagination and empathy. At their most basic level, Harris suggests, both abilities entail the practice of entertaining counter-factuals—thinking about things which aren't true. It is not true that I feel pain when looking at

[54] Ibid., 75-76

Bobby in pain (not the immediate physical pain of having been hit, at least), but I am able to draw on my understanding of what it's like to be struck in order to project the emotion of pain and be affected by it. Simulation theory entails this understanding. We rely on our previous experience, or the closest approximation in it, to identify with what others are experiencing now. Harris notes that children who engage in imaginative play with siblings when they are younger tend to become better at practicing cognitive empathy as they grow: "we have persuasive evidence that involvement in role play, whether in context of joint play with other children or pursued in a more solitary fashion through the creation of an imaginary character, is a correlate, and indeed an advance predictor, of later success on belief tasks."[55] Playing cops and robbers, for example, might seem an unlikely candidate for such a correlation, but adopting a role in play-acting, constantly imagining oneself with different goals, different emotions, can be very useful practice. And yet this is something children do naturally—"You will be the doctor, and I'll be the patient. You have to ask what's wrong with me, and then tell me what to do about it." In order to act out such a scene, both actors have to pretend that their characters a) don't know something which b) isn't true, but c) is postulated as true for the purpose of the scene. They must track what hypothetical people know and don't know, feel and don't feel, and differentiate it from their own real knowledge and feelings. The "doctor" doesn't get confused and think that something is actually wrong with her friend, any more than the "queen" gets confused and thinks that there is actually tea in her empty cup during a tea party. This is why children who engage in role play, as Harris notes, "turn out to be better able to view a situation from another person's point of view, as indexed by various measures of mental state understanding."[56]

2. Theories of minds in other bodies

Once we have acquired the ability to recognize others as having an identity which is separate and distinction from our own, we can begin to practice true cognitive empathy, imaginatively projecting

[55] Harris 2000: 45

[56] Ibid., 47-48

characteristics of that person's mind based on both our simulation of what we imagine they think and feel as well as our accumulated knowledge—our "theory"—of how other minds work. That being the case, I think it is important to consider for a moment how this projection works when it comes to separating a person from their body.

Possession—the occupation by one person of another person's body—is a familiar concept to most people. The film *Freaky Friday* told the story of a mother and daughter swapping bodies and trying desperately to play it off so that no one would realize. In the television show *Quantum Leap*, main character Sam Beckett leaps through time to land in the bodies of random people, faced with the task of solving some problem in their lives. In *Being John Malkovich*, a nerdy puppeteer discovers a portal that leads directly into the mind of the famed actor, allowing him to take over Malkovich's body and change his career. The term for this storytelling device when used in television and film is "body swap" (the popular web site Internet Movie Database uses the key word "soul transference"[57]) and it is an easy concept for the viewer to grasp. We can accept what has happened and move on with the plot without stopping to think "Wait a minute, so *which* traits of the original character will now be displayed in this other body?" We can easily understand what is happening when, in the 2003 version of *Freaky Friday*, the mother's character (who now occupies the body of her teenage daughter) suddenly realizes that she can eat French fries with impunity despite having denied herself the pleasure before, since her daughter's teenage body can metabolize them much faster and won't put on weight like her mother's body would.

Anthropologist Emma Cohen has been investigating folk concepts of spirit possession from a cognitive standpoint, in order to find out where inferences about theory of mind come in when we think about people switching bodies. In order to do so, she traveled to the city of Belém in northern Brazil to study a community of Afro-Brazilian (*culto afro*) cultists whose rituals involve trance and possession. In these rituals, the mediums would become possessed by one of the *orixás*, or personal spirits, for a temporary period. Mediums are capable of channeling different *orixás* for different purposes. Cohen was particularly interested in finding out if, during this period, the observers and the mediums themselves perceived the

[57] http://www.imdb.com/keyword/soul-transference/

medium's agency to be displaced by the possessing spirit, or if rather there is some fusion of the two which takes place. Cohen spoke to the *pai de santo* (the leader of the worship community where she was staying), who told her that when the spirit possesses a medium, it merges with the essence of that medium. This would explain why the same spirit could possess different people with differing results each time, as well as that if Cohen spoke to the spirit Ogum while he possessed person A, and then again the next day when he possessed person B, Ogum might not remember some of the things she had discussed with him.

When Cohen spoke with other mediums, however, who had been exposed to much less of the *culto afro* teachings than the *pai-de-santo*, the depiction of possession was notably different. Cohen reports that "a senior member clearly described possession as the joining of the body of the medium with the spirit of the entity. These two parts, he claimed, make up the new (possessed) person. Another senior ranking member described possession as the moment in which one's own spirit withdraws 'and another spirit comes and throws him/herself into your body.' Drawing a clear demarcation between medium and spirit, another member describes her possession episodes as follows: 'I don't know where my spirit goes. I don't know. I only know that I switch off. I don't remain in me.' Another person stated, 'Possession for me is a state of unconsciousness… in which we are not answerable for our actions, our bodily movements …we don't have control of our bodies anymore. It's the total loss of control of the body and the mind. Something else controls – it is the spiritual being.'"[58] If you were unaware that a person was currently possessed, Cohen notes, and addressed them by the person's name rather than the name of the entity possessing them, the person would say "I am not [person's name]; I am [entity's name]." Mediums spoke of their spirits lying down or dreaming while possessed, allowing the possessing spirit to take control and dominate them. This would seem to indicate that when speaking non-reflectively, the mediums viewed possession strictly as a matter their spirit being displaced, even if the more "theologically correct" version of the event said differently. After describing this, Cohen notes an intriguing aspect of possession as displacement from anthropologist Erika Bourguignon: "W]hen the spirits take over, women can do unconsciously what they do not permit themselves to do consciously. The demands that are made, the

[58] Cohen & Barrett 2008: 9

orders that are given, are those of the spirits' doings and sayings. They are neither responsible for nor aware of what is going on and do not remember it after the fact. They have ultimate deniability."[59]

Cohen and Barrett then decided to examine beliefs about minds and bodies in a community which does not (presumably) practice rituals involving possession—undergraduate university students in Northern Ireland. They wanted to find out if there was a strong inclination either way concerning which aspects of a person are subverted when possessed and which are not. The participants read ten different scenarios about two characters, Ann and Beth, in which hypothetical mind-switching takes place. One example: "Ann is very good at maths. She regularly gets excellent marks on 7-point quizzes – usually around 6 out of 7 of her answers are correct. Beth is very poor at maths. She regularly gets poor marks on 7-point quizzes – usually around 2 out of 7 of her answers are correct. Once when the girls were in maths class, somehow Beth's mind went into Ann's body. How well do you think that the girl will do in the maths test?"[60] Each scenario included typical Ann behaviors and typical Beth behaviors, as in this example. Subjects could then give their answer to each question on a seven point scale.

What Cohen and Barrett found from this experiment appears consistent with what initiate mediums in the *afro culto* told Cohen about being possessed. They treated possession as a kind of displacement *when talking about behaviors with a strong mental component* (such as doing well on a math test). When asked about behaviors with a strong physical or biological component, such as seeing with precision, respondents were much less likely to treat that behavior as being subverted in the possessee by the possessor. "These results suggest a tentative conclusion," reported Cohen and Barrett. " Northern Irish young adults tend to spontaneously infer that when one person's mind is transferred into another person's body, the normal 'host' mind is displaced. Displacement was spontaneously inferred significantly more frequently than fusion, even though both options were equally available as valid responses. This suggests that participants' responses were

[59] Quoted in Ibid., 11

[60] Ibid., 13

guided by a tacit one mind-one body principle."[61] The reason for this, Cohen and Barrett went on to suggest, may be that dualism truly is intuitive and therefore children come to understand a principle that only one mind is responsible for the behaviors exhibited by one body. This might explain why displacement theories are advocated even by people who have been given authoritative teachings to the contrary.

3. Common-sense dualists

In addition to having certain intuitions about what aspects of a person are transferred when their spirit possesses someone else, we have similar intuitions regarding what aspects of ourselves "belong" to our bodies and which do not. The manner of speaking that people naturally use concerning their bodies is one of ownership—"my" body or "my" brain, as we would refer to "my" car or house. Psychologist Paul Bloom argued in his recent book *Descartes' Baby* that we are common-sense dualists, which means that we automatically act and speak as if the self and body are separate, and what's more that the self is immaterial. The accumulating evidence about how infants react to different stimuli, Bloom says, gives credence to the idea that our dualism starts right from the beginning. Contrary to previous thinking that the infant's mind is a "blooming, buzzing confusion," as William James put it, there is in fact a very noticeable kind of differentiation going on which betrays a type of binary thinking we will continue to entertain for the rest of our lives—dividing the world into either agents or physical objects. Using props and performing myriad experiments to see where infants direct their attention during different kinds of stimulation, psychologists Philip Kellman and Elizabeth Spelke found that three-to-five month olds were able to apply certain principles when viewing physical objects. The four object principles found were:

1. Cohesion: babies expect objects to remain whole. If one end of an object is pulled, infants expect the entire

[61] Ibid., 18

object to go in the direction of the pull. If the object comes apart instead, they are surprised.

2. Continuity: if a stick is moved back and forth behind a box, with the ends of the stick appearing from behind the box at both the top and bottom, infants expect the ends to move together in the same direction, since they are perceived as a single object.
3. Solidity: place an object behind a screen, and tilt the screen backward—a baby watching this should expect the screen to be stopped in its tilting by the object in the way.
4. Contact: if an object moves toward another, but the second object moves away before it is hit by the first, this is surprising for babies just as it is for adults. This is because the expectation of contact entails that objects can only influence each other by touching.[62]

Contrast this with the way that babies react to people—they expect peoples' faces to move in response to their attempts to interact, and get upset if a person's face instead remains still and expressionless. Babies prefer to look at human faces compared to nearly everything else. They can distinguish happy faces from sad faces, and are able to imitate the facial expressions of the person they are looking at. When watching a hand move toward one of two objects, a baby will expect the hand to "chase" after the object it was originally heading toward if the two objects change places. "Hands are attached to people," notes Bloom, "and people have goals, and babies seem to understand that a reasonable goal for a person is to reach for a particular object, not go to a specific location."[63]

He acknowledges, however, that the ability to distinguish people from objects does not in itself demonstrate that babies are commonsense dualists. There is also the way that small children think about their own thinking. Once they learn about the existence of the brain, Bloom argues, they tend to assign it certain tasks while denying it others. A brain is involved in perception—seeing, tasting, smelling, and hearing.[64]

---

[62] Bloom 2004a: 12

[63] Ibid., 15

[64] Bloom does not mention touch, which I find interesting—all of the senses he mentions are those which belong to the head. The head can experience touch, of course, but the entire body can as well, and we tend to associate touch most strongly

What brain does *not* do, report children, is love Mom or pretend to be a kangaroo. Those are things the self does. "Once children learn that the brain is involved in thinking, they don't take it as showing that the brain as the source of mental life; they don't become materialists. Rather they interpret 'thinking' in a narrow sense, and conclude that the brain is a cognitive prosthesis, something added to the soul to enhance its computing power."[65]

Research from Paul Harris and Rebekah Richert supports this. They performed a study in which American Protestant children aged 4-12 were asked about what the effects of a religious ritual might be upon the mind, brain, or soul. The children, especially the older ones, said that when a baby is baptized a change happens on the inside in a way that cannot be seen or touched, and were much more likely to say that the soul changes than the mind or brain, but also were more likely to say that the mind changes than that the brain does. When asked more specifically about the nature of the brain, mind, and soul, children agreed that the mind and brain change over time, but were divided as to whether the soul does as well. When asked about whether the soul, mind, or brain confer identity, children responded that the child would not be the same if any of the above were removed, but the older children were most likely to respond that the removal of the soul would change the baby. "Surprisingly," Harris and Richert say, "even though children claimed across the board that babies had souls, they typically did not see the soul as being critical for cognitive, non-cognitive, or biological functioning, suggesting that they do not see the soul as being the same as either the mind or the brain. So what do children think the soul actually *does*?"[66] As it turns out, the children responded that the soul gives life, it goes to heaven when you die, that it is a "ghost inside your body," that it makes it possible to love, and that it gives them contact with God. By contrast, they gave tasks more biologically and cognitively oriented to the brain and mind, respectively.

---

with hands. Though Bloom does not address this, my guess is this is because the brain is imagined to be a kind of container for the self, which is why we may ascribe the ability to see to the brain, but at the same time get the strong impression that the self is located behind the eyes, looking out.

[65] Bloom 2004b

[66] Harris & Richert 2006: 419

This type of evidence, along with data from the experiments of Jesse Bering which shows that children intuitively ascribe certain mental and emotional characteristics to people after they are dead (as described in the last chapter) convince Bloom that there is at least an intuitive division between mind and body, though he does not develop that into a tri-part distinction as Harris and Richert do. Bloom's primary concern is what kind of effect the idea of a soulless body has on people. His choice of title for his book on this subject, *Descartes' Baby*, comes from a story told about Descartes after his death. The story goes that Descartes was rather a loner, but loved his illegitimate daughter Francine devoutly. Francine died at five years old, however, and Descartes was so distraught in his grief that he felt compelled to create a replacement for her. The replacement was an automaton, resembling Francine uncannily in many ways. Descartes came to love his mechanical "daughter," and took it everywhere he went, storing it in a small trunk. When Descartes took a journey by ship across the Holland Sea, he brought his mechanical doll with him in the trunk. The captain of the ship noticed the trunk, and the care with which Descartes treated it. One night when Descartes was asleep, the captain's curiosity got the better of him and he sneaked into the room and took the trunk. When he opened it, he was horrified to discover the automaton inside, and in his revulsion threw the doll overboard. "I like this story," Bloom says,

> because it captures how disturbing—in some cases, revolting—we find a body without a soul. It is a nice illustration of the emotional pull our common sense could have. But it also raises a serious problem. Common-sense dualism is wrong. There is no consensus as to precisely how mental life emerges from a physical brain, but there is no doubt that this is its source. If by "soul" then, you mean something immaterial and immortal, then souls do not exist. All of us are soulless bodies, no less than the robot Francine. In this sense too, we are all Descartes' babies.

Here is where, for Bloom, it gets frightening. If people are walking around with an intuitive psychology which tells them that everyone is a soul who is separate from their body, what does that mean for politics? If every person is an embodied

soul, then what will that leads us to conclude about abortion? Cloning? Euthanasia? Presumably there must be a point at which the soul enters the body, and a point at which it leaves—how do we determine that? Religion does not require dualism, Bloom maintains, but "a religion such as Judaism or Catholicism. . . cannot survive the rejection of the immaterial soul."[67]

Bloom wrote a brief editorial expressing these concerns for the New York Times.[68] A couple of weeks later, an interesting response to it was published in the National Review Online by Patrick Lee and Robert P. George—George being a member of George W. Bush's President's Council on Bioethics (to be discussed at greater length in Chapter 5). The essay, titled "Delusions of Dualism," takes Bloom to task for thinking that science has demonstrated that there is no immaterial aspect to people. But that was not their main point of contention with his editorial. The notion that people are made up of bodies and souls, Bloom had suggested, makes it theoretically easy to determine whether and when abortion and euthanasia are wrong—if it were demonstrated that souls do not exist, then we are faced with more complexity when having to make these moral decisions. But actually, Lee and George argue, it would be the other way around:

> If science *did* show that all human acts, including conceptual thought and free choice, are just physical processes, then science *would* provide answers to important moral questions. It would just provide different ones from those we have traditionally accepted. . .it would mean that human beings lack any special, intrinsic dignity worthy of full moral respect. Thus, it would undermine the norms that forbid killing and eating human beings as we kill and eat chickens, or treating them as beasts of burden as we do horses or oxen.[69] (Emphasis in original)

They go on to say that humans are intrinsically valuable due to their status as humans, rather than because they were endowed with a soul which conveys value upon them. Human rights

---

[67] Bloom 2004b

[68] Bloom 2004c

[69] Lee and George 2004

begin at conception, they clarify, but not because that is the moment when they receive their soul. It is because that is when a new human is created. Lee and George allow that humans are physical organisms, a type of animal, functioning through sensation and perception, physical responses to their environment. Still, however, they are unwilling to grant that materialism is true: "On the other hand, conceptual thought and free choice are actions that cannot be material actions: that is, they are spiritual or non-material actions." They go on to say that neither soul/body dualism nor materialism is true, but that "we are more than just the latest product of blind evolution, since there is an immaterial aspect of us that could not have emerged from lower material forces." Lee and George note that their position matches that of the "world's largest religious body, the Catholic Church."[70]

This statement strongly implies that they have not read *Descartes' Baby*, because then they would realize that this "third alternative" is still what Bloom would call soul/body dualism. This, it should be noted, can be labeled a flaw in Bloom's theory—it is too simplistic. As Harris and Richert said, Bloom does not clearly enough articulate which aspects of a person are intuitively immaterial. However, the idea that there *is* an intuitively immaterial aspect to people, and that this aspect is somehow unique to humans, is enough for him. He uses "soul" as the shorthand term for this even if Lee and George refuse to do so. Interestingly, the qualities of a human which they select as being immaterial or "spiritual" are conceptual thought and free choice. There may be more, but they do not elaborate on what else might qualify. Still, it seems appropriate to say that Lee and George believe that if materialism is true, then conceptual thought and free choice are not possible. Bloom, presumably, does *not* believe this, since he believes that materialism is true and yet his concern is for the reaction of outright hostility or denial with which he expects most people would receive this news. Even in the editorial to which Lee and George responded, Bloom disqualifies Pope John Paul II from being able to accept that "our souls are flesh" because he maintained that theories which "consider the spirit as emerging from the forces of living matter, or as a mere epiphenomenon of

[70] Ibid.

this matter, as incompatible with the truth about man."[71] This being the case, Lee and George should have seen that the Catholic Church's view of the soul still constitutes dualism according to Bloom. It seems that in attempting to argue against his position, they instead talked past him in order to preserve a concept of something about humans which gives them dignity, without which they are fit only to be killed and eaten or used as pack animals. It seems clear that this is precisely the dichotomy that Bloom is concerned about. Lee and George do not even specify that a materialist portrayal of human thought and behavior would make conceptual thought and free choice impossible, but rather that simply revealing them to be "just physical processes" would be enough to rob humans of their humanity and make killing them for food justifiable. This topic foreshadows two topics that will be discussed later in the dissertation—the idea that dualistic thinking seems intuitively necessary to provide for free will, and that dualistic thinking causes the concept of a "soulless body" to seem repellant, obscene, and undignified, rendering empathy difficult if not impossible. But for now, it is important to return to the original development of empathy.

4. Phylogeny: empathy in evolution

It is instructive in trying to get a hold on empathy to examine cases in which it might be expected occur, but does not. As mentioned, humans appear to be the only species capable of joint attention from infancy, and who naturally begin to point declaratively, sharing the experience of an outside object with conspecifics simply for the sake of sharing. I say "simply," but in fact one can imagine myriad reasons for wanting to draw someone's attention to an outside object. Tomasello performed an experiment in which a mother sits in an office with her infant on her lap. At the desk in front of them, another woman is stapling papers. When she leaves the room briefly, someone moves her stapler to another location. She returns and looks around the desk with a bewildered expression, and the infant pointed to the location of the stapler for her. This is a version of the Sally-

[71] Quoted in Bloom 2004c

Ann false belief task using real people, and demonstrates a quite sophisticated understanding of another person's goals and state of knowledge.[72] If we posit that the ability to empathize is biologically inherited, the question that naturally follows is, when did it evolve? Do our closest living relatives—chimpanzees and bonobos—have it, or something approximating it? Don't they need it? As noted, there is a faction, including Tomasello, whose answer to the latter two questions is "no." Tomasello suggests that the reason for this may be that chimpanzees spend a lot of time competing, and comparatively little time sharing. A chimpanzee has no need to point out for another where coveted fruit might be in the tree, for example. Sharing amongst chimpanzees does occur, but it happens in the form of directly handing food to a conspecific who is injured, elderly, or otherwise feeble. Tomasello does allow that chimpanzees are capable discerning what conspecifics want based on their gaze, and how to use that knowledge to their advantage:

> Thus, Hare et al. (2000, 2001) placed a dominant and a subordinate individual into competition with each other over food – with some pieces of food visible to both individuals and some visible only to the subordinate chimpanzee. By pursuing most often the piece of food hidden from the dominant's view, subordinates demonstrated that they knew what the dominant could and could not see. And, importantly, the subordinates knew what this meant for the dominant's goal-directed action: if the dominant could see the food or had seen it just before, subordinates could infer that she would go for it (whereas they would not make this inference if what she saw was instead a rock).[73]

This type of result indicates that apes have some level of registering the goals of conspecifics by ascertaining what they are able to perceive, which is not surprising to learn about an animal which competes with members of its own group while foraging, as well as occasionally with other groups while competing for territory. Being able to know what other chimpanzees are thinking about when they are *not* trying to gain access to something, however, requires a higher level of intentional thinking. In his book *The Intentional Stance*,

---

[72] Tomasello presented a video of this experiment at a presentation he gave at the University of Aarhus Theology Faculty on February 9th, 2006

[73] Tomasello et al 2004: 18

Daniel Dennett argued that cognitive ethologists must take care not to over-state the level of intentionality that an animal is capable of exhibiting. Dennett advocates an Occam's Razor approach to reaching conclusions about animal perception of intentionality—go for the simplest answer, even if it is a "killjoy." It seems Tomasello abides by this rule as well. Considering that it is very difficult, if not impossible, for most of us humans to imagine what it would be like to be forced to interpret other people's behavior strictly in terms of noticing what they perceive and chase, rather than also including thoughts about things which happened in the past or might happen in the future and what they think about those things, it seems doubly important not to apply a standard of "when in doubt, assume they think like us."

This view has its detractors, however. Primatologist Frans deWaal. deWaal has been a prolific advocate for chimpanzees and bonobos as representative of empathy's precursor—if not empathy itself. He argues that by studying other higher-order primates, we may discover how morality evolved. To a person who views empathy as the origin of true morality, this is an attractive proposition. But even given that, the question of what constitutes true morality betrays certain requirements for how we are to recognize true empathy. The primary criticism of deWaal's viewpoint, expressed in books such as *Chimpanzee Politics* and *Good-Natured*, is that he relies on anecdotes to illustrate the capacity of chimpanzees and bonobos to read each other's mental states. And indeed, he does—part of what makes his books so interesting to read is that deWaal describes his subjects in detail, allowing the reader to know them individually as if they were characters in a play. deWaal's writing invites anthropomorphism, and this also is not a charge he rejects. Rather than accept criticism for anthropomorphism, deWaal accuses critics of anthropodenial—"the *a priori* rejection of shared characteristics between humans and animals."[74] deWaal argues that our simplest explanation is to accept similarity between their behavior and ours as authentic, to "read" their behavior as we do automatically for other humans. "Anthropomorphism is a possibility among many," deWaal writes, "but one to be taken seriously given that it applies intuitions about ourselves to creatures very much like us. It is the application of

[74] deWaal, 2006: 65

human self-knowledge to animal behavior. What could be wrong with that?"[75]

With this perspective in mind, deWaal's theory on empathy follows: he refers to it as the Perception-Action Model, or PAM. deWaal put forth this model along with psychologist Stephanie Preston, and it entails that "attended perception of the object's state automatically activates the subject's representations of the state, situation, and object, and that activation of these representations automatically primes or generates the associated autonomic or somatic responses, unless inhibited."[76] So it is important to note that for them, empathy's essence is affective and responsive. Preston and deWaal do not deny that cognitive empathy exists (though they suggest that it might be more properly called "perspective-taking," since one can argue that all empathy is cognitive), but it can be difficult to see how it fits in with their model. The degree to which this model is based in response and affect can be seen in how it emphasizes that these representations are activated *automatically*, and their existence *automatically* generates responses.

Though Preston and deWaal present their model as uniting the many disparate approaches to explaining empathy that currently exist, it may be that they are uniting at the cost of losing important details. Though they allow for the possibility that an empathetic response may be inhibited, they mention only empathetic disorders such as autism and sociopathy as possible sources of inhibition, not recognizing possible barriers at the cognitive level. Preston and deWaal acknowledge cognitive empathy as taking place based on the perception-action mechanism, "but requiring additional cognitive capacities that develop with the pre-frontal cortex."[77] The critical requirement for these cognitive capacities appears to be self-awareness. In order to take the more abstract step that is empathizing with an object who is imagined, or who is otherwise not immediately available, we need first to have a self-other distinction. This distinction needs to be strong enough to maintain representations first of our own mental and emotional states, and secondly to represent those states in a removed other. Inhibition of the representations and the accompanying responses is not simply a matter of a breakdown in some motor capacity, but can be deliberate.

---

[75] Ibid., 67

[76] Preston & deWaal 2002: 4

[77] Preston & deWaal 2002: 5

The psychologist Gordon Gallup theorizes that animals with the ability of self-recognition use their own experience as a means to make inferences about that of others. In order to test which animals have self-recognition and which do not, Gallup initiated the "mirror test." This test involves marking the face of a sleeping animal, and then presenting it with a mirror when it awakes to see how it will react. If the animal touches the mark on its face and tries to remove it, this is understood as a sign that it is viewing *itself* in the mirror, as opposed to some other animal. Higher-order primates have passed this test, as one might expect, but there is also evidence that dolphins and elephants also have the ability. Gallup theorizes that self-recognition is a sign that a creature is capable of conceiving of itself *as a self*—that is, having a sense of continuity, a sense of personal agency, and a sense of identity.[78] Self-recognition means understanding that you have a degree of control over your own actions. In turn, this opens the door to attributing intentions and goals to others. Because we are aware of ourselves as intentional beings, beings with minds, we assume that others who are like us do as well.

The discovery of mirror neurons in macaque monkeys in the early 1990s,[79] which fire in "recognition" of a conspecific's goal-directed behaviour, has sent neuroscientists into a flurry of excitement regarding what this may mean for human theory of mind. Perhaps they are the gateway to empathy, as V.S. Ramachandran exclaimed eagerly on the pages of the cognitivist think tank web site edge.org, perhaps the driving force behind the "great leap forward" in human evolution. Ramachandran even went so far as to call them the single most important unreported story of the decade. Others have been more hesitant. How much can some neurons discovered in monkeys really tell us about *us*?

According to neuroscientist Vittorio Gallese, we have good reason to be cautious. Mirror neurons are instrumental in terms of action understanding, though there is no necessary reason to believe that they may allow us to practice so-called "mind-reading." Rather, they may be restricted to "behaviour-reading," which we shared with the monkeys in whose brains they were found originally. Like a monkey, we can recognize by watching someone grasp for something that we are witnessing a goal-directed behaviour—motor simulation is allowed to take place in our brain, which may have set the stage for

---

[78] Gallup 1996

[79] http://www.crossroadsinstitute.org/newsletter/nlarticles/sept05/autismmu.html

more advanced forms of simulation in humans. At any rate, it presents a strong argument for the simulation theory as opposed to "theory theory," of mind, which by contrast presents empathizing from a detached, rationalist perspective. This is not to suggest that "theory theory" cannot be viable, but rather than it may be a process which humans have learned to apply only once they were able to "mentalize." Gallese's goal is to use mirror neurons to reach back in phylogenetic time toward the gap between our empathic capacity and that of our more ape-like ancestors, stressing that "empathy is grounded in our lived-in body, and it is this experience that enables us to directly recognize others not as bodies endowed with a mind but as persons like us."[80]

So it is perception of *similarity* which enables an empathic connection with the other. This resonates strongly with accounts provided by deWaal and colleagues, which is not surprising since his emphasis is primarily on perception as well. So the quickest route to empathy, as they portray it, amounts to "I see that you are like me, therefore I *feel* for you." Note that this form of empathy is affective. Cognitive empathy does not happen on its own, Gallese explains, which has been a mistake of past theories—it is supplemental to and continuous with the affective form, which is based primarily on behaviour. A handicap in the ability to practice affective empathy, such as in a person with autism, might result on them having to rely completely on explicit theorizing that is associated with the more cognitive route. This would not result in giving them Vulcan-like powers of logic, but rather represent a significant impairment which makes the emotional displays of others difficult to decipher, and cause it to be difficult as well to determine what to do about them.

Gallese even goes so far as to say that when emotion enters the picture, we likely "tag" aversive motor schemas to experiences associated with negative feelings. According to Gallese, "the coordinated activity of sensorimotor and affective neural systems results in the simplification and automatization of the behavioural responses that living organisms are supposed to produce in order to survive."[81] So this coupling leads to association of a particular emotional charge through the simulation process of empathy.

---

[80] Gallese 2001: 43

[81] Ibid., p. 46

5. Understanding, misunderstanding, and refusing to understand

In this section, I will address two conditions which have been described as disabling empathy, autism and sociopathy, and discuss what effect that has on social relationships. Then I will talk about what it means to refuse to empathize, and why people might choose to do so under certain circumstances.

Autism is a developmental disorder of great interest to cognitive scientists because of what it may teach us about Theory of Mind. Psychologist Simon Baron-Cohen, a pre-eminent researching on autism, describes an autistic person as one who has a deficient or complete lack of ability or inclination to be aware of other people's mental states—which he calls "mindblindness."[82] One of the things that can lead to misunderstandings about the nature of autism and how to know if a person is autistic is the fact that autism exists on a spectrum—on one end are people who are have "low functioning," "classic," or "profound" autism and who may never speak and may suffer from mental retardation, and on the other end are people who are "high functioning," which means that though they have trouble functioning socially they are nonetheless capable of living independent lives, finding jobs which support them and in which they find pleasure, and forming relationships and establishing families. The latter have Asperger's Syndrome. In an article for Wired magazine entitled "The Geek Syndrome," Baron-Cohen suggested that the conglomeration of people with high technical skills in Silicon Valley, starting families, and passing on their "math-and-tech genes" might have something to do with the increasing numbers of children with autism. He has also argued that men are more likely to have autism than women because women have brains which are more inclined toward empathy, whereas men's are more driven toward systematizing—constructing, exploring, and analyzing systems.

Uta Frith, another leading voice in psychology and autism, published a seminal book on the subject entitled *Autism: Explaining the Enigma* in 1989, and updated in 2003. Its cover illustration is a painting by Georges de la Tour entitled "The Cheat with the Ace of Diamonds." It depicts two young men and a woman sitting at a table playing cards, and a servant girl standing next to the table. Frith uses

[82] Baron-Cohen 1995

this painting to illustrate the things we automatically infer from the image, but which a mindblind person might not see. The eyes of each of the people in the painting tell the entire story. The young man on the left is holding a pair of aces behind his back and looking steadily at his cards. The servant girl, standing to his side, is looking askance at his hand concealing the cards. The sitting woman points at the cheating young man while looking up warily at the servant girl. The other young man at the table has eyes only for his own hand of cards, appearing oblivious to the whole matter. It is possible that the cheating man is aware that the woman knows of his cheating, or perhaps he doesn't. Regardless, the impression is that the oblivious young man is likely to be cheated soon. It takes several sentences to describe what is going on in the painting, but only a moment's look at it would convey to most people everything that I have just mentioned.

However, this will not necessarily be the case for someone on the autism spectrum. Frith notes that she received an email from a young woman called "AC" with high-functioning autism who said that she spent quite some time gazing at the painting, pondering the quality of the brushstrokes, marveling at how precise the details were, contemplating the realism and skill of the artist who must have painted it. "And then," the young woman wrote, "I read the inside of the book, and I was like, What the hell? There's this whole 'soap opera' that the 'normal' person is supposed to pick on first, and this person cheating, and that person knows, and that other person doesn't, etc. it's nuts!"[83] Frith remarks that even though AC's example demonstrated a marked difference in the way that a mindblind person might view such a painting, it also suggests that the difference is not necessarily negative—a non-autistic viewer might fail to notice the skill of the artist entirely, or to a lesser extent, because she was too caught up in the story being played out in the image. Of course, this difference in emphasis on what is important can lead to a great deal of social awkwardness and frustration. Frith relates an excerpt from a 1983 paper by Sir Michael Rutter, who referred to a young man who

> …complained that he could not mind-read. Other people seemed to have a special sense by which they could read other people's thoughts and would anticipate their responses and feelings; he knew this because they managed to avoid upsetting people whereas he was always putting his foot in it;

[83] Frith 1989: 79

not realizing that he was doing or saying the wrong thing until after the other person because angry or upset.[84]

As mentioned earlier in this chapter, children with autism experience delays in acquiring language or may never begin speaking at all. The image of an autistic child is of someone who keeps to herself, who seems to exist in her own private world separate from the one around her. She is noticeably less affectionate than other children, and may be unwilling to look people in the eye or avoid physical contact, even hugs from her parents. As she grows up she will have a hard time functioning socially, and will find herself hurting people's feelings by her apparent callousness and awkwardness. She will prefer to communicate literally and become confused or annoyed when others do not do so. She may have trouble understanding jokes, irony, or metaphors. Frith characterizes autism as primarily a disorder of communication. By this she does not mean that people on the autism spectrum suffer from impairments in ability to use grammar—even in those who have had language delays show no problems in that area. Rather, she means that autism is characterized by a theory of mind deficit which makes it hard to learn language, and even after mastering language to be poor at understanding the point of questions, to have insight into the minds of other people in order to be able to know what to say to them.[85]

Frith uses the term *mentalizing* to describe "the ability to predict relationships between external states of affairs and internal states of mind."[86] Mentalizing is what one does when examining the previously described painting in order to notice and reflect upon the subterfuge it depicts. It is what is required in order to conclude that Sally is going to look in the wrong place for her ball after Anne has hidden it—the scenario of Sally and Anne is a mentalistic story. If the story instead depicted a man going to the store and giving the clerk some money and the clerk handing him a candy bar, that would be a behavioral story. Behaviorism is a type of psychological approach in which behavior is explained without reference to internal mental states—mentalizing is irrelevant to this approach.

In order to determine whether children with autism are behaviorists rather than mentalists, Simon Baron-Cohen conducted an experiment in which he presented children with cartoon pictures

---

[84] Ibid, 79

[85] Frith 2007

[86] Firth 1989: 77

which told different stories, only the pictures were scrambled and the children were asked to put them in order. In addition to a behavioral story and a mentalistic story, subjects were also presented with a mechanical story, which did not depict humans acting at all but rather a balloon gradually floating toward a tree and popping on its branches. If mentalizing is a problem for autistic children, went the hypothesis, then they should be able to put the mechanical and behavioral stories in order but not the mentalistic one. As it turned out, that is exactly what happened. The mechanical story was understood by the children with no problem. Likewise, the behavioral story, which depicted a girl going to a shop to purchase something, was not a problem for them—it could be described easily without any reference to what was going on inside anyone's head. The mentalistic story, however, presented problems—the children just could not understand it. That story was modeled on the Sally-Anne experiment, and the same children who had been involved in that experiment also participated in this one. This story depicted a boy placing his chocolate in a box, and then going outside to play. While he is gone (and therefore without his knowing), his mother sneaks in and eats the chocolate. Upon his return, the boy opens the box and is shocked to discover that his chocolate is not waiting for him. The same children who had had trouble in the Sally-Anne experiment, tellingly, had trouble here as well. "Their 'stories,'" says Frith, "were told without the attribution of mental states or appreciation of the naughtiness of eating the chocolate. For example: 'A boy plays football. He puts some chocolate into a box. His mother eats the chocolate. He opens the box. It is empty."[87]

What does all of this mean in terms of the differences in the way autists see their social world as compared to how non-autists see it? It means that there is something invisible which has an important effect on the way people behave, but which autists have trouble recognizing. They can't literally see inside people's skulls, but neither can they decipher the inscrutable part of other people that makes them act and react in unexpected ways. This means that an autist's theory of mind really is more likely a scientific theory—it has to be acquired explicitly. Simon Baron-Cohen has developed resources for people on the autism spectrum to help them be able to "read" facial expressions better, especially eye movements. Books by famous people with autism such as Temple Grandin attempt to help others in their lessons

[87] Ibid., 88

on acquiring a skill that for most other people is an inborn talent. I have spent some considerable time watching and interacting with an online community for people with Aspergers, and one continuing source of frustration with "neurotypicals" (their name for people who are neurologically average, non-autists) is that they *don't say what they mean*. That is, the "NTs" use euphemisms, colloquialisms, and other expressions that confuse them, or get offended when an autist insists on taking them literally. One young woman complained about how as a receptionist, her job occasionally required her to go to the office of people in her business and announce the arrival of someone who had come to visit them. Upon hearing that Joe Johnson had come to see her, the colleagues apparently would frequently reply "And who is that?" This mystified and irritated the woman, because after all, she had just *announced* who it was.

Research on beliefs concerning supernatural agents for people on the autism spectrum is for all intents and purposes non-existent, as far as I am aware. This is strange considering that finding out how people who tend toward behaviorist interpretations of human behavior rather than mentalizing in order to "read" invisible goals and beliefs in other people's minds think about gods and spirits could be instructive in gaining a better understanding of the cognitive roots of belief in such agents. If autists do not have the same theory of mind as non-autists, does that have an effect on which religious concepts they find compelling? Are some harder or easier to believe in that others? The information we do have on this subject so far is mainly anecdotal. In a 2002 paper which discussed the subject, Jesse Bering quoted Temple Grandin:

> In nature, all particles are entangled with millions of other particles, all interacting with each other. One could speculate that entanglement of these particles could cause a kind of consciousness for the universe. This is my current concept of God.[88]

I performed a brief casual survey amongst the people in the Aspergers online community and found that while nearly 20% reported that they were Christian, 35% refused to categorize themselves as adherents of any particular named religion, though they did describe themselves as "spiritual." Another 30% said that none of the terms "spiritual," "religious," or "mystical" described them. 37% reported that they did not believe that any aspect of a person's mental states continues

[88] Quoted in Bering 2002: 14

beyond death. 57% said that they never imagined what a person who had died might be thinking or doing right now.[89] I find these responses intriguing, and suspect they indicate that people on the autism spectrum are not as likely to believe in personal gods or are more likely to put their own spin on religious teachings in order to make them "work." Exactly how this happens, however, remains to be seen.

Sociopaths are also people who are described as being deficient in empathy. However, their story is quite different in this regard since they are well aware that another person can have different beliefs, especially false beliefs. It is necessary to have a good grasp of the mental states of others in order to deceive, but also important to lack the affective element which causes you to sympathize with them—at least long enough to lie. It is for this reason that psychologist Hank Davis calls deception the "evil twin of empathy."[90] A sociopath may not have anything close to the same feelings as other people around them at a given time, but must act as if they do. A high-functioning autistic person must not only do this, but also take pains to learn what certain social cues mean in order to act in a way that doesn't confuse or offend those around them.

But a person need not be either autistic or sociopathic in order to be inhibited in practicing empathy. There are a number of top-down techniques, conscious and unconscious, that humans practice in order to *avoid* empathizing with each other.

In response to deWaal and Preston's paper detailing their PAM theory of empathy, psychologist Albert Bandura had this to say:

> "Deficient empathicness is a pervasive phenomenon rather than confined to pathologic types, as evident in the widespread inhumanities that people perpetuate on each other. Otherwise considerate people selectively disengage empathic restraints and moral self-sanctions in executing destructive activities in the name of religious doctrines, righteous ideologies, and nationalistic imperatives. . It is accomplished by cognitively restructuring injurious conduct so that it can be done free from the restraint of empathy and self-censure. . . The prevalent failures in empathic control stem from ideology rather than impaired biology."

---

[89] Unpublished survey of 82 people with unconfirmed diagnoses of Aspergers

[90] Davis 2002: 33

When we consider that familiarity with and similarity to someone else is a powerful predictor of our ability to have empathy for them, the question of what happens when we *refuse* to empathize inevitably should arise. Refusing to understand someone-- or at least refusing to admit that you understand them-- can be taken as a statement that you are *not* familiar with, not similar to them. A refusal to empathize draws a clear line between self and other. To illustrate how this takes place, I will give an example of one such refusal.

Recently Christopher Hitchens, author of *God is Not Great* and vociferous critic of religious faith, appeared on the radio interview show Open Source to discuss his book. He was confronted there by Princeton associate professor of religion Eddie Glaude Jr. with the idea that religions are not properly presented when reduced to a set of truth claims, and that doing so makes nonsensical such phenomena as black slaves in America adopting the same faith as their captors. Most religious people, Glaude Jr. pointed out, are not literalists, which means that taking stabs at the literal truth value of their doctrines should not be taken as a refutation of the value of religion per se. Rather, people have motivations which cause them to reach different interpretations of the text from which they derive meaning, rather than adopting it wholesale along with the interpretation handed to them. The response from Hitchens was one of contemptuous bafflement: "You can make up anything like that as you go along as far as I can see, but it's just white noise. White noise. White noise, what you're talking to me, absolutely meaningless…Nothing you say…I just have to say that I make my living by scrutinizing words. It's what I do. Not a single thing of what you just said made a word of sense to me." I found this both comical and very revealing. Incomprehension makes for a subtle weapon, though Hitchens is rather blatant here about the move being made—from "I don't understand" or "I don't get it" to "Your position is nonsensical and not worth considering."

There are two interesting psychological elements involved here—attribution bias and the inhibition of cognitive empathy. Attribution bias is commonly understood as taking credit for your successes while blaming your failures on others. Humans tend to be overzealous detectors of agency, to the point of asserting intentionality on the part of others for the purpose of assigning blame. The actor-observer bias is a kind of attribution bias involving the diminishment of our own responsibility for behaviour while over-emphasizing that of others ("You did that on purpose!"). The straw man of asserting that religious believers—or the only religious believers that matter—

take the claims of their sacred texts to be fundamental and the most important aspect of their belief is an example of actor-observer bias. It seeks to hold people accountable to what they don't actually believe, whether they have claimed to do so or not. Cognitive empathy (the capacity to know the content of another person's mental states) is inhibited through the denial of understanding. Not understanding someone's position, or pretending not to understand it, makes the implicit statement that you refuse to play their semantic game, to place yourself in their system of meaning. Comedian Bill Maher suffered the results of this tendency following the 9/11 attacks when he attempted to accurately represent the content of the terrorists' mindsets, saying that he didn't think "cowardly" was a fitting term to describe an individual willing to go to his death for his beliefs. Understanding the terrorists was, for the time being, off the table—it got too much in the way of blaming them as emphatically as possible. In order to avoid the perception of being similar to someone whom they despise, it's not strange to consider that a person would desire to increase the distance between him or herself and someone considered wrong-headed in some way by emphasizing a *lack* of empathy for them. Condemnation and empathy, therefore, appear to be difficult to practice simultaneously. Certainly, a person who wishes to convict someone entirely will avoid making any statement which betrays a common ground, seeking instead to make the difference between them as apparent as possible.

The primary trigger of both affective and cognitive empathy, according to the simulation model, is our perception of similarity with others. Through imaginative projection, we observe or contemplate the state of another person, and through introspection manage to create in ourselves a similar affect, mental state, or both, which allows us to identify with them. A combination of the fact that another human's mental state requires mentalizing, rather than simply observing behaviour, and the fact that our introspection makes it seem that we are rational people who make calculated judgments, makes it easier for us to attribute to others an immaterial soul just as we presume ourselves to have. We recognize a similarity with them on that basis, which may stir us to make empathic moral judgments. To preserve ingroup versus outgroup boundaries, we may portray the "others" both foreign and lesser. When we have successfully relegated a person to this status, we need not be troubled by the stirrings of sympathy, because the perception of similarity to us has been removed.

"Negative empathy" is a term psychiatrist George Ainslie uses to describe the savoring of someone's suffering—not simply the dispassion that allows some instances of cruelty in war, but rather the actual enjoyment of another's misfortune. Ainslie characterizes this move as an effort to draw a firm line between ourselves and others who exhibit characteristics which we fear manifesting in ourselves. It is only possible for humans, that is, possessors of theory of mind, because it requires the ability to acknowledge the potential of our mental states coming to resemble others and becoming aroused with fear and dread of such an event. The experience of negative empathy is a commonplace event, according to Ainslie—he gives as examples "the boor getting his comeuppance, the driver who cut us off stopped by the police, and the pretensions of the poseur punctured, as well as less respectable examples such as *schadenfreude* and our minor persecution of people whom we hope we do not resemble."[91] Negative empathy requires us to consider the fact that we *could have been* the one suffering, but some difference means that we are not—and that recognition gives us pleasure, even if we ourselves are the ones who cause the suffering.

6. Conclusion

What I have tried to convey in this chapter is that our empathic response, our imaginative projection of other people's agency is possible because we are able to recognize in them a source of very real but invisible thoughts and feelings which affect how they behave in the world. The relationship of the development of empathy to that of language acquisition is closely linked, because the fostering of joint attentional scenes from a very young age makes it possible for children and their parents to focus their common attention on something outside of them, even something neither of them can see. By attaching words to these outside items, a child can learn words. Even before that point, establishing joint attention is a pleasurable activity for most children—it is something they begin to instigate without prompting, and are disappointed when a parent does not respond as they expect. As they grow, children begin to understand that the contents of their own minds and those of everyone else are not

[91] Ainsle 2006: 225

the same, and that other people can have false beliefs about things as well as different emotional states. The child can grasp moral rules such as not to hit by recalling what it is like to be hit oneself. Cognitive empathy requires a self-other distinction and an ability to mentalize rather than strictly viewing the acts of others in behaviorist terms. Chimpanzees are behaviorists, and to a certain extent that appears to be characteristic of autism, though autists can develop their "theory theory" of mind explicitly over time.

These deficiencies when it comes to practicing empathy do not make autists bad people, though they may be socially awkward and occasionally say the wrong thing because they have not intuitively picked up on what the "right" thing is. Sociopaths, on the other hand, tend to know what the "right" thing is, but lack the affective element of empathy which would prevent them from using others for their own purposes. The deliberate refusal to empathize can occur when a person wants to establish how very different he is from someone else—to deny similarity, as similarity and familiarity are two primary triggers of empathic reactions. Acknowledging that you are like me indicates that I think we are similar on the inside, and so the invisible things that drive our goals and our emotions are made of similar stuff. Refusing to acknowledge this suggests that you are deficient in some way, incomplete or lower. Later in this dissertation, I will outline a particular form that refusal to empathize can take.

Chapter 3

Supernatural Responsibility

In this chapter I will be focusing on how moral evaluation can affect attributions of cause and intention in our folk psychology, and what this indicates about our intuitive idea of the self.[92]

1. Causal attribution and morality

Causal attribution in morality entails assigning blame—if an individual can be determined to be the primary cause of a negative event, then that person is deemed blameworthy for the event. Negative causation, or omission, can also be an identifier of blameworthiness. If my job is to push a button every day which will prevent an island science station from imploding (to take an example from the television show *Lost*), then I can be held blameworthy if one day I neglect to push the button and the station implodes. There are, of course, many different causes for this event that can be articulated—you could say that the cause was the fact that a sudden demagnetization of the station was what caused the implosion, but that would be like identifying the presence of oxygen as the cause of a fire. Both are, of course, critical factors in the occurrence of the event, but neither is an appropriate candidate for moral blameworthiness. If we hold that moral responsibility entails causation, and I as an agent am judged to be the primary causal factor, then I am morally responsible for the implosion. If the implosion were judged to be a good thing, then I would not be deemed blameworthy.

Causation is one factor in determining blameworthiness, but it is not the only one. We generally do not hold people responsible for

[92] Portions of this chapter have been taken from Koch 2008b

negative events they cause by accident, unless it is through neglect. You most likely would blame me for deliberately withholding food from my dog, but probably not if I simply forgot to feed him. I am not responsible for feeding every dog in the world, but because I have adopted this dog and taken on the role of caretaker for him, I am morally obligated to see to it that he gets fed. If I forget to feed him for a sustained period and he starves to death, then I am guilty of negligence and would be blamed alongside someone who intentionally starved their dog to death. But here causation comes in again—let's say that I share responsibility for feeding the dog with my housemate, Bob. We alternate monthly, and it's my month to feed the dog. I do not, and the dog starves. What is the cause? Most people would say it is my not feeding the dog, even though Bob did not feed him either. Likewise, if Bob and I share the task of giving the dog medication and it is Bob's turn to give it to him but later I give him some as well, then I would be held the primary cause of the dog having received an unhealthy dose of medication even though Bob also gave him some. My moral responsibility for the dog's poisoning makes me, intuitively, the primary cause of it.

2. We're blaming him; he must have done it

A number of experiments have been performed to date which substantiate this tendency to assign causation according to moral blameworthiness, and I will be discussing them in this chapter—as well as what this indicates for folk intuitions of responsibility. M.D. Alicke, for example, asked subjects to consider a young man named John who gets into a car accident. John has been speeding, and Alicke supplies different reasons for which this is the case—one is that John is rushing home to hide an anniversary present which he bought for his parents but has been left out in the open. Another is that John is trying to get home in time to hide a vial of cocaine from his parents before they have the chance to see it. Alicke offered a number of other factors which caused the accident, including an oil spill which counteracted John's attempt to brake the car before hitting the other vehicle, a tree branch which obscured a stop sign and therefore prevented John from realizing that he needed to stop, and the other driver having ignored a stop sign intended for him. Alicke found, however, that in each case John's having a socially undesirable motive

for speeding (hiding the vile of cocaine)—the more "culpable" one--caused subjects to be far more likely to identify him for the primary cause of the accident than those scenarios in which John's motivation was socially desirable (hiding the anniversary present). As Alicke put it, "With causal necessity, sufficiency, and proximity held constant, the more culpable act was deemed by subjects to have exerted a larger causal influence."[93]

Philosophers Joshua Knobe and Arudra Burra performed a study to examine perceptions of intentionality in scenarios with a moral valence. The results indicated that when an action taken by an actor in a hypothetical situation has harmful consequences, respondents are more likely to attribute intentionality to the actor than otherwise. To give an example:

> The vice-president of a company went to the chairman of the board and said, "We are thinking of starting a new program. It will help us increase profits and it will also help the environment.' The chairman of the board answered, 'I don't care at all about helping the environment. I just want to make as much profit as I can. Let's start the new program.' They started the new program. Sure enough, the environment was helped.

Now compare:

> The vice-president of a company went to the chairman of the board and said, "We are thinking of starting a new program. It will help us increase profits and it will also harm the environment.' The chairman of the board answered, 'I don't care at all about harming the environment. I just want to make as much profit as I can. Let's start the new program.' They started the new program. Sure enough, the environment was harmed.[94]

If you're like the majority of the respondents, your reaction is to attribute greater intentionality to the vice-president in the latter scenario. They also assessed the vice-president as more deserving of blame in the latter scenario than he is deserving of praise in the former. Knobe and Fraser presented another group of subjects with a scenario in which the receptionist of a philosophy department kept a supply of pens at her desk, which administrative assistants were allowed to take. Professors, however, were obliged to buy their own.

---

[93] Quoted in Driver 2008: 427

[94] Knobe & Burra 2006: 117

One day, an administrative assistant comes by the desk to take a pen, and Professor Smith takes one along with her. Later that day the receptionist needs to take down an important phone message, but can't find a pen. Upon being asked who caused the problem, subjects agreed that the professor had caused the problem and disagreed that the administrative assistant had caused it. This was the case even though the scenario had specified that professors routinely took pens from the receptionist's desk, despite repeated admonishments against it—this addition intended to disqualify the atypicality of a behavior as a possible explanation for why it was viewed as the problem.[95] A similar response was given to the following vignette:

> Lauren and Jane work for the same company. They each need to use a computer sometimes. Unfortunately, the computer isn't very powerful. If two people are logged on at the same time, the computer crashes. So the company decided to institute an official policy. It declared that Lauren would be the only one permitted to use the computer in the mornings and that Jane would be the only one permitted to use the computer in the afternoon. As expected, Lauren logged on the computer the next day at 9:00 am. But Jane decided to disobey the company policy. She also logged in at 9:00 am. The computer crashed immediately.[96]

Knobe points out that in this scenario, Jane is the one we identify as the cause of the crash even though the crash would not have happened if Lauren did not log on. The normative quality of Jane's behavior—the fact that she is the one who broke a rule—causes her behavior to become the most salient. As Knobe puts it, the moral dimension of Jane's behavior causes it to obtrude all other causal factors in the event.

3. Moral nativism? Maybe

This tendency to view normative violations as particularly salient and special has been treated as a possible indicator of an innate moral faculty. Moral nativism is a hot topic these days, to the degree that it seems there is no scientist or philosopher unwilling to comment

---

[95] Knobe & Fraser 2008: 443

[96] Knobe 2005: 6

on it. A recent issue of New Scientist was even prompted to ask on its cover, "If morality is hard-wired in the brain, what's the point of religion?"[97] Moral nativists don't merely assert that we are innately capable of being good people, but that we have a dedicated system or tendency to make moral judgments. Depending on which moral nativist you ask this may take the form of explicit moral rules which have contributed to our survival and therefore have been adaptive in our evolution ("Don't have sex with close relatives"), or it may take the form of generalized domains of moral judgment such as Haidt et al advocate. Marc Hauser has argued that the moral faculty is similar to the language faculty, in that it involves rules which we apply regularly (as we apply grammatical rules) without being able to articulate them or even necessarily be aware of their existence. He and colleagues attempted to illustrate this by using different permutations a philosopher's thought experiment called the trolley problem. The basic details of the trolley problem are as follows:

> Denise is a passenger on a trolley whose driver has just shouted that the trolley's brakes have failed, and who then fainted of the shock. On the track ahead, there are five people; the banks are so steep that they will not be able to get off the track in time. The track has a side track leading off to the right, and Denise can turn the trolley onto it. Unfortunately there is one person on the right track. Denise can turn the trolley, killing the one; or she can refrain from turning the trolley, letting the five die. Is it morally permissible for Denise to switch the trolley to the side track?[98]

Most respondents will reply that yes, it is morally permissible for Denise to do so. However, if presented with a slightly modified scenario in which a man standing on a footbridge over the trolley's track, Frank, has an opportunity to stop the train by pushing a fat man who is standing next to him down onto the track so that the trolley will kill him but save the five, most people will reply that that it is not morally permissible for him to do this. This has been the case for English-speaking respondents from many different countries in which Hauser and colleagues have presented the problem, regardless of age, ethnicity, gender, religion, or level of education. They note, however, that only 30% of respondents could offer a "sufficient justification"

---

[97] 1 September 2007

[98] This is the phrasing used by Hauser et al in 2008: 128

for their differing opinions of the moral permissibility of Denise and Frank's behavior. More likely they tended to refer to gut feelings or say something like "It just struck me," which sounds like what Jonathan Haidt has called "moral dumbfounding." Hauser takes this as sign of the workings of an unconscious moral grammar. Neurologist and philosopher Joshua Greene suggests that we may have evolved an aversion to directly physically harming someone (such as pushing the fat man off the bridge) which kicks in and causes us to judge that action as morally impermissible even though it would save the lives of five other people, but which does not prevent us from judging it permissible to harm someone in a detached way, such as redirecting the trolley to hit one person rather than five. Technological innovation has created myriad ways for us to both harm and help someone at a remove that weren't possible in our evolutionary past, which means that we simply may not have adapted to have the same intuitions about such behaviors as harming or helping someone who is standing right in front of us.

It is also possible that the hypothetical scenarios tug at our intuitions in different ways which aren't being taken into account. In response to the trolley scenario as presented on one blog, a woman commented "No, I would not push a bulky person in front of a train in attempt to stop it to save five. Why? Physics: there is no way in God's green earth that a 200- or 300-pound person is going to stop several thousand tons going even 30 miles an hour. Children should be raised with that understanding, and anyone who does not understand that should go back to grade school. . . Yes, I am not in the 'spirit' of the question. Yet the moral dilemma was presented to me to THINK about, to evaluate and respond. So I did."[99] She has a point--if we are trying to discover unconscious moral intuitions, respondents cannot be expected to frame their answers in the "spirit" of the question. Most people do not contemplate moral hypotheticals on a regular basis in the manner of logic students, being careful to take every statement as true and add no extra assumptions. Even if they did, that would not necessarily prevent a small voice in the back of their mind from saying "What if the fat man doesn't actually stop the train? What if by pushing him, that just means that Frank is killing six people instead of five? Isn't Frank doing something wrong even in assuming that won't happen?" This is an inevitable problem with trying to discover moral intuitions by presenting hypothetical

[99] Hatcher 2008: 34

scenarios. It would still seem that the best way to determine a person's moral judgment of a situation is to ask them outright, but added assumptions can't be discounted if neither the tester nor the respondent realizes that they are there to begin with.

Moral empiricists such as philosopher Jesse Prinz are skeptical of the nativist argument. Prinz argues that unlike language, there is no critical period at the beginning of life in which one's moral capability must be developed. Even feral children can learn to judge right from wrong and obey rules. Children demonstrate an ability to distinguish between moral and prudential infractions, but Prinz argues that this can easily be explained by the very different reactions that parents show to these infractions from the beginning. Children may agree that it's okay to sit on the floor rather than in chairs at school if the teacher says it is okay but that it's not okay to hit someone even if the teacher says it is, but this can be explained by the fact that from birth their own parents have treated such behaviors very differently, and a child can extrapolate from this conditioning which behaviors will fall into which categories. If hitting your brother is wrong, and hitting the cat is wrong, then hitting Sarah at the next table is most likely wrong as well, and will stay wrong no matter what the teacher says. Prinz grants that emotions certainly back our moral intuitions, but maintains that this doesn't mean there are necessarily *moral emotions*. We can get angry, disgusted, and contemptuous without moral components to those emotions. They take on a moral valence when applied to someone else, constituting blame for something which would inspire guilt if directed toward ourselves. This makes moral emotion and judgment an *inevitable* part of living socially, but not necessarily innate.

Regardless of whether we have an innate moral faculty or not, New Scientist's question of what the point of religion is if morality is "hard-wired in the brain" seems odd on its face, and reveals a problematic assumption about the thesis of moral nativism. Nativist Susan Dwyer frames it nicely: "People are apt to ask, 'If morality is innate (read: if moral nativism is true), why do people behave so badly?' But this is like asking 'If language is innate (read: if linguistic nativism is true), why do people speak and write so badly?"[100] Moral nativism is not the thesis that people are innately disposed to behave morally. It is the thesis that we are innately disposed to make moral *judgments*. One may argue that we are unable to truly be moral if we

[100] Dwyer 2008: 409

can't make judgments about morality, but that would apply equally to our capacity to be *im*moral. So New Scientist might just as well have asked, "If morality is hard-wired in the brain, what's the point of punishment?"

I am going to refrain from taking a position on moral nativism here. At this point, I will not be surprised if it turns out that we have some form of innate moral faculty, and I will not be surprised if it turns out that we don't. It may even be that we once did, but through the course of evolution have to some extent "exported" the task of moral judgment to our moral social enforcers, the keepers of the mores, whether they be our parents, the elders, the judges and juries, the bioethicists. An evolved sense of right and wrong may tell us that we should not strike someone who has not harmed us, but it cannot be expected to tell us how to feel about global warming. And this may explain a good deal about why a politician's sex scandal can make national front-page news, but his stance on the ozone layer will barely make a two-inch column on the tenth page. Humans have been grappling with issues like sexual morality for many thousands of years, but only recently in our history have we become aware that there *is* something called the ozone layer, let alone what should be done about it.

4. Folk fears about responsibility, part 1: determinism and reductionism

As we learn more about the cognitive side of moral decision-making, the fear that reductionism will eliminate true responsibility continually rears its ugly head. The libertarian magazine Reason asserts that "You can't see why on an fMRI"[101] in response to the question of whether the insanity defense, inability to determine right from wrong, will see an increase in usefulness if you can point to a blotch on a brain scan to show that your client has a condition which means he can't be held responsible for his crime. Even though the very small numbers of defendants who are deemed not guilty for reasons of insanity are escorted to a mental hospital and not the front door, the belief that we can come to know enough about the workings

[101] Doherty 2007:34

of the mind to negate responsibility is a continuing specter. "The fix is in!" Tom Wolfe proclaimed in a recent essay on neuroscience and evolutionary psychology entitled *Sorry, But Your Soul Just Died*. "We're all hardwired! . . . Don't blame me! I'm wired wrong!"[102] In a text rife with exclamations, Wolfe describes brain imaging research, population genetics, and modern psychiatry as a part of an age "in which science is a court from which there is no appeal." Science, he reports, is stealing the inner self—complete with self-control, conscious determination, free will. First we were told that social conditioning makes people what they are and explains everything they do, and now we're being told that a little investigation into the brain accomplishes that task. Either way, we can no longer properly hold people responsible for their actions. But at least social conditioning can be changed—evolutionary psychology and neuroscience are far more threatening because they imply that our personalities are set in stone. "My brain made me do it," as Paul Bloom puts it.[103] As Bloom's phrasing implies, these fears rely on the implicit assumption that the self is something *other* than the brain, something which can own or inhabit the brain, but is not identified with it. This self is what is responsible for one's actions—not the brain, and not the genetic recipe which built it, and which was in turn altered and shaped over millions of years of evolution. Eliminate this self, this soul, and you can't blame (or praise) anyone for anything.

The problem with this kind of thinking is that, when you hold it down and examine it, it doesn't actually leave the soul anything to do. If the soul is something independent of both brain and environment, then it becomes nonsensical to speak of it as being affected by or causing anything. Our decision-making requires us to make decisions *about things*—things in the world which we encounter, things which pass through our brain in the form of chemical messages between neurons (even if we're unaware of them). If the soul is divorced from both this meaty electrical environment and the much more airy one outside of the skull, then it effectively becomes a mute pundit. There is nothing upon which it can even comment. Those who fear epiphenomenalism, the belief that our conscious thoughts have no effect on our actions, should properly regard belief in a soul as the biggest threat.

---

[102] http://www.orthodoxytoday.org/articles/WolfeSoulDied.php

[103] Bloom 2006: 209

Some take solace in the notion of quantum mechanics, and its implication that on the microscopic level, there is randomness. Randomness seems like an escape route from the prison of mechanistic determinism portrayed by the rest of science. But true randomness doesn't imply freedom—it implies unpredictability. A person whose mind was truly random would be completely unresponsive to positive or negative incentives to behave in any way. Rewards and punishments would have no effect on his actions. He would have no particular motivation to be a good person or bad. There is no reason why a random or mysterious brain should be more easily held responsible than one which is biologically deterministic, since randomness means true loss of control, and we can't hold people unable to control their actions responsible for them.

Neuroscientist Michael Gazzaniga points out a problem with using the insanity defense for a schizophrenic person, for example—"Their rate of violent behavior is not above that of the normal population, especially when they're on their medication. So, if that's true, how can you use that as a defense, that they're doing something because they're insane?" Steven Pinker explains it thusly in *How the Mind Works*: "The ethics game treats people as equivalent, sentient, rational, free-willed agents, and its rules are the calculus that assigns moral value to the behavior through the behavior's inherent nature or its consequences. Free will is an idealization of human beings that makes the ethics game playable."[104] To put it another way, the concepts of free will and responsibility are decisions we make in order to get along. They are useful on that level, which does not preclude them being informed by science. This debate is really over a form of the naturalistic fallacy, since it implies that once the "is" of human cognition is understood, it becomes the "ought" as well—we cannot question or judge what was determined to happen through the interaction of mind and environment. But if a person's behavior can be predicted from a chain of prior events, that does not mean that he or she is no longer responsible for them. It just means that we may have a better insight into what causes harmful behavior, and how to stop it.

One fear, as has already been discussed, is that explaining the causal chain of events that lead to our decisions down to the neuronal or genetic level—that is, a reductive explanation of decision-making—will compromise our ability to hold people responsible for

[104] Quoted in ibid, 43

their behavior. Philosopher Eddy Nahmias[105] argues that this fear is an intuitive one, but that determinism is a separate factor which we *don't* fear, when properly described. Most people don't view the behavior of others as causeless, or random. It is not disturbing to contemplate that John refused to speak to his brother Bob because Bob did something that angered him earlier. We do fear several things which determinism does not, in fact, entail: fatalism (the belief that no matter what we do, we can't change what will happen), or coercion (the belief that the past, or natural events, force our decisions). Or, for that matter, epiphenomenalism (the belief that we are effectively on auto-pilot; that our conscious choices do not matter) which is often taken as the result of reductionism. The problem is therefore, Nahmias argues, in misconstruing determinism. Determinism is simply the belief that "there is at any one instant exactly one possible future."[106] But this doesn't entail, for example, that nothing we do matters, that our decisions are meaningless, or that we are not responsible for our actions. Determinism doesn't prevent consciousness from being causal. What we really fear is not that our actions are part of an unbroken causal chain of events, but rather than in the process, our role as a conscious decision-maker is somehow usurped.

Of course, if it's really true that determinism is not what we fear, we may well ask how come it can be so easy to mistake it for reductionism. The idea that a prophet or a fortune teller can say what will happen for us in the future is not disturbing—indeed, many people pay great sums of money to have their fortune told, and are comfortable with the idea that there are categories of people whose behavior is similar because they were born in a certain month or year. But the thought that certain behaviors may be genetically or neurologically determined is frightening, and seems to compromise the role of the will. Why the disparity? I would suggest that the difference is a matter of bottom-up versus top-down, skyhooks versus cranes. Teleological thinking can be comforting existentially whether the plan is God's or the universe's, but "genetic determinism" is an epithet because it portrays us as biological machines, piloted by our mindless chromosomal overlords. Richard Dawkins did not intend his book *The Selfish Gene*, more than thirty years old now, to send the message that our genes "want" us to be selfish, or that our goals are

[105] Nahmias 2006

[106] Dennett 2003: 102

merely a deluded impression that we're the ones holding the reigns. But that did not prevent the book from being taken for such. The difference between proximate and ultimate causation (our own reasons for doing things, as opposed to an evolved motivation to do them) appears to be confusing and threatening, but we find meaning in the idea of a transcendent plan for our lives which comes from above. The need for such a plan most likely explains a good part of opposition to the concept of evolution. The thought that modern humans have not existed for the first few million years of life on Earth, and that our ancestors were much simpler creatures, is antithetical to the belief that our existence is specially planned.

5. A thought experiment

For those who are skeptical or even offended by attempts to describe human cognition with reference to evolutionary origins, the blurring of the line between humans and animals is a constant source of concern. Looking at human nature as something which may not have always existed as it does now, but rather developed over time from something simpler, seems to imply a serious limitation—that as animals, we too exist on a short leash of genetic possibility, resulting in a similarly short cultural leash. But when discussion of languages comes into play, the distinction between humans and other animals—even those who are our near genetic cousins, chimpanzees and bonobos—the difference becomes startlingly obvious. Though the question of what evolved first between self-awareness and the ability to manipulate symbols is still hotly contested, the fact that we have both means that the environment in which we as humans evolve is a cultural one as well as a physical one.

In his book by the same name, Richard Dawkins argues for a concept he refers to as the "extended phenotype." In contrast with the traditional conception of phenotype which refers to the outward physical manifestation of one's genetic code, Dawkins widens this to include those products of the behavior of an organism which arise because of its genetic makeup, and which affect its chances of survival positively or negatively. By this standard a beaver's extended phenotype would include not only its sharp yellowed teeth for gnawing, but also the product of its labors—the dam it builds, the flooded lake in which it lives. It is important to consider that the

extended phenotype as conceived here is not limited to the individual beaver, but may be a communal behavior among several. For humans, the extended phenotype may include the artifacts on which we depend in day-to-day life. But what of language itself, of the semiotic jungle through which we habitually navigate without a second thought? Is it an element of our extended phenotype, or has it become the environment itself, in which our phenotype is challenged and our survival ability tested? It would seem that the answer is both, for reasons to be discussed below.

To be human is, effectively, to think outside your own head. This is the argument of situated cognition theory, otherwise known as the extended mind thesis. Indeed, philosophers Andy Clark and Rob Wilson go so far as to say that the traditional individualistic conceptions of cognition, which stops at the boundary of one's own skull, do so arbitrarily. There is no reason to define cognition as internal only, considering that the toolkit no longer is. If the space in which cognition takes place includes memory aids, records, and other forms of "scaffolding," in the sense in which Vygostky proposed, then those artifacts are not simply external props, but part of the very structure of thought itself. Exploitation of external resources is a primary theme in Clark's book *Being There*, though he does not go so far as to agree with Daniel Dennett that public language can be considered as much artificer as artifact. Cognition may extend beyond the skull for Clark, but he himself remains an individualist to the extent that for him, the individual human remains the agent who acts and thinks, and language a tool of the mind. But why not follow this thesis to its logical conclusion? The vast majority of artifacts which aid the mind are not created by the individual who uses them. The only reason that humans have the luxury of becoming "dumb," as Clark puts it, by outsourcing so much of their cognition, is the fact that semiotic artifacts represent an accumulation of knowledge of a vast number of people over time. It is why a student may learn more by age thirteen today than one the same age would have known even in the last generation. If the individual skull and skin are an arbitrary place at which to believe cognition ends, then what is the justification for stopping at the boundary of the skulls and skin of other people?

It is not necessary to argue for the existence of a "hive mind" in order to recognize that more cognitive energy is devoted to discerning and influencing the state of mind of others than anything else. This becomes immediately evident from the behavior of those whose theory of mind abilities are impaired, causing them to be

unable to accurately attribute beliefs and desires to others beyond those which are plainly obvious. Simon Baron-Cohen, who was discussed in the previous chapter, theorizes that this lack of ability to read minds may be the reason that they are notoriously slow to begin talking and reading, because the acquisition of such skills requires the ability to attend to external things—symbols, shared referents—with other people, in this case the parent who is earnestly trying to convey information which is simply not absorbed. According to cognitive scientist Mark Turner, language is "an instrument used by separate brains to exert influence on each other, creating through biological action at a distance a virtual brain distributed in the individual brains of all participants in the culture."[107] But this description of a "virtual brain" is troublesome, because it implies that individual brains work in harmonious agreement in pursuit of a common goal, when obviously real life does not reflect such an arrangement. Indeed, an argument could be made that our ability to pick up clues about other people's mental states from their behavior was an adaptation which arose just as much for the purpose of competition as cooperation. Influence exerted by one person on another may include efforts at intimidation, deception, or other forms of manipulation which are far from sympathetic. Other people can act as "external cognitive resources" against their will, or simply without their knowledge.

One of the most useful elements of the extended mind thesis, I find, is its questioning of the implicit assumption that the individual mind is what does the thinking (active), and the world is what it thinks about (passive). As Wilson puts it, "Meaning is not out there in the world waiting for us to detect it, but created by the mind in its interactions with the world; meaning is intrinsic to the in-the-head mental representation."[108] But we need to go beyond this, since it is not simply a mind/world dichotomy at play. Each person's "world" includes countless *other* minds, each with its own meaning and intentionality. Hence each mind is literally both mind and world, though one of the two is strongly emphasized over the other depending on whether you are considering the matter from the inside or the outside of the mind in question.

What Clark calls "second order cognition," the ability to reflect on one's own thoughts, depends on the thinker being able to externalize her own thoughts from herself, to make them public to

---

[107] Quoted in Jensen, in press

[108] Wilson, "Meaning Making and the Mind of the Externalist," in *The Extended Mind*, ed. Richard Menary, in press.

some extent in order to reflect on them. When we are unable to do this effectively, the next rational step is to find someone else and have them join you in examining your thoughts. Brain-storming, advice seeking, and psychological counseling are all examples of situations in which an individual actively involves another mind (with its own intentionality) into the process of cognition. This "thinking about thinking" cannot be neatly divided into a mind/world dichotomy. Discussion for the purpose of increasing general understanding, as opposed to debate or one-sided attempts at manipulation, requires both empathy in the sense of being able and willing to see another's point of view, and a perception of value in involving another mind in order to expedite one's own knowledge acquisition process.

Having demolished, or at least radically modified, the mental house to include artifacts in the world, I would say that Clark and Wilson should go the next step and recognize the degree to which humans are themselves artifacts—of evolution, of accumulated cultural knowledge, of their own self-reflective thoughts—so as follow the causal chain rather than stopping with material references as if they had sprung out of nowhere, independent of human involvement as opposed to the direct result of human action. The problem with this approach, however, is that it threatens to muddy even further the distinction between individual and shared intentionality. If the mind of each living human today exists to make its way within the environment of culture—that is, the environment of other minds—to what extent can we say that individual intentionality really exists? It is our default state to think of ourselves as autonomous agents, doing things for our own reasons rather than someone else's. Additionally, concern might be raised for the question of how to properly study cognition if it is not longer restricted to the individual mind, or even the individual plus the artifacts with which he or she interacts, but rather extended to include scaffolding that takes the form of other people. But this, it would appear, is unavoidable. Meaning itself is a cumulative social product, rendering the decision to stop our concept of cognition at the gates of other people's house to be itself an arbitrary one.

Wilson rightly points out that the problem of intentionality is exacerbated once meaning-making is not considered to be "in the head." But what of communal meaning-making? The practices of creating and telling myths and performing rituals derive their meaning entirely from the fact that they are social. A myth is an authoritative narrative that gives an explicit interpretation of how to view the

world—quite often, this means the social world. The effectiveness of ritualistic actions depends on the roles performed by participants, just as much as on the sacralized physical artifacts which they use. The emergence of meaning in this context depends upon humans acting as scaffolding for each other. Without using the language of cognitive science, the sociologist Peter L. Berger pointed this out eloquently in describing religious belief and behavior as a "sacred canopy," a set of externalized symbols and meanings which are themselves social products, and are subsequently given the appearance of objectivity and internalized by participants.

Can it be useful to speak of group cognition in religion? Indeed, it has already been done to a certain extent by evolutionary biologist David Sloan Wilson, as part of his greater thesis that it is meaningful to speak of group evolution in the sense that religious belief and behavior is functionally adaptive—that is, it contributes to the survival of its members in some sense. For Wilson, the religious group is metaphorically an organism in itself, competing with other religious groups for survival in an environment of human social interaction. Though the question of whether religion itself is an adaptation is of course controversial, Wilson's portrayal of religious tradition as heritable (though not in the Larmarckian sense, but rather in terms of knowledge passed down through generations) in order to examine how different beliefs and behaviors are "naturally selected" by social forces could be very useful in light of the extended mind thesis. Just as individual humans both cooperate and compete against each other for survival, religious groups do so in a very real and analyzable sense as well. Envisioning group cognition in this way may well aid understanding of how individuals function within a religious community, as well as why communities often clash and find it difficult to see eye to eye.

6. Folk fears about responsibility, part 2: Victims and perpetrators

The social psychologist Roy Baumeister attempted to identify the psychological mechanisms at work when people who commit evil, and people who are the victims of evil, describe the causal factors at play in the act. To examine the validity of a common claim that violence and cruelty are born of low self-esteem, he conducted a meta-

analysis of numerous perpetrator and victim reports from violent crimes. Contrary to the prevailing idea that perpetrators of violence have low self-esteem, Baumeister found that they were more likely to have artificially *high* self-esteem. The qualifier "artificially" is important here because a person whose image of himself is high because he has *earned* special status in some area is generally recognized as such by others. When this is not the case, however, a person's high opinion of himself is likely to be challenged on a frequent basis, causing inner conflict and the need to prove himself by retaliating violently against the challenger. It is often the nature of violence to escalate—an accidental shove is repaid with a harsh word, which is repaid with yet harsher words, which are repaid with blows and possibly death, both parties in the altercation viewing themselves as having been attacked by the other, and therefore merely defending themselves or their honour. Studies have found that people who live in so-called "cultures of honour," such as the American South and other places with a historical background of occupations such as herding, in which a person's resources can be stolen from them in full, tend to resort to violence at much smaller infractions than in cultures with agricultural backgrounds. This is because in such cultures, a man's reputation is integral to his economic survival.[109] Failing to respond to challenges with the "appropriate" force can make him look weak to others, lowering his status in the group and making it more likely for him to be a target.

Evil is the creation of the victim, or more specifically of those who claim victim status. If there were no victims, then in a very real sense there would be no evil, because the perpetrators of evil deeds almost never view their behaviour in those terms. Baumeister has surveyed massive amounts of documentation of crimes, interviews, and theoretical work to examine that very issue, while making the controversial move of not simply taking the accounts of victims at face value.[110] In so doing, he discovered some novel and notable things about the psychology of people on both sides. Perpetrators tended to emphasize their circumstances—even while acknowledging guilt, they would talk about the social situation, factors in their background, coercive or provocative elements, etc. while victims mainly ignored these things. Victims tended to emphasize the perpetrator's responsibility, saying things like "There was no reason

[109] Nisbett & Cohen 1996

[110] Baumeister 1997

for him to do that." Victims, and those who sympathize with them, are often unwilling or unable to acknowledge that someone who hurt them had understandable reasons, and for others to suggest as much would amount to sympathizing with the enemy. Instead, they commonly rely upon what Baumeister refers to as the Myth of Pure Evil. This myth holds the following:

1. Evil is the intentional infliction of harm on people.
2. Evil is driven primarily by the wish to inflict harm merely for the pleasure of doing so. Evil is either gratuitous or simply senseless.
3. The victim is innocent and good—minding her own business, going about her day and not bothering anyone else. The evil one bears all of the blame.
4. Evil is the other, the enemy, the outsider.
5. Evil has been that way since time immemorial.[111]

He draws attention to the difference in the way victims and perpetrators characterize the *cause* of the act—perpetrators are much more likely to describe environmental circumstances, such as being tired or angry because of some previous event, the behaviour of the victim, or their physical surroundings. The victim, by contrast, is more likely to specifically address the perpetrator's mentality, saying things like "He did this for no good reason," or "He did this just to hurt me." To suggest that anything about the victim's behaviour had something to do with the act of evil amounts to blaming the victim. As Baumeister puts it, "It is more comforting to think that the world contains evil, malicious people who attack innocent victims for no reason than to believe that one's sufferings are the result of one's own poor judgment and ill-advised actions that provoked a violent response from someone else."[112] In addition, when people have a limited capacity for helping, they are much more likely to give priority to innocent victims than people who appear to have brought their unfortunate situation on themselves. Note that in this account, the victim is removing the evil act from its social context—outside circumstances are not allowed to be a factor. Evil must be intentional so as to hold the perpetrator fully responsible. Now the victim has a legitimate grievance against the perpetrator, and uses that as grounds to strike back. In striking back, the victim then emphasizes context—the former perpetrator's actions-- and we have a reversal of roles. It's

[111] Ibid., 73-74

[112] Baumeister 1997: 90

perfectly natural, then, for perpetrators throughout history to have viewed themselves as victims—whether of the specific person that they are attacking, or from members of that person's group against whom the victim is striking back generally. This out-group may be members of a different class, race, gender, sexual orientation, religion, political faction, and so on. Over a period of time, and in times of war and other mass movements of aggression, the humanity of those being attacked must be minimized, a practice which is performed especially effectively when the movement is impassioned with idealism. People cease to be ends unto themselves, and become instead means to an end.

So we have a victim with an interest in making him or herself look good, and the perpetrator look bad—and not only bad, but evil. This is achieved by divorcing the perpetrator from his surroundings (including his own body—hormones and genes are not allowed take the fall), and assigning him full intentionality and hence responsibility. After all, a person must be capable of good, at least in theory, in order to be blamed for evil. Everyone has their own reasons for acting the way they do, which is not to say that those reasons are always rational, but violence committed for no reason, or out of sheer sadism, is very rare in the world. But you wouldn't know it from the stories told about *that group over there* when you're a member of *this group right here*. People tend to automatically and inevitably think of their own group as good, and if we are the good guys, that leaves only one alternative for the other guys.

This model of evil fits with the description of the supernatural I gave earlier, in that it represents a causal mind which is independent of and precedes matter—in this case, the behaviour of the body. The ascription of evil is made possible by the dualistic portrayal of the perpetrator's intentionality, and thereby his or her moral responsibility. A reductionistic approach seems to compromise our ability to properly hold a person responsible for their own actions—why is the *person* responsible, rather than his genes, or neurons, or environment? Why shouldn't we just let a criminal go if we can find that he has a brain abnormality which in many people seems to cause aggression?[113] The prognosis of evil allows us to stop questioning, allows the buck of moral responsibility to stop in one place, with one person. Evil allows us, as Baumeister says, to distinguish ourselves from the perpetrator as well as the perpetrator

[113] http://www.sciammind.com/article.cfm?articleID=3124278A-E7F2-99DF-37D9F7F0D409C398

from any exculpatory circumstances. It acts as an effective blocker against empathy, which can undermine proper condemnation. It might seem odd to consider the *inhibition* of empathy adaptive since so much work has been done lately to think about what adaptive functions it would serve, but in some circumstances it may well be more expedient to forming quick, decisive conclusions about other people's behaviour in order to decide whether they should be punished or ostracized, or simply not trusted in the future. The primary elements which have been shown to stimulate empathic reactions are perceptions of similarity and familiarity. If these perceptions can be diminished or eliminated by emphasizing that *he* is responsible for the bad thing whereas *I* am not, a hasty barrier can be constructed which cuts off any thoughts of identification with him which might lead to an unwillingness to condemn or punish.

7. **On moral levitation**

"Moral levitation" is a term used by philosopher Daniel Dennett in his book *Freedom Evolves*. He defines the idea of moral levitation as, in effect, saying the following:

> If your feet are on the ground, the decision isn't really yours—it's really the planet Earth's decision. The decision isn't *made by you* but is rather a mere summation of causal trains intersecting in your body, a mobile bump on the surface of the planet, buffeted by influences, answerable to gravity. Real autonomy, real freedom, requires that the chooser be somehow suspended, isolated from the push and pull of all those causes, so that when decisions are made nothing causes them except *you*![114]

The moral levitator, Dennett argues, is the inevitable result of a particular kind of thinking regarding free will: agent causation. People can only truly be held responsible for their actions, this view holds, if they act as a "prime mover unmoved," according to proponent Roderick Chisolm.[115] A person who is a libertarian with regard to free will (which is not to be confused with the political term) believes that in order for true responsibility to exist, determinism must

[114] Dennett 2003: 102

[115] Quoted in Dennett 2003: 100

be false, as discussed earlier in this chapter. A dualist, however, is a kind of libertarian who can take things one step further and maintain that determinism is false because the part of us which is responsible is wholly separate from material existence. This separation is what makes moral praise and blame possible, since otherwise we risk ascribing credit to the wrong sources.

Dennett's book is a defence of a compatibilist view of free will—that the kind of freedom which matters, and (luckily) the one we have, is not threatened by determinism. The fact that so many people automatically conclude otherwise is telling. There are multiple ways to be a non-compatibilist. A person can maintain that free will and determinism are not compatible, and hence free will does not exist, or if they are a libertarian they can maintain that free will and determinism are not compatible and therefore determinism is necessarily false, at least when it comes to matters relating to human responsibility. But the depiction of free will which seems to predominate most in our folk psychology is the kind that entails moral levitation, as I have tried to show in this chapter.

This portrayal of responsibility is perceived to be endangered by evolution, as previously mentioned. Evolution entails that we are in a very real sense connected to everything else on earth—we are made using the same "recipe" for life, DNA, because we all come from a common ancestor. Through our evolution we have retained many aspects which we still have in common with other animals, which in the eyes of some anti-evolutionists leads to the conclusion that "if you teach people that they're descended from animals, they will act like animals." When biologist Randy Thornhill and anthropologist Craig Palmer came out with their book *A Natural History of Rape: Biological Bases of Sexual Coercion* in 2000, the idea that rape could be viewed as evolutionarily adaptive was greeted with a great deal of controversy. The conclusions that evolutionary psychology reaches about why humans behave the way they do are frightening to both moral levitationists and social constructivists because they imply inevitability, whereas social constructivists want to maintain that behaviour is shaped mainly by environment. Moral levitationists, on the other hand, want to maintain that behaviour is determined by sheer will. There is nothing, either in society or in one's genes that could compel a person to rape— or steal, or murder, or anything else immoral unless the individual consciously decides to do so. The refusal to countenance the idea that there is a genetic component to homosexuality well may stem from this sort of thinking.

A genetic basis for a behaviour implies that the shining locus of free will is blemished, compromised, allowing for doubt to enter in concerning whether it is appropriate to blame homosexuals for their desires and behaviours.

Chisolm acknowledges the supernatural aspect to this portrayal of human responsibility, that it entails a "prerogative which some would attribute only to God."[116] Baumeister's depiction of evil as a quality which divorces perpetrators from outside (or inside) causal influences meshes well with a moral levitationist notion of responsibility. In order to pin blame on someone, and pin it solidly enough that you can brand them as evil, it takes a strong intuition that humans have essential natures which levitate, separate from the body, separate from all material existence. I would argue that this kind of move is entirely intuitive, especially when we consider the responsibility of those who strike us as highly immoral.

8. **Conclusion**

A recent article in Newsweek magazine, describing American 2008 presidential candidate John McCain, said of him, "He believes deeply that America is a force for good, and he recognizes that there are certain people in the world who send their children off to be suicide bombers or repress their citizens viciously whom you can't use any word other than evil to describe."[117] At a time in history when America struggles in its relationship with the people of Iraq, Iran, Afghanistan, and Pakistan, these are interesting terms to use. What does one do with people one can't describe with any other word than evil? What does it mean to label one's entire nation as good (or even one's entire government), in opposition to the evil of these people? This sort of comment is an illustration of good and evil being ascribed as immutable characteristics, pitting the good ("us") and against the evil ("them") even though McCain is characterized as specifically using the term concerning people who commit specific acts as evil as opposed to entire countries. This method of labelling certainly makes the battleground simpler, the goals more clear-cut, but it also serves to divorce of agents from both their human nature and their social

[116] Ibid., 100

[117] Kushner 2008: 64

situation, and cut off triggers of empathy which might otherwise have prevented strong condemnation and promoted humanization.

In order to characterize a person or people as evil, we must follow certain steps, as described in this chapter. We must:

1) Emphasize their causal role in creating suffering for others, which requires that we
2) Over-attribute their intentionality in such acts so that we can effectively blame them, thereby
3) Identifying a victim (which will likely be us) and thus
4) Separate their agency from worldly influences and isolate it as evil.

This is the process by which evil is made supernatural, isolating agency from the body and society in order to emphasize responsibility by portraying responsibility using moral levitation. I have attempted to demonstrate in this chapter both how intuitive this move is, but also offer some explanation as to why it is not necessary, and even counter-productive.

Chapter 4

Disgust and Divinity: How the vulnerable body creates the soul

Near the end of his book *The Symbolic Species*, Terrence Deacon suggests that our evolved ability to use symbols has carried with it the symbolization of the self itself, creating an inevitable discomfort and loathing of the cessation of this symbol. "Knowledge of death, of the inconceivable possibility that the experiences of life will end, is a datum that only symbolic representations can impart." (Deacon p. 436) Later, he extrapolates that this way of thinking leads to the intuition of a distinction between the self and the body as a way of coping. In this chapter I am going to investigate a particular angle on how this might take place, presenting a sketch of the ways in which symbols of our biological existence—that is, our animal nature and mortality—may function in our moral psychology in a way that is particularly relevant for scholars of religion.[118]

The title of the chapter suggests that I am going to focus on the body—that is partially true. The main purpose is to describe the ramifications of a particular *kind* of focusing on the body—the kind that creates symbols of things which portray the body as vulnerable, and therefore suggest an escape route, something independent of the body which cannot be tarnished, that thing being the soul. The particular emotion I am going to focus on is disgust, and how it manages to perform this neat trick. Disgust is an evolved response that we have toward certain things in the world that are bad for us. It can be distinguished from fear as not all disgusting things are

[118] Portions of this chapter have been taken from Koch 2008a

dangerous, nor are all dangerous things disgusting. This chapter discusses the ways in which we are capable of recruiting this particular response to carve out dividing lines between the body and the self, and further between the self and other selves.

1. **Disgust as an intuition**

Disgust represents rejection in its most vehement sense; I call it an "otherizing" emotion. It is one of the core emotions—other animals have a form of it as distaste, a valuable trait to have in a world where bitter things may be poisonous. This is what psychologist Paul Rozin calls "core disgust," the reaction to ingestion or possible ingestion of something that seems plainly foul and may be actually dangerous. From the extension of this comes a more social form of the emotion, which is called "animal reminder" disgust because the things which trigger it appear to be things which suggest that we, like the rest of the natural world, get sick, decay, and die. So even things which do not represent a direct danger, such as any bodily fluid independent from the body (except tears, possibly because they appear to be unique to humans) or the sight of someone with an obvious physical deformity, or even the thought of certain sexual acts, provoke a similar reaction to what you'd get if you'd just eaten a terrible-tasting food. This kind of disgust can be taught by using terms which elicit a disgusted reaction metaphorically, which then create an association in the mind between the visceral reaction and a new subject, as will be discussed further later. In those who experience it, the neurological response is strongly similar to that for core disgust—a recent study by Moll et al. using fMRI has revealed that there is considerable overlap in the brain between disgust and moral indignation.[119]

The psychologist Marc Hauser noted that if empathy is the emotion that brings people together, then disgust is what drives them apart. This seems to be a critical element in our examination of disgust and how to distinguish it from other emotions such as anger, contempt, or shame. Shame is an indicator of something wrong within the self, whereas disgust indicates something wrong with the other. It indicates a need for distance, of risk for contagion or

[119] Moll et al. 2005

contamination. Disgust, interestingly, has the two features of sympathetic magic originally described by George Frazer in the *The Golden Bough*: contagion and similarity. The law of contagion says that contact makes a particular item suspect—test subjects would refuse to drink a beverage which had been swashed with a sterilized cockroach, and said that they would refuse to wear a sweater which had been worn by an immoral person. The law of similarity says that resemblance should raise suspicion—subjects refused to eat chocolate fudge which resembles dog feces.[120]

The word "disgust" comes from the Latin for "distaste," and Rozin identifies as "core disgust" those things which elicit such visceral reactions as described above. But beyond this, he describes a realm of social disgust in which this same reaction is expanded to include members of certain groups and people who practice certain behaviours. These people are viewed as animalistic, base, subhuman. In human history violent in-groups that have come to power have often had a puritan bent, and portrayed members of their out-groups as filthy animals. Hitler described the Jews as maggots, hiding inside a cavity in the body of society and cringing from the light when suddenly exposed, while maintaining the highest of standards for the S.S. police force. The Ku Klux Klan in the United States stressed sexual purity, and woman campaigned against liquor and vice,[121] while the group described African-Americans as "apes," "dogs," and "monkeys."[122] One salient recent example is Colorado's Amendment 2 to the state constitution in 1996, which denied local communities the right to make nondiscrimination laws on the ground of sexual orientation.[123] Proponents of the amendment (that is, opponents of nondiscrimination laws for sexual orientation) embarked on a PR campaign preceding the statewide vote in which they circulated pamphlets describing gay men as being fond of eating feces and drinking blood.[124]

Further, inducing the emotion of disgust appears to carry a distinct effect which causes people to make moral judgments which they otherwise would not. The experimental philosopher Shaun

---

[120] Rozin et al. 2000

[121] Baumeister, 1997

[122] Billig, 2001

[123] Ultimately declared unconstitutional by the U.S. Supreme Court in *Romer v. Evans*, 116 S. Ct. 1620 (1996)

[124] Nussbaum 2004: 101

Nichols found that violations of norms which are disgusting (what he calls "norms with feeling") evoked judgments which were much more harsh than violations of social convention. Testing subjects on their response to scenarios which involved moral infractions (such as hitting someone or pulling their hair), conventional infractions (such as wearing pajamas to school or drinking soup out of one's bowl at a party) and disgusting acts (such as picking one's nose or spitting in a glass before drinking it), he found that subjects were actually prone to judging the disgusting violations as being very similar to the moral ones in terms of permissibility, seriousness and authority contingencies (that is, whether the act would be okay if the appropriate authority says it is).[125] What's more, it appears that disgust can *emerge* according to the decision that a particular behavior is immoral, such as the disgust some vegetarians develop for meat-eating.[126] Nichols' conclusion was therefore that disgust has a uniquely moralizing effect—it turns violations of conventional rules into moral infractions, which are considered quite a bit more serious. Psychologist Frederik Björklund found that people with high disgust sensitivity demonstrated an increased tendency to judge an act as immoral when primed with what he describes as "morally irrelevant disgust."

Social psychologist Jonathan Haidt and colleagues (including Björklund) have performed a number of experiments using disgust to examine its relationship to moral judgment. In one, they hypnotized subjects to react with disgust upon hearing the word "often." They then presented the subjects with vignettes containing the word "often" or a synonym of it, some of which described morally reprehensible characters and some which were morally admirable. Amazingly, subjects hypnotized to feel disgust when hearing the word "often" would judge the morally admirable character as morally wrong when the word came up in the vignettes.[127] In an experiment which identified the phenomenon to which Haidt refers as "moral dumbfounding," he and colleagues presented subjects with a scenario of brother-sister incest in which testers systematically negated every objection offered by the subjects—i.e., if the subjects objected to possible birth of a child with defects, the tester would respond that birth control was used. When left with no such arguments to make,

[125] Nichols 2002

[126] Rozin et al. 1997

[127] Wheatley and Haidt, forthcoming.

some subjects would conclude that the act was permissible. But the majority continued to maintain that it was not, even when left with a response that was essentially "I don't know, I can't explain it, it's just *wrong*!"[128]

Here is a moral quandary for you to consider:

> *Frank's dog was killed by a car in front of his house. So he cut up the body and cooked it and ate it for dinner. How wrong was that?*

Imagine that while you are considering this scenario, you're sitting at a nice, clean desk. Then imagine that instead of the clean desk, you're considering the morality of this question while sitting at a very messy desk. There's a crusty drink cup in front of you, along with a chewed-up pencil, a used tissue, and a greasy pizza box. Do you think your evaluation might come out differently in the latter situation? If you're like the subjects presented with the question by Haidt and colleagues,[129] you're more likely to judge Frank eating his dog as immoral when sitting at the messy desk. And what's more, you probably wouldn't be doing so consciously. Nevertheless, provoking a response of disgust in you to something you already may find disgusting in the first place—the scenario itself—may well cause you to be harsher in your moral judgment.

There are different ways to interpret these kinds of results. Haidt's own conclusion is that intuitions are the primary basis of our moral judgments, and reasoning really only comes in *post hoc*, when we're trying to convince others of a conclusion we've already reached through other means. This is his explanation of why people attempting to debate a moral issue often feel as if they've hit a brick wall. After all, if we share the same moral values, then why isn't the other person persuaded by my argument? The inference is that they must either be a) evil or b) stupid. Haidt argues that the intuitive domains from which our values stem are partially evolved and partially enculturated—the enculturation explains why some of us may be more concerned with avoiding suffering and emphasizing reciprocity (American liberals), while others may be more concerned with respect for hierarchy, or concerns about purity and pollution (American conservatives).

---

[128] Murphy, Haidt, & Björklund, 2004

[129] Haidt, Schnall, & Clore, forthcoming

This conclusion is derived from what Haidt, McCauley, and Rozin call the "The CAD Triad Hypothesis."[130] CAD stands, in part, for the "big three" moral codes—community, autonomy, and divinity. They were originally described as such by anthropologist Richard Shweder et al,[131] as an attempt to identify the primary types of moral causal ontology in different societies. This ontology in folk psychology, explained Shweder at al, is different from what might be produced by a logician, such as John Stuart Mill, or an empiricist, such as David Hume. Rather, causation in folk psychology is highly concerned with maintaining normality, assigning blame, and controlling future events. To this end, a folk psychologist (that is, a human being functioning in non-reflective, "common sense" mode) will not be satisfied with a dead end concerning the causation of an event with moral ramifications—a cause *must* be found, in order to resolve the matter and relieve the morally condemning emotion. They suggest that "a transgression (for example, disregard for one's parents, incest) is most readily moralized if it is imbedded in an intellectual framework which carries with it the implication that the transgressor has violated the sacred order of things, as manifest in nature, society, or the self."[132]

Haidt et al. took this idea and from it developed a corresponding set of moral emotions, which align to the CAD triad and, conveniently, share the same acronym—contempt, anger, and disgust. According to their mapping, violations of the community ethic arouse the emotion of contempt, violations of autonomy arouse anger, and violations of divinity arouse disgust. Though they also acknowledge Shweder's attention to what he calls "other-suffering" emotions, such as sympathy and empathy—those with which Hume and Adam Smith were concerned—they do not articulate a corresponding triad for them. There may not be one which so neatly aligns as the one they propose, but the relationship between disgust and empathy—or rather; inhibition of empathy—is a focus of continued examination in this paper. Though, following Shweder, they identify contempt, anger, and disgust as having in common the character of arousing disapprobation in others, the defining characteristic of a moral emotion according to Hume, the corresponding emotions of approbation are not addressed within this

[130] Rozin, Lowery, Imada, and Haidt 1999

[131] Shweder et al, 1997

[132] Ibid., 127

consideration. Haidt considers them elsewhere on his own, a treatment I will be addressing in a later chapter.

Haidt et al. trace anger and disgust back to their animal precursors. Numerous studies have been performed on the two emotions in non-human animals, but not necessarily as *moral* emotions. Anger is mainly a response to frustration or goal blockage as a means of rousing the energy and resources necessary to overcome an obstacle. Disgust as a precursor, on the other hand, is more accurately described as distaste, a reaction that correlates to what Rozin calls "core disgust"—the avoidance of that which may enter the mouth and contaminate it. The elaboration of this emotion into the social domain is called *sociomoral disgust*, a reaction which can be found in cultures the world over, but with varying triggers—sometimes actions performed by a third party, such as when disgust aroused in an uninvolved observer. The word for disgust, Haidt et al. report, exists in this heavily moralized usage in many languages, including French, German, Hebrew, Russian, Japanese, and Chinese.[133] Sociomoral disgust, they argue, invokes the moral domain of divinity. It applies when "to decide if an action is wrong, you think about things like sin, the natural order of things, sanctity, and the protection of the soul or the world from degradation and spiritual defilement."[134]

On the assumption that moral systems may vary between countries, Rozin et al. tested subjects from both Japan and the U.S. by giving them a list of moral violations based on Shweder's moral codes and asking them to say which was the most appropriate word to describe their feelings—contempt, anger, or disgust—or to choose the most appropriate from a series of photos depicting people displaying those emotions. They found that in both cultures, there was support for the hypothesis of a correlation between the moral emotions chosen and Shweder's codes, with somewhat of an exception lying in the emotion of contempt. The concept turned out to be somewhat nebulous lexically. The study reports that "the mapping from the contempt face to community is much more substantial than the mapping to the word in both cultures."[135] They also found that sometimes people placed moral violations in different domains than had previously been expected. For example, pushing ahead of

---

[133] Haidt et al, 1997

[134] Rozin et al 1999: 576

[135] Ibid,. 581

someone in line was in both cultures classified as more of a violation of community than of autonomy. More interestingly, they also found that "the divinity code is so foreign to the moral system used by educated Americans that most of these participants assigned what we took to be such violations to the nonmoral category."[136] Unfortunately, the study only tested subjects on five different scenarios which were judged to be in the category of violations of divinity, as compared to twelve for community and ten for autonomy, and no mention is made of which violations were judged by educated Americans not to be moral violations at all. Of the five, three appear to constitute core disgust: eating a piece of rotten meat, touching a corpse, and watching someone bite into a wormy apple. The other two are more clearly sociomoral: shaking hands with someone who is in an incestuous relationship, and hearing about a 70 year old male who has sex with a 17 year old female.

Rozin refers to disgust as the "body and soul" emotion, because according to their research, the things which most often evoke disgust are those which stand as unpleasant reminders of our own animal nature and mortality. Corpses, faeces, and creatures we associate with the above such as rats and cockroaches. Bodily fluids, once they are separated from the body. "Violations of the body envelope," such as injuries and surgery, which remind us that we are physical beings, just like the animals, and our lives are also messy.

2. Playing Frankenstein

Shweder's moral domain of divinity, as articulated by Rozin et al., is connected to the disgust response. Earlier I mentioned some of the triggers which seem to be more or less primary in their disgust-eliciting nature, such as injuries, fecal matter, decay of flesh, etc. Associating these qualities with human beings can elicit moral condemnation of those people by those we wish to be our friends— if you join with me in revulsion of him, then we appear to be comfortably higher. We can affirm our human dignity. We can, in a very real sense, remove ourselves from materiality. It is in this sense that disgust makes us "supernatural"—not in the sense of violating

[136] Ibid., 581

intuitive expectations, but actually in an apparently *intuitive* way. Haidt has called the emotion "the guardian of the temple of the body,"[137] the "paradigmatic emotion of spiritual pollution"[138] and in his recent book *The Happiness Hypothesis*, he contrasts disgust with what he calls elevation (an attempt at transposing divinity into something a bit more secular) and argues that the potential for human flourishing is inhibited when we lack this dimension on our lives.

There are those who do earnestly, consciously, believe that disgust is a reliable predictor of behaviors which are actually immoral, or at least that such a reaction is a justification for taking action on behalf of the welfare of society. One of its most famous defenders of the latter view was Lord Devlin, a British judge, who wrote *The Enforcement of Morals* in 1965. Devlin argued that the disgust of the common person, whom he described as "the man on the Clapham omnibus,"[139] constitutes a strong reason to make a behavior illegal even if it doesn't actually harm anyone. The idea was that the disgust of the common man is an indicator of a hazard to society, and therefore measures taken to prevent this emotion are justified in the interest of preserving society.

There was quite a lot of outcry and furor when the moral philosopher Peter Singer, seeking to make a case for consistency in our treatment of animals, published an essay in Nerve Magazine in 2001 entitled "Heavy Petting" arguing that sex with animals is not necessarily immoral, especially considering our willingness to put them through great hardship on their way to being killed and eaten. While he sought to make vivid some of our hypocritical attitudes regarding humane treatment toward selected species, the reaction was of incredible disgust that he could speak of such a subject without condemnation.[140]

Former head of the (U.S.) President's Council on Bioethics, Leon Kass, wrote an essay, first published in The New Republic in 1997, and entitled "The Wisdom of Repugnance," claiming that disgust serves as a deep moral guide within us, alerting us to acts which are immoral by means of sheer revulsion.[141] He was speaking

---

[137] Haidt, 1997

[138] Haidt 2003

[139] Nussbaum 2004: 4

[140] See: http://www.dailyprincetonian.com/archives/2001/03/08/opinion/2591.shtml, http://www.animalrights.net/archives/year/2001/000040.html, http://www.all-creatures.org/sof/rolemodel-singer.html

[141] Kass 1997

in that instance on the subject of human cloning, for which he endorsed a worldwide ban. The idea of creating a clone whose organs might be substituted for those which have failed, he explains in the essay, is an example of "Frankensteinian hubris" and "man playing God." "We are repelled by the prospect of cloning human beings," Kass writes, "not because of the strangeness or novelty of the undertaking, but because we intuit and feel, immediately and without argument, the violation of things that we rightfully hold dear."[142] As support for this thesis, Kass notes that even the creator of Dolly the famous cloned sheep, Ian Wilmut, has said that he "would find it offensive" to clone a human being. Compare this to the words of Louise Brown, the world's first baby to have been conceived by in vitro fertilization, on how she would feel if she had been the product of a donated egg: "I would be disgusted."[143] Kass has come out against the donating of eggs, on the grounds that it "represents another large step into turning procreation into manufacturing. It's the dehumanization of procreation."[144]

Clearly Kass and Wilmut, and presumably Brown, are not opposed to science generally. Nor are they likely opposed to efforts in medicine and biology to help humans to be healthier and live longer. But each of them has a point at which applying such an approach to the human reproductive system becomes repellant, disturbing. Why? Kass's language is revealing. The claim that we are "playing God" has come up again and again in the history of medical advancements, including in 1954 when the first kidney transplant was performed—by Dr. Joseph Murray, who eventually won a Nobel prize for his efforts.[145] No doubt that to those who made this allegation, the idea of moving an organ from one body to another seemed gruesome and "Frankensteinian" at the time. Today, of course, we think nothing of it. When a South African doctor performed the first heart transplant thirteen years later, similar refrains resounded—the human body will be treated like an old car, with parts replaced at will. The comparison continues today, using the same terminology: "Just like we have car body shops today, we'll have human body shops in the future where we go to replace any worn-out parts," said Dr. Michio Kaku, host of a

---

[142] Ibid.

[143] Howley 2006

[144] Lopez 1998

[145] Wighton 2007

BBC Four series on medicine and technology, in an article last year's Times Online. "I went into the world's leading centre in regenerative medicine, in North Carolina. It was a little like going into Frankenstein's laboratory; there were body tissues – noses, ears and bladders – growing on the laboratory benches."[146]

The human body reduced to a machine, man playing God, man playing Frankenstein—the language repeats itself in a new cycle with each new innovation. Mary Shelley's Dr. Frankenstein, though he had built his monster with the parts of dead bodies retrieved from local graveyards, was horrified nevertheless when his project came to fruition and the monster became alive. Perhaps the idea is that even if we don't yet see the horror of a particular suggested innovation, we will do so, like Frankenstein, only when it is too late. But given the number of times this rhetoric has been invoked, it's hard to take Kass completely at his word when he insists that the strangeness and novelty of human cloning really has nothing to do with his objection to it. Who can say but that not too far in the future, therapeutic cloning may become as much of a non-issue as organ transplantation? That is, unless Kass and like-minded others succeed in getting it banned, as is his goal. Compare the images of being Frankenstein and playing God—why should playing God be disgusting, repellant? Divinity should be the opposite of disgust in this model, but the idea is that we are, after all, merely *playing,* pretending. Like Frankenstein, we have stepped over some invisible boundary, unintentionally defiling what we find there. We have tried on God's sacred vestments, but they are ill-fitting. So, like the magician's apprentice, we stumble around, mucking up the procedure through our hubris. The revulsion lies in our brazen grasping for God's toolkit. This is what Kass intends to warn us against. In a later interview, he complained at being viewed as merely an old-fashioned moralist: "In the biomedical area, the people who are bringing you all of the novelties occupy the moral high ground. They are humanitarians. They are interested in curing diseases, ending suffering, extending life. If anybody says 'Let's go slow here,' it looks like the imposition of a narrow religious view on what is a pluralistic society, and the response is 'Get your morals off my science.'"[147] But there is a difference between saying "This new technology might bring these detrimental results, so let's be cautious" and simply invoking the

[146] Wighton 2007

[147] Wilkinson, Francis 2008: 62

"wisdom of repugnance" to suggest that there is something questionable about the technology because it inspires feelings of inner discomfort. Objecting to an innovation because it is relatively new and untested, one might argue, is actually *more* tenable a justification than pointing to feelings of repugnance and invoking nebulous concepts of human dignity. Often innovations which contribute significantly toward lessening human suffering and extending our lives aren't pretty—heart transplants, for example, are not something most people would like to watch during lunchtime. But Kass's standard of relying on inner revulsion as a moral indicator does not offer a reliable guide by which to determine whether such a procedure is in fact moral or immoral.

On an edition of BBC World News Today recently,[148] presenter Philippa Thomas interviewed bioethicist Arthur Caplan on the ethics of inserting human DNA into a cow's egg in order to create hybrid embryos. Stem cells would then be drawn from this embryo and studied for the purpose of learning about the development of dehabilitating diseases such as Alzheimer's and Parkinson's. Earlier the broadcast had depicted protests taking place in London in November to protest the permission having been given to two British universities to do such research. Josephine Quintavalle of CORE (Comment On Reproductive Ethics) made the following comment: "We don't agree with the fertility regulator in principle because we don't think that there is ever any justification for breaking this great moral taboo of combining human and animal gametes in this way. We're not talking about adding an animal cell to a human or vice versa. We're talking about actually creating an embryo."[149] Questioned on the ramifications of the procedure, Caplan acknowledged that "some people might find it distasteful or even taboo." However, he noted that though an embryo which was part human and part cow (approximately 99.9% of the former and 0.1% of the latter) would in fact be created, there is no intention to create an actual chimera, or to "take it anywhere out of the dish." Therefore, he said, we are not "crossing a inviolate line, even though we're talking about a human and animal mixture." He then noted that perhaps legislation should be paused in the future to prevent anything involving cross-species manipulation from "leaving the dish." But currently, he said, nothing is being done that is actually "harmful or

[148] January 17, 2008, BBC America

[149] Ibid.

disrespective [sic] to human dignity." The newscaster pushed him further, questioning the possibility of "sliding" from the creation to the implantation of embryos. Upon his reply that the prospect of anything viable could be created in this way was remote, her response was, "So do you think it's okay for us to play God if we're simply not very good at it yet?" Caplan laughed, and noted humans have the ability to manipulate DNA, and that if there is a divine being, "he wouldn't have armed us with these tools lest we were intended to find out the secrets of life."

A look at CORE's website stating their views on cloning reveals them to be highly reminiscent of Kass's arguments. It claims that cloning is "unnatural," that it is a "colossal intervention in the natural order to change from sexual to asexual reproduction."[150] The use of the term "natural" to mean "good" or "normal" is widespread, but in this case it appears to refer to a violation of the essential nature of being human, in the sense of being a crime against nature. It is a crime against nature to aspire to the category of the divine, and it is likewise a crime to disdain, disregard, or alter our status in the category of humanity. This is not merely done by imitating animals, but also by committing acts which bring shame to the species—they make others wish that they did not have to share membership with us. Rozin's observation of the rules of contagion applying to social disgust resurfaces here, when we think about other instances when the rhetoric of "crime against nature" is used for moral condemnation. In many states in the U.S. this is the kind of language that was included in legislation against homosexual sodomy, until such legislation was overturned by *Lawrence v. Texas* in 2003. The term "crime against nature" has been used in American case law since at least 1814, when it was invoked in a case concerning bestiality.[151] It has been variously recruited to indict sodomy (whether homosexual or general), cunnilingus, and necrophilia as well. The sexual behavior of a human appears to have a particular relevance to our essential status in the order of things. Our sexual impropriety is an offense to that order, and that impropriety can be recognized by deviant behavior. And what's more, such sexual impropriety of another is an offense to others by its very existence. We are tainted by it through its contagious properties.

---

[150] http://www.geneticfutures.com/ltltg/info/sheet11.asp Accessed Jan. 19, 2008

[151] Andrews v. Vanduzer, N.Y.Sup. 1814

The belief that a response of disgust indicates that something is immoral can be seen with clarity if one chooses to broach the subject of bestiality. Prominent public figures in the United States such as the late evangelist Jerry Falwell and former Pennsylvania senator Rick Santorum have compared homosexuality to bestiality in order to evoke a response of disgust and thereby moral condemnation. Recently, Republican presidential candidate Mike Huckabee joined their ranks in proclaiming that the desire to legalize gay marriage in the states is akin to legalizing marriage with non-human animals.[152]

The disparity in which things are regarded as disgusting by different cultures appears to point strongly to the utility of the emotion as identifying and ostracizing members of the out-group. Attributes and behaviors which are identified within your particular group can be taught and passed down, so that they can function as a sort of secret code which bonds a group together and makes the in-group/out-group boundary that much more clear and firm-- as well as identifying the out-group clearly as the bad guys, whereas we, of course, are the good.

Philosopher Martha Nussbaum wrote a book entitled *Hiding From Humanity: Disgust, Shame, and the Law*, in which she makes the case that while emotion is not, and should not be, an issue entirely divorced from the law, disgust is not a valid consideration in legal matters in a society which aspires to Millsian liberalism—that is, a society which does not punish individuals who are not directly harming others. But without connecting disgust directly back to religion, she argues that the emotion functions to help ourselves forget, or obscure, our own capacity for evil. But I think the most powerful argument she makes is that disgust pushes its object so far from us as to deny it moral responsibility altogether, and thus to deprive it of its humanity in a significant way. To accuse a person of being evil is at least to affirm their humanity in the sense of suggesting that they know right from wrong and yet chose wrong and are culpable for such a decision. Anger, she says, is an attractive emotion in that it compels active engagement with the object of one's anger. To be disgusted by a person, on the other hand, is to eject them from the human family altogether, and thus render them an immutable "other." To quote Nussbaum:

> "No matter how we define sanity for legal purposes, when we turn someone into a monster we immediately raise the issue of

[152] http://www.beliefnet.com/story/228/story_22873_2.html Accessed Jan. 19th, 2008

> sanity. Aristotle already held that certain individuals . . . were so weird that they were not even vicious because such extreme and bizarre pathology shows that someone is not really a chooser of ends after all.[153] No matter what psychological concepts we use, we have a hard time not getting into similar difficulty when we try to combine a strong ascription of moral responsibility with an a account appealing to disgust at the alleged monstrousness and inhumanity of the person's deeds. . . Disgust, far from shoring up the moral borders of our community, may actually make them harder to police."[154]

Nussbaum argues that the fact that small children are not easily disgusted by vomit, feces and the like means that disgust is taught. Rozin allows that children perhaps react to parental cues to develop full blown disgust reactions. But it's equally possible that disgust is a developed reaction which will emerge regardless of what children see in their parents. Babies, who spend a good deal of their waking life excreting in one way or another without any control over it, would presumably suffer a great deal if the processes disgusted them as much as it does adults. But unless we can find a society in which parents are devoid of disgust reactions in front of their children from the beginning, it will be difficult to test this.

Nussbaum comments briefly on the theoretical approach of anthropologist Mary Douglas in the latter's seminal work *Purity and Danger*, but ultimately discards it in favor of Rozin's (though she does acknowledge that the former approach may be useful for analysis of prohibition and taboo). Nussbaum has a number of reasons for doing this, which I think deserve a closer look to determine whether Douglas really deserves such short shrift. Nussbaum's principle complaint about the usefulness of Douglas' approach for an analysis of disgust is that Douglas tends to conflate her terms—she does not clearly distinguish the categories of the impure, the disgusting, and the dangerous. There is obviously not a direct mapping between any two of the three. Douglas, however, does not see this as a problem in addressing ritual impurity because her explanation focuses on social contextualized symbols—another problem for Nussbaum. Douglas' theory is sociological and culturally specific to the complete exclusion of individual psychology, allowing for no appeal to human nature to

---

[153] Nichomachean Ethics VII.5, 1148b24

[154] Nussbaum ibid., p. 165

account for why certain objects (feces, corpses) appear to be universally disgusting—even if they are not universally ritually impure. Indeed, Douglas even allows that in some cultures, those objects become ritually pure *because* in all other contexts they are considered disgusting. For Douglas, the main concern is anomaly—the fact of not fitting into a given category for one reason or another, which she alleges makes an object stand out as dangerous, impure, or disgusting. But which of these becomes the actual manifestation, for her, is apparently completely culture-specific. Douglas, speaking as an anthropologist, emphasizes cultural variance without any mention of some aspect of human nature contributing to the *content* of what we may find dangerous, disgusting, and/or impure. As Nussbaum points out, anomaly by itself just isn't strong enough to typify or predict any of the above in a given culture. Feces and corpses are disgusting, but not anomalous. A duck-billed platypus is surely anomalous as an egg-laying mammal, but there's no particular reason to suspect that it should therefore be considered disgusting, dangerous, or impure. Given these problems, Nussbaum finds Rozin's notion of animal-reminder disgust a more plausible theoretical grounding.

Douglas may have an emphasis on anomaly that does not address disgust adequately for Nussbaum's needs, but she also does a very interesting treatment on the notion of boundary-crossing—what Rozin and colleagues would call "body envelope violations." When discussing sexual pollution and the boundaries of the body, Douglas' language closely overlaps both Nussbaum and Rozin concerning the perception of the exchange of fluids. "The function of the different body parts and their relation," she writes, "afford a source of symbols for other complex structures. We cannot possibly interpret rituals concerning excreta, breast milk, saliva and the rest unless we are prepared to see in the body a symbol of society and to see the powers and dangers credited to social structure reproduced in small on the human body."[155] For Douglas, the analogy works from the outside in. The society is transposed onto the body, rather than the body onto society. Restrictions on the body, especially those pertaining to female and male sexuality, represent the society attempting to regulate itself via the vehicle of the body.

[155] Douglas, pg. 142

3. Gender and sexual disgust

Nussbaum draws a connection between sexual disgust and misogyny. Following William Miller, she identifies semen as a fluid that men find (once it has left the male body) inherently disgusting, leading to a view of the receivers of semen being inherently debased through contamination. She maintains that men are disturbed by birth and thereby come to associate birth-givers, and the animal continuation of life that they represent, with aversion. Women, according to Nussbaum, are universally but often implicitly identified as the sexual aspect of men. Engaging in sex with women is a man's way of indulging his animal nature, and only women are guilty for this dalliance with animalism because once intercourse is completed, they remain women. They remain in the role of the temptation of animal nature toward man, disrupting male autonomy but only temporarily. They represent the bodily functions, and thereby the body and its frailties generally—including mortality.

A specific example of this manifestation in a religious context can be found in Liz Wilson's analysis of Buddhist hagiography. In her book *Charming Cadavers*, Wilson pours over Pāli, Sanskrit, and Tamil Buddhist texts to look for redactors' use of disgust invoking rhetoric as a device to "evoke a salutary sense of aversion toward carnal pleasures"[156]—and later, to women themselves. Wilson argues that graphically descriptive passages of gruesome scenes involving corpses, especially female corpses, were used to cultivate a sense of impermanence concerning the body and its relationship to *samsāra*, mundane material existence. Contemplation of the female corpse as a practice contributed to the sustaining of the *sangha*, the renouncing community of monks and nuns, in their attempts to avoid the folly of lust and its resulting creation of attachments to the world—babies, spouses, families. In particular, two aspects of the hagiography described by Wilson jump out as illuminating: scenes from the story of the Buddha's awakening to suffering and subsequent efforts toward enlightenment, and tasks assigned by him to future monks and nuns concerning their mental training.

One very poignant example of this is the story of Sirimā. Sirimā is a former courtesan who converts to following the teachings of the Buddha. She becomes a patron of the *sangha*, and spends a good deal of money on food for the monks, organizing a fine meal for

[156] Wilson (1996), pg. 11

them at her house every evening. She entertains many different monks and, over time, word of her beauty and generosity spreads among them. One particular monk learns of her from a fellow monk who has been a guest at her house, and becomes entranced with the depiction and desires to meet her. Sirimā's appearance and charms represent something of a threat, however, to those monks whose appreciation for the impermanence of beauty and life is somewhat tenuous. To them Sirimā, regardless of her intentions, acts as a servant of Māra, the god of temptation, desire, and death. So when the monk in question decides to venture off to the home of this ex-courtesan whom he loves but has never met, he is giving into attachment in the form of lust and putting himself in danger.

But luckily for him, as it happens Sirimā has fallen gravely ill. She has a high fever, yet decides to attend her customary dinner and greet the visiting monks regardless. Her diseased appearance still does not detain our monk, who reasons than a woman who looks so lovely even in illness must look so much more beautiful when she is at full health.

That evening Sirimā dies, and word is brought to the Buddha of her demise. When he learns of this, he sends a message to the king that Sirimā's body will not be cremated. Rather, it will be brought to the charnel field (a repository for the deceased, where the bodies of criminals and the like are summarily dumped to be devoured by crows and dogs and those of the more fortunate are shrouded, cremated, or both). The Buddha commands, however, that guards will be posted to prevent the animals from disrupting Sirimā's body. After three days, her body has become bloated, and from the nine openings on her body maggots have begun to pour out. Her body by this point had "burst open like a cracked vessel of rice."[157] The Buddha tells the king to have a messenger sent out throughout the town proclaiming that everyone must come to see Sirimā, else incur a penalty of eight *kahāpanas*.

The monk who had fallen for Sirimā has meanwhile fallen into a desolate state, unable to eat his food and ignoring all those who try to speak to him. But when some other monks inform him that they are on their way with the Teacher to see Sirimā, he jumps up excitedly and decides to set off with them. Having not been informed of Sirimā's death, he is eager to see her again. The monks then journey to the burning ground, where the Buddha's community of monks and

[157] Wilson (1996) pg. 85

nuns has gathered along with regular citizens and the king himself. The Buddha tells the king to have a messenger go out and announce to the people that anyone who is willing to pay a thousand *kahāpanas* may have Sirimā. When the king replies that no one will rise to the occasion, the Buddha suggests that the price be lowered. It becomes clear, after successive reductions in price, that no one will have Sirimā—including the lovelorn monk. The Buddha then bids the monks to look closely at the former courtesan's body, and to consider what prices she used to command, whereas now no one will have her for free. "Look at this decorated image, an elevated mass of wounds. This diseased thing is highly fancied, [although] it's neither permanent nor stable."[158] The monk is then delivered from the Buddha from his attachment to Sirimā through this deliberate effort to portray the transience of beauty for even the most charming woman. Though the monk's passion for the woman initially drew him excitedly to view her, in the end he was compelled to contemplate instead her decaying corpse.

The story of the Buddha's own path to awakening also entails coming to view the female form as a disgusting object, though not yet actually a corpse. As the Sanskrit biographies tell it, the prince Gotama Siddhartha was kept sequestered in the family palace by his father the king Suddhodana, because Suddhodana had been told by a seer that Gotama would become either a great king or a great renouncer. To avoid his son choosing the latter option, the king housed him in a beautiful suite of apartments where he would have no opportunity to be exposed to the harshness of everyday reality. In these grand, spacious apartments, the king ensured that Gotama would be entertained by sensuous, enchanting women who would sing and dance for him, lightly playing gold-rimmed tambourines and smiling languid smiles as they laughed and cavorted. These ladies kept Gotama happy for quite some time.

However, as we know, it was not to last. During an outing, despite the king's best efforts Gotama was exposed to the sight of a wizened old man, and on subsequent outings he would view in turn a sick man and then a corpse, and then a man who described himself as a renouncer. Having glimpsed the suffering that truly exists in life, the *bodhissatva* determines that he will follow the monk's example and become a renouncer himself. He returns home and informs his father of his decision, which predictably devastates the man. Gotama

---

[158] Ibid.,pg. 86

then retires to his apartments, in no mood to be entertained by any ladies of the evening. They play beautiful music for him, but it has no pleasing effect. At this point, the Akanistha gods, supporters of the Dharma who wish to encourage Gotma in renouncing the world, decide to step in by casting a magical sleeping spell over the harem women, such that they are suddenly compelled by sleepiness to collapse to the furniture and the floor and drift out of consciousness. When this happens, Gotama looks around at the sleeping women and is struck by how they appear to them. Their sensual beauty gradually gives way to an impression of consciousness having been lost permanently—their bodies messily scattered about, alluring eyes lidded, perhaps never to open again. Their ornaments and garlands have fallen to the floor and broken, and their mouths and legs hang open in a carnal display having turned macabre. Gazing around him, the *bodhissatva* is overcome with a feeling of contempt, and says to himself, "Such is the real nature of women, impure and monstrous, yet a man, deceived by dress and ornaments, succumbs to passion for women."[159]

Later when the Buddha was meditating alone on his way to awakening, the god Māra sends his three daughters to tempt him. The daughters are Lust, Thirst, and Discontent, and they attempt to dissuade the Buddha from enlightenment through seduction and reminders of the kind of pleasures he could be enjoying had he chosen to follow in his father's footsteps. They take on the form of young, attractive, female renouncers and beg Gotama to teach them. When he fails to open his eyes, they assume the form of elderly women, thinking that this will arouse the Buddha's pity. This doesn't work either, but when they try to assume their previous forms they find that they are unable to do so. The Buddha has counter-acted the magic of Māra and confined them to their manifestations as withered women. The three daughters go hobbling off to beg Māra's help in restoring them, but he is unable obscure the way in which the Buddha has revealed the impermanence of Lust, Thirst, and Discontent. Later, the Buddha would go on to portray all *human* women as minions of Māra, god of desire and death. "Womankind," he notes, "is entirely in the snare of Māra."[160] Women thus portrayed become agents complicit in the attempt to confine men to the realm of samsāra. The only way that

[159] Ibid., pg. 67

[160] Ibid., pg. 36

they may become renouncers themselves is to consider *themselves* as disgusting, old, and even dead.

4. **Disgust and reductionism: boundary-crossing**

The nature of contamination is such that the body is pervious in many places. Each orifice constitutes an opportunity for the outside world to enter in. The nature of disgust is such that it is border-transcending. All of the senses are involved—touch, taste, sight, hearing, and smell—and each resists the intrusion. As Nussbaum puts it, borrowing from David Kim, the key concept is that of "crossing a boundary from the world into the self."[161] The body is host to both harmful and helpful bacteria, though the idea of bacteria being helpful is counter-intuitive at first. Our mitochondria have evolved along with us as passengers who kept their baggage packed, keeping their DNA separate from ours. Nevertheless, we navigate the world as though there is something inside which is unified and impenetrable, homogenous, though the environment outside is varied and heterogeneous. It's a kind of essentialism which doesn't bear up to reminders that we are in fact complexly vulnerable creatures. Some of our alien passengers bring disease. The sight of someone with a disability can be discomforting, though familiarity and increased awareness can alleviate this.

The parasite is arguably the ultimate border-transcender. Stories of strange and particularly disgusting parasites make excellent urban legends—a woman consuming octopus eggs and giving birth to baby octopuses (not possible), or spiders laying eggs in a lady's beehive hairdo, or in someone's nose or ears (possible, but unlikely). Stories of tapeworm eggs being used as diet pills teach an object lesson on the price of vanity, and the lengths to which people will go to pursue it.[162] True stories of parasites are often stranger than fiction. The emerald jewel wasp, *Ampulex compressa*, stings a cockroach when she is ready to lay her eggs, injecting venom directly into its brain that inhibits its fear response. The cockroach then becomes compliant to the will of the wasp, and can be guided by its antennae,

---

[161] Ibid., 92

[162] http://www.snopes.com/horrors/vanities/tapeworm.asp Accessed Jan. 23, 2008

used as a kind of leash, to the wasp's underground lair. There the wasp lays an egg in the roach's abdomen, where it will remain simply resting as the wasp's egg hatches after three days. The resulting larva will then eat its way out of the roach, consuming its internal organs in an order which allows for the roach to stay alive as long as possible. The larva then enters the pupal stage and forms a cocoon in the roach's body, from which it will emerge as an adult wasp.[163]

Parasites which perform one or both of those horrifying roles in a human—taking control of the victim's central nervous system, or growing from an implanted egg into a larva stage by consuming the victim's body from the inside—are popular in science fiction. In the film Star Trek 2: The Wrath of Khan, the villain inserts a burrowing worm into a character's brain through his ear canal, which take's over the victim's brain and subverts his will. In the Alien film series, the alien species, called xenomorphs, reproduces by laying eggs which can detect the presence of potential hosts. When a human potential host draws near, the larval stage xenomorph, called a "facehugger," launches itself at the face of the host and grabs onto it firmly (as can be deduced from its name), inserting a tube-like organ into the mouth of the host and thereby implanting an embryo in the host's chest cavity. The facehugger then detaches and dies, but the embryo remains growing in the host's chest, presumably feeding off of him/her in some way until it is ready to emerge and begin adult life, which it does by bursting from the chest of the host and thereby killing him/her. If the xenomorph's designers, American film writer/director Ridley Scott and Swiss artist and surrealist H.R. Giger, did not have the emerald jewel wasp in mind, the resemblance is uncanny.

The lancet fluke is a parasitic worm which takes over the brain of an ant and causes the ant to climb a blade of grass. If it slips and falls before it reaches the top, it will doggedly try again, and attach itself there firmly by its mandibles. The reason that the fluke has taken over this ant vehicle is to get into its terminal host, which is the stomach of a sheep or a goat. When the animal eats the grass, it will consume the ant along with the fluke occupying it, delivering the fluke to its destination where it will reproduce and then eventually lay eggs which are excreted by the animal. Philosopher Daniel Dennett used the lancet fluke as an analogy to religion in his recent book *Breaking the Spell*, trying to convey the concept from memetic theory

[163] Haspel et al (2003)

that ideas may take over our minds in order to benefit themselves (that is, to spread more efficiently) rather than to benefit us. Though he later clarifies that concepts such as democracy and liberty may be spread in a similar manner, it's not surprising that a comparison of one's most cherished ideals to a brain-invading, mind-controlling parasite which leads us to our death might not be very comforting. Charles Darwin used the parasite as a counter-example to the idea that the design of the universe can be viewed as demonstration of God's benevolence: "It is derogatory that the Creator of countless systems of worlds should have created each of the myriads of creeping parasites." Of wasps in particular, he said, "I cannot persuade myself that a beneficient and omnipotent God would have designedly created the Ichneumonidae [one group of parasitic wasps] with the express intention of their feeding within the living bodies of Caterpillars."[164] If the idea of a parasite feeding off the body of its host is intuitively terrifying, the idea of it feeding off of the host's mind is even more so. As Carl Zimmer notes in his book *Parasite Rex*, Hitler used every facet of the notion of the parasite as a metaphor to arouse disgust and enmity for the Jews, portraying society as a healthy host which becomes endangered with the "infestation" of a Jewish population. Marx and Lenin saw the bourgeoisie as the parasites of society, and a parasitologist named Horace Stunkard would go on to describe parasitism in 1977 as typifying the nature of the welfare state. Konrad Lorenz used the existence of parasites as a particular warning to humans: "If one judges the adapted forms of the parasites according to the amounts of regressed information, one finds a loss of information that coincides with and completely confirms the low estimation we have of them and how we feel of them. . . A retrogression of specific human characteristics and capacities conjures up the terrifying specter of the less than human, even of the inhuman."[165]

The fear that parasites elicit in us is deep down, Zimmer says, a kind of fear of reductionism. Whether in biology or science fiction (and the two are not always easy to distinguish), parasites taking over our bodies and minds represents a reminder that we are not somehow separate from the world, or even from the other animal species on it. We are flesh, even our thoughts could not exist without the chemical and electrical transactions of neurons in the brain, and that brain is susceptible. If a wasp can take over the mind of a cockroach, then

---

[164] Quoted in Zimmer, Carl *Parasite Rex* (2000), pg. 14-15

[165] Ibid., pg. 22

who is to say something there is a similar parasite who can do a similar thing to us? We recoil in horror, our privileged position taken from us. People want to be more than the sum of their parts, which is why Francis Crick's "astonishing hypothesis" that we are "no more than the behavior of a vast assembly of nerve cells and their associated molecules"[166] can be so discomforting. If Terrence Deacon is right and our propensity to create symbols brings with it the byproduct of mind/body dualism, then it's not surprising if elements that problematize that view should be cordoned off and given a social category. A commitment to the symbol that is the self may well carry with it the baggage of disdaining some parts of our human nature. So if there are reminders hanging around that we are, after all, just vehicles made of meat traversing the world until we die—just like the ants are doing-- perhaps sociomoral disgust presents an attempt to control and compartmentalize that fear.

5. **Conclusion**

The principle point I have attempted to make in this chapter is that disgust is a very deep-seated response which we have evolved to protect us from contamination by toxic or parasitic things in the world, but when elaborated by symbol-using human beings, it can become a tool for visceral moral condemnation. Disgust can be described as that impulse which drives us to differentiate ourselves from, and close our boundaries off to, the ugly side of biological existence. Returning to Deacon's sentiments which began this chapter, disgust aids us in making symbols of ourselves, establishing personhood as something which is beyond material reality, beyond mortality. Moral condemnation becomes more severe when disgust is involved in the scenario being considered. Reactions of moral disgust are therefore hard to justify reflectively, because they stem from intuition rather than conscious moral reasoning. It closes the door to empathy, making it harder or even impossible to recognize what we have in common with other human beings.

Feelings of disgust are aroused when we consider strange new medical advances, when those advances entail crossing boundaries at the end of life, the beginning of life, or the notion of a person being

[166] Crick, Francis

contained within his or her own body and going throughout life with all of the same parts he or she was born with—no more, no less. These are the domains which are perceived as belonging to God, who is perceived as having created us as we are for a reason. When a disgusting behavior is described as a "crime against nature," it is this divine order which is viewed as being violated. Homosexuals arouse disgust because their existence challenges established roles of gender, calling into question the issue of whether penetration is really what establishes who is male and who is female. The Buddha taught his followers to regard the female body with disgust in order to release themselves from earthly desires entirely, believing that attraction to women is a tool of the world to prevent access to enlightenment. The notion of parasites causes us to shudder, calling into question whether we can really think of ourselves as separate from our own bodies. This idea is so compelling that it enters our storytelling again and again. Disgust, I would argue, is perhaps the most compelling argument for humanity's discomfort with being identified as wholly material, wholly biological. It represents the walls of a panic room in which we can hide from material existence when it becomes too frighteningly real.

Conclusion

1. Agency

I began this dissertation with a discussion of what it means to be an agent and to think about agents. Agency is defined here as the state of being *intentional*—of possessing goals and making efforts to achieve them. As agents ourselves, we have an overwhelming interest in perceiving agency in the world around us, since other agents have goals of their own which may be in accordance or in conflict with our own. It also seems that the world is easier to live with when we can talk to it—even if we know full well that our car isn't listening when we shout at it for not starting, or that the computer doesn't care in the slightest that we have to print this document or else miss a deadline, for some reason we do it anyway. This inexorable impulse to humanize the world, when it becomes earnest, serious, and ingrained, may become the basis of religion. This portrayal of the origins of gods and spirits is natural, continuous with our other attempts to figure out how the world works and use it to our benefit as possible. But in order for the concept of a divine being to "stick," we must be able to have certain expectations of it. Such agents must be like other people in certain ways—they must be able to witness human behavior as humans do, and they must care about human behavior as humans do. They have access to knowledge which we do not, strategic knowledge. And they may be prone to act on that knowledge, which is why it is important to curry their favor, or possibly avoid them altogether. Whichever is the case, this knowledge and interest make them relevant to human existence.

The inquiry into the cognitive origin of beliefs in supernatural concepts—and what constitutes a supernatural concept in the first place—is ongoing. But the fact that agency should be centrally located in this inquiry is easy to acknowledge. What is less obvious, however, is whether the term "supernatural" should apply specifically to non-human religious agents. The qualification that in order to be supernatural an entity must be minimally counterintuitive requires that in order for an agent to be supernatural, it must violate inferences which stem from our mental ontological categories. But is immaterial and invisible agency really counterintuitive? The argument of this

dissertation has been that it is not. Its suggestion is that the matter of what is really counterintuitive about human psychology and biology needs to be examined more closely, in order to push at the borders of supernatural agency. What if we intuitively treat other humans as being supernatural, in the sense of their agency being "unnatural" or preceding nature, every day? My intention here is to suggest that this in fact what we do, and that it can be seen in our moral psychology.

2. Empathy

Our ability to practice empathy, to understand the goals and motivations of others, can be aided (or sometimes inhibited) by the idea that those come from a place we cannot see.

Empathy is a means of connecting with other people by identifying with them, understanding their thoughts and feelings through imaginative projection. Explicit, cognitive empathy is an ability that only humans seem to have, though other higher order primates appear to have a precursor. This is because much of our ability to interact socially, and thus empathize on an advanced level, appears to rely on our willingness and capacity to engage in joint attention scenes as children. Joint attention requires the ability to comprehend facial expressions and eye direction, as well as the fact that the contents of another person's mind are not necessarily the same as our own. Small children delight in pointing declaratively in order to share the experience of noticing, and are disappointed when the person for whom they are pointing does not actively share in the experience. This kind of exchange is important for the acquisition of language. Through joint focus on an object, and then confirming the name for the object, a parent can teach the child the names for things in the world and then later converse with the child about things they saw earlier in the day, or even things in the child's imagination. This understanding that other people may see and believe differently, even falsely, allows us to mentalize—to formulate working theories about how and why people behave the way they do based on unseen factors. Children can then come to interpret all of the social stories going on around them in mentalistic terms. They know that when Sally returns to the room she will look for her ball where she left it, rather than where Ann has hidden in it in her absence.

We have intuitions about what happens when one person's spirit possesses another person's body. When this happens, the spirit

displaces the one which would normally have been in occupation—spirits do not fuse to form one spirit, but rather generally speaking there is a rule of one spirit to one body. What's more, those attributes which are transferred tend to be those which are mental rather than biological, mental states which are connected to beliefs and desires of the possessing spirit. Evidence from early development suggests that humans may be mind/body dualists from birth. Infants automatically seem to see a clear distinction between objects and agents, and treat them very differently—they expect agents to react to their attempts to interact, and they expect people who are reaching for things to continue to reach if the object they are grasping for is moved away. An infant who is attempting to communicate with a parent will become frustrated if he/she notices that the face of the parent is still and non-responsive. As children grow up, they intuit different responsibilities for the mind as compared to the brain. The mind's job is to imagine, to love, to exhibit the child's personality, whereas the brain has more mundane cognitive duties. Children appear to hold onto this division into adulthood.

A person with a mentalizing deficiency such as autism, however, must take what is usually a more tedious route of relying on behaviourism and trial and error, at least until they have gained considerable experience in the world (for a quite a few, deliberate training) in how to "read" other people's intentions and emotional states. This has been described as a lack of empathy, but it is not just empathy in the sense of a sympathetic reaction to someone who is suffering. Though it may include this, the deficiency is more a matter of being unable to practice perspective taking. It seems possible that those with autism are less likely to share our intuitive dualism, however to my knowledge no intensive research has been conducted in order to establish or refute this

Perception of similarity to other people is a strong elicitor of empathy, as well as is familiarity. Accordingly, denial of empathy sends a strong signal of dissimilarity. Negative empathy, a kind of pleasure in the suffering of others, is "there but for the grace of God go I" sort of emotion. We make ourselves safer by identifying in others traits which we might have had, but have avoided—and then we are able to take pleasure when those traits bring others down. Refusal to understand someone is a subtle kind of negative empathy. It says "I refuse to validate you by speaking the same language." Disorders such as autism and sociopathy can inhibit the practice of empathy from the bottom up, but the desire to differentiate oneself from others allows us to inhibit it from the top down.

3. Responsibility

Our ideas about the locus of human responsibility, our ability to apply blame, and our propensity to believe in the concept of evil are supported by a concept of the self existing separate from the body.

Living socially means having to learn how to cooperate, and cooperation requires moral responsibility and the assignment of praise and blame when necessary. Human moral reasoning functions to try and determine who caused negative events so that responsibility can be assigned where appropriate. Experimental evidence shows that our tendency to identify people as the cause of an event increase when that event has negative consequences. The impulse to blame is strong, and it is not necessarily always—or even mainly—a matter of conscious reflection and careful reasoning. Our intuitions strongly guide moral decision-making as well. We seem to be more resistant to harming one for the benefit of many when that harm requires direct, physical action. Our moral emotions guide us to non-reflective moral decisions in a sort of auto-pilot fashion, unless we are directly asked to justify those decisions—in which case we may be caught in the spotlight, morally dumbfounded. The kind of moral scenarios which our ancestors have confronted for thousands of years have helped to determine what kinds of issues of responsibility are easier for us to process, and hence which are most likely to attract our attention. The idea that at any particular moment, there is one future—that is, determinism--- is counter-intuitive because it implies that there is no free will, and hence no one can properly be held responsible for their actions. This is an incorrect conclusion, however, as our conscious decision-making is a causal factor along with all other factors we are faced with, internally and externally. Removing all of these other causal factors from the scenario does not increase freedom, but rather robs us of it—we would not longer have anything to make decisions *with* or *about*. Agency surrounded by nothingness or randomness is agency with nothing to do.

Evil is a supernatural concept in that it presents as stark a disparity between victim and perpetrator as possible by portraying the perpetrator as having supposedly absolute free will—that is, portraying the perpetrator as an unmoved mover. The notion of evil essentializes people, removing them from both the causal influences of their physical and social surroundings as well as those of their

neurological and genetic underpinnings. It over-attributes their intentionality in order to make evil the ultimate reason for their actions, prohibiting further inquiry. The attribution of evil marks the perpetrator as an outsider, one in complete possession of free will and therefore one who can be held fully responsible and punished. Evil paints only in black and white. This moral levitation, this hovering over any and all possibly exculpatory circumstances, allows for an easy, assured negative empathy. It closes the door to attempts at understanding, and therefore identifies who is the victim and who must be punished with stark, clean lines.

4. **Purity**

Our beliefs about human dignity and the distinction between humankind and the rest of the animal world, as well as the barriers between self and non-self, rest in part on a portrayal of a pure soul untainted by worldly sources of disgust.

Disgust is rooted in our need to avoid things in the world which poison, pollute, and invade. The primary elements which arouse disgust have to do with death, decay, parasites, and reminders of those things. Animal reminder disgust is so called because humans intuitively tend to view themselves as divorced from the rest of the animal kingdom, especially when those animals with which we are confronted are perceived as feeding on death or suffering. The natural urge when disgust is aroused is to remove oneself, as quickly as possible, from the source of that arousal—or else, as in the case of parasites, remove that source from oneself. Disgust is a moral emotion because when inserted into a social scenario, it creates a moral valence where none previously existed. The fact that an otherwise benign behavior is disgusting is often enough to cause it to be perceived as immoral—another instance of moral dumbfounding, and another barricade to empathy. In order to prompt their dehumanization, groups of people whose members are of different religions, races, genders, and sexual orientations throughout history have been associated with disgusting objects and animals. Unlike the categorization of evil which affirms humanity but also absolute blame, disgust denies humanity and hence dignity. The concept of purity is undoubtedly associated with religious identity because it connotes remove—humankind is removed from the rest of existence by virtue

of our relationship with the divine. Humans who are made or make themselves disgusting are incapable of maintaining this relationship, and hence become profane.

Body-envelope violations arouse disgust in the form of aversion to open wounds, disabilities and sickness as well as some medical practices intended to rectify such things—particularly medical practices which challenge the boundaries at the beginning and end of life as well as those of the body itself. The classic objection when this kind of disgust is provoked is that the provocateurs are Dr. Frankenstein re-invented, over-stepping their allotted role into God's territory.

"Crimes against nature" are violations of this divine order. Homosexuals challenge roles of gender concerning the issue of whether penetration is really what establishes who is male and who is female, and therefore are labeled as disgusting. The Buddha taught his followers to regard the female body with disgust in order to remove their earthly desires which tied them to the world, preventing access to enlightenment. Parasites call into question whether we can really think of ourselves as separate from our own bodies. This idea is re-occurring in our stories because it is simultaneously repellant and entrancing, like good horror movies are. Disgust represents our discomfort with being identified as wholly material, wholly biological. If our dual mind-body nature makes us unique in the world, then disgust and the intuitive moral inferences which stem from it serve to create the barrier around that uniqueness, bearing it up and firmly establishing the boundary between ourselves and everything else.

5. **Duality**

The significance of intuitive belief in a soul is that our affirmation or denial of it in others has a dramatic effect on our ability to empathize with them.

Anthropomorphism is a means of empathizing with the non-human world—making that world seem more like us so that it is easier to understand, more comfortable to live in. To some extent it is also a means of empathizing with the human world, since we cannot truly experience what it is like to be someone else but rather must imaginatively project as best as we can, attempting to emulate what is going on inside their heads using what we have in ours. In that sense, we are anthropomorphizing other *people*. Dualism, assuming that

there is an invisible, immaterial essence to all of us which is made of similar stuff, can provide a short cut to empathy. I may not know the neurological substrates for the feeling of betrayal you had when God did not save your brother's life in his car accident, but our common humanity allows me to approximate it and even feel it myself. But the flip side of that coin is that it comes accompanied by a sense of free will and responsibility which depend on immateriality, remove from the world, which makes it possible to believe in evil people. Evil is the simultaneous acknowledgement of someone's humanity and condemnation of them on that basis—in contrast with disgust, which is elicited by the identification of someone as *entirely* material, soulless, inhuman. I refer to this depiction of the self in the dissertation as a skyhook kind of agency, one in which intention precedes all else, all material causes. In this way we intuitively see others as supernatural, in the sense of being above nature, *before* nature. In this way we intuit the soul.

Abstract

The intent of this doctoral dissertation is to explore the origins of soul belief by examining the traits and responsibilities we intuitively give that entity. Using empirical data from cognitive inquiries into folk psychology, I will argue that there exists a need to re-examine the theoretical view of the soul as a supernatural concept—that is, a concept involving top-down agency. Our folk psychology invokes mind-body dualism, the notion of an aspect of human nature which is separate from the body, and I will explore some of the ways in which this happens. Our ability to practice empathy, to understand the goals and motivations of others, can be aided (or sometimes inhibited) by the idea that those come from a place we cannot see. Our ideas about the locus of human responsibility, our ability to apply blame, and our propensity to believe in the concept of evil are supported by a concept of the self existing separate from the body. Our beliefs about human dignity and the distinction between humankind and the rest of the animal world, as well as the barriers between self and non-self, rest in part on a portrayal of a pure soul untainted by worldly sources of disgust. I will argue that this notion of agency which is separate from and precedes the body and the world is intuitive, causing our everyday thinking about the self to, in effect, be a kind of supernatural thinking—which means that souls and gods may be similar in our cognition in ways that deserve further consideration.

Danish

Afhandlingen undersøger sjæletroens oprindelse ved at analysere de træk og egenskaber, som vi intuitivt tillægger sjælen. Med inddragelse af empiriske resultater fra kognitive undersøgelser af folkepsykologien, hævder jeg, at der er et behov for grundige overvejelser over den teoretiske anskuelse af sjælen som et overnaturligt begreb. Menneskets intuitive psykologi forudsætter en krop-og-sind-dualisme, dvs. ideen om at noget i den menneskelige natur er adskilt fra kroppen. Jeg undersøger nogle af de måder, hvorpå denne idé fungerer. Vores evne til at praktisere empati, at forstå andre menneskers mål og motivationer, kan forøges eller forhindres ved ideen om, at disse mål og motivationer kommer fra et sted, vi ikke kan se. Den

opfattelse, at vi har et selv adskilt fra kroppen, gør os i stand til at placere det menneskelige ansvar et sted, at tillægge mennesket skyld og at tro på ideen om ondskaben. Opfattelsen af den rene sjæl ubesmittet af verdslig smuds gør det muligt for os at tro på menneskets værdighed og på adskillelsen mellem mennesker og dyr. Det er min påstand, at denne idé om en aktør, som er adskilt fra og som kommer før kroppen og verden, er intuitivt. Denne idé er årsagen til, at vores dagligdags forståelse af selvet udgør en slags overnaturlig tænkning, og at vores kognitive repræsentation af sjæle og guder udviser en række fællestræk, som fortjener en nærmere undersøgelse.

Bibliography

Ainslie, George
2006 “Cruelty may be a self-control device against sympathy” *Behavioral and Brain Sciences* Vol. 29, Iss. 3; 224-225

Atran, Scott & Ara Norenzayan
2004 “Religion’s evolutionary landscape: Counterintuition, commitment, compassion, communion.” *Behavioral and Brain Sciences* 27

Bailey, Ronald
2003 “Pulling Our Own Strings: Philosopher Daniel Dennett on determinism, human ‘choice machines,’ and how evolution generates free will” *Reason Online*
Accessed 8/28/07 <http://www.reason.com/news/show/28782.html>

Bandura, Albert, Bill Underwood, & Michael Fromson
1975 “Disinhibition of Aggression through Diffusion of Responsibility and Dehumanization of Victims*” Journal of Research in Personality* 9, 253-269

Baron-Cohen, Simon
1995 "First lessons in mind reading," *The Times Higher Education Supplement*, July 16

Baron-Cohen, Simon, Alan Leslie, & Utra Frith
1984 “Mechanical, behavioural and intentional understanding of picture stories in autistic children” British Journal of Developmental Psychology, 4, 113-25

Baron-Cohen, Simon
1997 *Mindblindness: An Essay on Autism and Theory of Mind.* MIT Press

Baron-Cohen, Simon

2004 *The Essential Difference: Male and Female Brains and the Truth About Autism.* Basic Books: New York, New York.

Baron-Cohen, Simon

2001 "The Geek Syndrome" *Wired magazine*, 9.12

Baumeister, Roy

1997 *Evil: Inside Human Violence and Cruelty* W.H. Freeman and Company

Bering, Jesse & Becky Parker

2006 "Children's Attributions of Intentions to an Invisible Agent" *Developmental Psychology* Vol. 42, No. 2. 53-262

Bering, Jesse

2002 "Intuitive Conceptions of Dead Agents' Minds: The Natural Foundations of Afterlife Beliefs as Phenomenological Boundary." *Journal of Cognition and Culture* 2.4

Bering, Jesse

2002 "The Existential Theory of Mind" *Review of General Psychology* Vol, 6, No. 1, 3-24. Educational Publishing Foundation

Bering, Jesse

2006 "The folk psychology of souls" *Behavioral and Brain Sciences 29*, 453-498.

Bering, Jesse

2003 "Towards a cognitive theory of existential meaning" *New Ideas in Psychology* 21, 101-120

Billig, Michael

2001 "Humour and Hatred: The Racist Jokes of the Ku Klux Klan." *Discourse and Society*, Sage Publications. Vol. 12 (3) 267-289

Bloom, Paul and Deena Weisberg
2007 "Childhood Origins of Adult Resistance to Science" *Science* Vol. 316, 996-997

Bloom, Paul
2004a *Descartes' Baby: How the Science of Child Development Explains What Makes Us Human.* Basic Books

Bloom, Paul
2006 "My Brain Made Me Do It*" Journal of Cognition and Culture* Vol. 6, Issue 1-2

Bloom, Paul
2004b "Natural-Born Dualists" posted on *Edge Third Culture* (5/13/04)
Accessed 10/27/2005 <http://www.edge.org/3rd_culture/bloom04/bloom04_index.html>

Bloom, Paul
2004c "The Duel Between Body and Soul" *The New York Times* Editorial Desk, Sept.10

Boyer, Pascal
2000 "Functional Origins of Religious Concepts: Ontological and Strategic Selection in Evolved Minds" *The Journal of the Royal Anthropological Institute* Vol. 6, No. 2, 195-214

Boyer, Pascal
2002 *Religion Explained*. Vintage

Bourguignon, Erika
2004 "Suffering and healing, subordination and power: Women and possession trance" *Ethos, 32*(4), 557-574.

Bulbulia, Joseph
2004 "Religious Costs as Adaptations that Signal Altruistic Intention" *Evolution and Cognition* Vol. 10, No. 1, 19-42

Clark, Andy
1997 Being There*: Putting Brain, Body, and World Together Again* MIT Press

Cohen, Emma
2007 *The Mind Possessed.* Oxford University Press

Cohen, Emma & Justin Barrett
2008 "When minds migrate: conceptualizing spirit possession." *Journal of Cognition and Culture*, Volume 8, No. 1-2, 23-48

Davis, Hank
2002 "Too early for a neuropsychology of empathy," *Behavioral and Brain Sciences* 25 (1) 32-33.

Dawkins, Richard
1989 *The Extended Phenotype*. Oxford Paperbacks

Dawkins, Richard
1976 *The Selfish Gene.* Oxford University Press

Dennett, Daniel
1985 *Darwin's Dangerous Idea.* Simon & Schuster

Dennett, Daniel
2003 *Freedom Evolves*. Penguin Books

Dennett, Daniel
1987 *The Intentional Stance*. MIT Press

deWaal, Frans
1999 "Anthropomorphism and Anthropodenial: Consistency in our thinking about humans and other animals." *Philosophical Topics* 27: 255-280.

deWaal, Frans
2006 *Primates and Philosophers.* Princeton Universtiy Press

di Pellegrino, G., L. Fadiga, L. Fogassi, V. Gallese & G. Rizzolatti
1992 "Understanding motor events: A neurophysiological study," *Experimental Brain Research* 91: 176-80.

Doherty, Brian
2007 "You Can't See Why on an fMRI: What science can, and can't, tell us about the insanity defense." *Reason* July issue

Driver, Julia
2008 "Attributions of Causation and Moral Responsibility." in Moral *Psychology Volume 2*, *The Cognitive Science of Morality: Intuition and Diversity.* Walter Sinnott Armstrong, ed. MIT Press, 423-439

Driver, Julia
2008 "Kinds of Norms and Legal Causation: Reply to Knobe and Fraser and Deigh" in *Moral Psychology Volume 2, The Cognitive Science of Morality: Intuition and Diversity.* Walter Sinnott Armstrong, ed. MIT Press, 459-461

Dwyer, Susan
2008 "How Not to Argue That Morality Isn't Innate: Comments on Prinz" in *Moral Psychology Volume 1, The Evolution of Morality: Adaptations and Innateness.* Walter Sinnott Armstrong, ed. MIT Press, 407-418

Eisenberg, Nancy & Janet Strayer, eds.
1987 *Empathy and Its Development*, Boston: Cambridge University Press.

Frith, Uta
2003 *Autism: Explaining the Enigma*, Wiley-Blackwell

Frith, Uta
2007 "Language and Theory of Mind" Presentation at the "Language in Cognition, Cognition in Language" conference at the University of Aarhus, Denmark October 13

Gallese, Vittorio
2001 "The 'Shared Manifold' Hypothesis: From Mirror Neurons to Empathy," *Journal of Consciousness Studies* 8, No.5-7, 43

Gallup, Gordon
1998 "Self-awarenes and the evolution of social intelligence" *Behavioural Processes* 42

Greene, Joshua
2003 "From neural 'is' to moral 'ought': what are the moral implications of neuroscientific moral psychology?" *Nature* Vol. 4, 847-850

Guthrie, Stewart
1980 "A Cognitive Theory of Religion." *Current Anthropology* Vol. 21, No. 2

Guthrie, Stewart
1993 *Faces in the Clouds: A New Theory of Religion*. Oxford University Press.

Guthrie, Stewart
2008 "Spiritual Beings: A Darwinian, Cognitive Account." *The Evolution of Religion: Studies, Theories, and Critiques.* Collins Foundation Press, 239-245

Haidt, Jonathan & Joshua Greene
2002 "How (and where) does moral judgment work?" *TRENDS in Cognitive Sciences* Vo. 6 No. 12

Haidt, Jonathan & Craig Joseph
2004 "Intuitive ethics: how innately prepared intuitions generate culturally variable virtues" *Daedalus*, Fall 55-66

Haidt, Jonathan &Matthew Hersh
2001 "Sexual Morality: The Cultures and Emotions of Conservatives and Liberals*" Journal of Applied Social Psychology* 31, 1, 191-221

Haidt, Jonathan
2001 "The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment." *Psychological Review* Vol. 108, No. 4, 814-834

Haidt, Jonathan
2006. *The Happiness Hypothesis*. New York: Basic Books.

Haidt, Jonathan & Jesse Graham
(in press) "When Morality Opposes Justice: Conservatives Have Morals Liberals May Not Recognize." *Social Justice Research*.

Harris, Paul
2000 *The Work of the Imagination* Blackwell Publishers

Haspel, Gal, Lior Ann Rosenberg & Frederic Libersat
2003 "Direct Injection of Venom by a Predatory Wasp into Cockroach Brain" *Journal of Neurobiology* 56 (4): 287-292

Hatcher, Carol
2008 "When Morality is Hard to Like" *Scientific American Mind* Vol. 19 No. 1, Feb/March, 34

Hauser, Marc
2006 *Moral Minds* Harper Collins, New York

Hauser, Marc, Liane Young, & Fiery Cushman
2008 "Reviving Rawls' Linguistic Analogy: Operative Principles and the Causal Structure of Moral Actions" in *Moral Psychology Volume 2, The Cognitive Science of Morality: Intuition and Diversity.* Edited by Walter Sinnott Armstrong, ed. MIT Press. 107-143

Heider, Fritz & Marianne Simmel
1944 "An experimental study of apparent behaviour". *American Journal of Psychology*, 13

Hoffman, Martin
2000 *Empathy and Moral Development: Implications for Caring and Justice,* Cambridge University Press

Howley, Kerry
2006 "Ova For Sale: The art of the deal in the grey market for human eggs" *Reason* October

Johnson, Philip
1997 *Defeating Darwinism by Opening Minds* InterVarsity Press

Kass, Leon, ed.
2003 *Being Human: Core Readings in the Humanities.* New York: W. W. Norton & Company

Kass, Leon
1997 "The Wisdom of Repugnance." *The New Republic*. 2 June,17-26.

Knobe, Joshua
2003 "Intentional action in folk psychology: an experimental investigation." *Philosophical Psychology*, 16, 309-324.

Knobe, Joshua & Arudra Burra
2006 "The Folk Concepts of Intention and Intentional Action." *Journal of Cognition and Culture.* Vol. 6: Isssue 1-2, 113-132

Knobe, Joshua & Ben Fraser
2008 "Causal Judgment and Moral Judgment: Two Experiments" in *Moral Psychology Volume 2, The Cognitive Science of Morality: Intuition and Diversity.* Walter Sinnott Armstrong, ed. MIT Press. 442-447

Koch, Gretchen
2008a "Does Evolution Threaten the Soul?" *The Evolution of Religion: Studies, Theories, and Critiques.* Collins Foundation Press 371-377

Koch, Gretchen
2008b "Dualism, Moral Judgment, and Perceptions of Intentionality." *The Evolution of Religion: Studies, Theories, and Critiques.* Collins Foundation Press. 257-262

Koch, Gretchen (forthcoming) "Care of the soul: Empathy in a dualistic worldview." *Origins of Religion, Cognition and Culture.* Armin W. Geertz, ed. Equinox

Koch, Gretchen
2008c "Full of Sound and Fury: The Media Response to Dennett." *Method and Theory in the Study of Religion*. Vol. 20, No. 1, 36-44

Lawson, E. Thomas & Robert McCauley
1990 *Rethinking Religion: Connecting Cognition and Culture*. Oxford University Press

Lebo, Lauri

2008 *The Devil in Dover* Free Press

Lee, Patrick and Robert George
2004 "Delusions of Dualism." *National Review Online*. 9/30/04. Accessed 12/30/05, <http://www.nationalreview.com/comment/patrick_george200409300821.asp>.

Lopez, Kathryn
1998 "Egg heads: young women in need of cash are increasingly deciding to sell their bodies" *National Review*, Sept. 1

Loveland, Katherine & Susan Landry
1986 "Joint Attention and Language in Autism and Developmental Language Delay" *Journal of Autism and Developmental Disorders* Vol. 16, No. 3

Mash, Eric and David Wolfe
2002 *Abnormal Child Psychology*. Wadsworth Group

Moll, Jorge et al.
2005 "The Moral Affiliations of Disgust: A Functional MRI Study." *Cognitive and Behavioral Neurology* Vol. 18, No. 1

Murphy, Scott, Jonathan Haidt, & Fredrik Bjorklund
(unpublished) "Moral dumbfounding: When intuition finds no reason."

Nahmias, Eddy
2006 "Folk fears about freedom and responsibility: Determinism vs. reductionism" *Journal of Cognition and Culture*, Vol. 6 No. 1, 215-237.

Nisbett, Richard & Dov Cohen
1996 *Culture of Honor: The Psychology of Violence in the South*. Westview Press.

Nussbaum, Martha
2004 "Danger to Human Dignity: the Revival of Disgust and Shame in the Law" *The Chronicle Review* Aug. 6th

Nussbaum, Martha

2004 *Hiding from Humanity*: *Disgust, Shame, and the Law*. Princeton University Press.

O'Gorman, Rick, et al.
2005 "Altruistic Helping and Punishment," *Evolution and Human Behavior* 26 (5), 375-387.

Preston, Stephanie & Frans deWaal
2002 "Empathy: Its Proximate and Ultimate Basis." *Behavioral and Brain Sciences*, 25, 1-72.

Prinz, Jesse
2008 "Is Morality Innate?" in *Moral Psychology Volume 1, The Evolution of Morality: Adaptations and Innateness*. Walter Sinnott Armstrong, ed. MIT Press pg. 367-406

Prinz, Jesse
2007 "Can Moral Obligations Be Empirically Discovered?" *Midwest Studies in Philosophy*, XXXI, 271-291

Prinz, Jesse
2006 "The emotional basis of moral judgments*." Philosophical Explorations* Vol. 9, No. 1, 29-43

Prinz, Jesse
2008 "Resisting the Linguistic Analogy: A Commentary on Hauser, Young, and Cushman" in *Moral Psychology Volume 2, The Cognitive Science of Morality: Intuition and Diversity.* Walter Sinnott Armstrong, ed. MIT Press. 157-170

Prinz, Jesse
2008 "Reply to Dwyer and Tiberius" in *Moral Psychology Volume 1, The Evolution of Morality: Adaptations and Innateness.* Walter Sinnott Armstrong, ed. MIT Press 427-439

Pyysiäinen, Ilkka
2004 "Intuitive and Explicit in Religious Thought" *Journal of Cognition and Culture* 4., 123-150

Richert, Rebekah & Paul Harris
2006 "The Ghost in My Body: Children's Concept of the Soul" *Journal of Cognition and Culture* 6.3-4, 409-427

Rozin, Paul, Jonathan Haidt & Clark McCauley
2000 "Disgust." in *Handbook of Emotions,* M. Lewis & J. Haviland eds. Guilford Press, 637-653

Rozin, Paul, M. Markwith, & C. Stoess
1997 "Moralization and Becoming a Vegetarian: The Transformation of Preferences into Values and the Recruitment of Disgust." *Psychological Science* 8: 67-73.

Schunerman, Randy, quoted in Kushner, Adam
2008 "How McCain Sees the World," *Newseek* Vol. CLII, Sept. 15, 64

Shaw, Laura et al.
1994 "Empathy Avoidance: Forestalling Feeling for Another in Order to Escape the Motivational Consequences." *Journal of Personality and Social Psychology* 67 (5) 879-887.

Shweder, Richard, Nancy Much, Manamohan Mahapatra, & Lawrence Park.
1997 "The 'Big Three' of morality (autonomy, community, and divinity) and the 'Big Three' explanations of suffering." in *Morality and Health* A.Brandt & P. Rozin, eds. Routledge, 119-169

Shweder, Robert & Jonathan Haidt.
1993 "The Future of Moral Psychology: Truth, Intuition, and the Pluralist Way." *Psychological Review* Vol. 4, No. 6

Tiberius, Valerie
2008 "The Nativism Debate and Moral Philosophy: Comments on Prinze" in *Moral Psychology Volume 1, The Evolution of Morality: Adaptations and Innateness.* Walter Sinnott Armstrong, ed. MIT Press pg. 419-426

Tomasello, Michael, Josep Call, & Brian Hare
2003 "Chimpanzees understand psychological states – the question is which ones and to what extent" *Trends in Cognitive Science* Vol. 7, Issue , 153-156

Tomasello, Michael
2001 *The Cultural Origins of Human Cognition*, Harvard University Press

Tomasello, Michael & Josep Call
1999 "The effect of humans on the cognitive development of apes," *in Reaching into Thought: The Minds of Great Apes.* Cambridge University Press, 371-403

Tomasello, Michael, Malinda Carpenter, Josep Call, Tanya Benhe, & Henrike Moll
2005 "Understanding and Sharing Intentions: The Origins of Cultural Cognition" *Behavioral and Brain Sciences* 28, 675-691

Wighton, Kate
2007 "Welcome to the body shop" The Times, Nov. 3

Wilkinson, Francis
2008 "The Discover Interview: Leon Kass" Discover Magazine, Feb, 62

Wilson, Rob
2008 "Meaning Making and the Mind of the Externalist," in *The Extended Mind*, Richard Menary, ed. Ashgate

Wolfe, Tom
2007 *Hooking Up*, Picador

Zimmer, Carl
2001 *Parasite Rex: Inside the Bizarre World of Nature's Most Dangerous Creatures.* Free Press.