

Stony Brook University



OFFICIAL COPY

The official electronic file of this thesis or dissertation is maintained by the University Libraries on behalf of The Graduate School at Stony Brook University.

© All Rights Reserved by Author.

**Critique of a Homuncular Model of Mind
from the Neo-Vitalist Perspective**

A Dissertation Presented

by

Peter Vere Warden

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Philosophy

Stony Brook University

December 2007

Stony Brook University

The Graduate School

Peter Vere Warden

We, the dissertation committee for the above candidate for
the Doctor of Philosophy degree, hereby recommend
acceptance of this dissertation

(Dissertation Advisor)
Harvey Cormier
Associate Professor
Department of Philosophy

(Chairperson of Defense)
Edward Casey
Distinguished Professor
Department of Philosophy

(Internal Reader)
David Allison
Professor
Department of Philosophy

(Outside Reader)
Bernhard Flynn
Professor
Empire State College

This dissertation is accepted by the Graduate School

Lawrence Martin
Dean of the Graduate School

Abstract of the Dissertation

**Critique of a Homuncular Model of Mind
from the Neo-Vitalist Perspective**

by

Peter Vere Warden

Doctor of Philosophy

in

Philosophy

Stony Brook University

2007

This dissertation is a philosophical defense of the Bergsonian notion that that which separates the quick from the dead is a special solidarity of the past with the present and future. Named variously memory, spirit, or duration—depending upon the acuity of phenomenological focus—this solidarity in conscious beings involves a metaphysical economy whereby ever flowing images ultimately grounded in a sensorial body continuously serve to map out solutions for survival and well-being.

The integrity of thought and its warm and intimate substrate—the bodily, is what will be seen to stitch the most profound self together in the most primordial fashion. That the temporal aspect of this stitching is merely the obverse of its sensorial component will become clear after a contrapositive exposition of the temporality of the purely spatial and homogeneous.

Once this wholism has been demonstrated, as outlined above, through its progressive sytematicization in the thought of Henri Bergson, William James, and finally Antonio Damasio—we will proceed to turn it on its head. Returning to the homuncular nature of the operaton of the intellect hypothesized by Dennett, a final section of the essay will offer an interpretation of mind and the body with which it is always connected in light of pain, the survival instinct, and the connectivity of all of the elements of a thinking being, which is alive.

TABLE OF CONTENTS

1. Introduction; Wholism vs. Homuncularism	p. 1
2. William James; Thought and the Body	p. 6
3. Daniel Dennett and the Multiple Drafts	p. 19
4. Homuncularism and Verbal Reports	p. 40
5. The Jamesean Stream of Thought; Critiquing Verbal Reports	p. 52
6. More Homuncularism: Enter Wholism	p. 65
7. Antonio Damasio and Organic Wholism	p. 76
8. Wholism, Pain, and the Environment	p. 85
Bibliography	p. 90

CHAPTER ONE: INTRODUCTION; WHOLISM VS. HOMUNCULARISM

This work seeks to answer two main questions. Firstly: is there a categorical difference between living consciousnesses and consciousnesses that arise from purely artificial or man-made platforms? And secondly: is the temporality of a living consciousness necessarily tied in a causal manner to the wholism and the interdependence of the parts that make up its organic platform? Phrased in the negative, the second question might read: ‘can a robot ever have a stream of thought’?

More specifically, I will be starting off in the early chapters with foundational accounts of the closely related conceptions of the mechanics of the correlation of the phenomenal temporality of consciousness in humans to their very organic embodiments. I will be relying heavily here on the formulations of two turn of the last century thinkers: William James and Henri Bergson. The conceptual frameworks describing this phenomenal temporality have come to be known under the respective rubrics of the “Stream of Thought” (James) and “*la duree*” (Bergson).

While both James and Bergson were overtly engaged in the early formulation of the discipline of psychology (at least as the term was more broadly understood around 1900), it is quite clear that the one (Bergson) was much more invested in questions of abstract metaphysics, especially when it came to questions of the phenomenality of the

temporality of consciousness, while the other (James) was much more grounded in the concrete aspects of human embodiment and the dynamics of the physical substrates of Thought as he saw it. And perhaps this should come as no surprise since Bergson was classically trained as a mathematician, whose instruction was his first vocation—while James was first a medical physician and taught in the medical school at Harvard before coming to the wider fields of investigation of philosophy and psychology.

Both James and Bergson go to great lengths to elucidate and illuminate the essentiality of the nondiscontinuity of the streams of thought which occur in conscious minds. Again, while focussing on varying aspects of this phenomenon—James using a much more biologically based method of analysis, in comparison to Bergson’s more metaphysical ruminations—at the core of both treatments of the questions surrounding the ultimate construction and structure of consciousness lies an interest in and simultaneous hypostatization of the synergistic, open, and yet still somewhat circular economies which characterize the interactions evident amongst mind and body, thought and matter, reflection and action. For each, understanding the physical and specific interpenetration of thought and the body in which it arises, mind and the matter from which it springs, is essential to understanding not only the physics but the *metaphysics* as well which properly apply to a fuller understanding of how Thought comes to be in the human thinker.

So, while the dependence of thought and memory upon a Body intent upon an Action and engaged thusly in an environing situation are very clearly delineated by James and Bergson in their *magnum opera* (*The Principles of Psychology* and *Matter and Memory*, respectively) neither drives home the point that the indecompossibility of

thought and the body from which it arises is intrinsically due to the organic inseparability of all of the parts of the complex machinery which comprise the only heretofore known platforms upon which might arise consciousness and the thought and memory which are nearly apposite to what it is to be human.

Hence I take my departure in this project with William James' theses on the organic dynamism implicit in the thought-organism-environment complex. And once this has been outlined, I would like to take this newfound, or refound, understanding of the organicism which lies at the back of all conscious thought as we know it—and apply it as a meta-hermeneutic to what has been called the “cyborg” nature of our modern, post-modern, or even “non-modern” existence, as the various schools would have it.

In particular I will be looking at how James' understanding of the ‘consciousness experiments’ upon some of the so-called lower vertebrates evolved into his understanding of the relation of brain function to Thought itself, and how the dynamics of his rather biologically based model worked. This will be the subject of the second chapter.

Moving on from James' description of the operation of thought in Chapter 2 in a biological organism, and the crucial connectivity evidenced therein between thought, its more immediate platform the mind/brain, and its more distant bases in the bodily apparatus—in a third chapter I will begin the investigation as to how consciousness might operate and or arise in an artificial system. This will begin with an in-depth analysis of the model of mind and consciousness of one of the better-known modern proponents of machine consciousness—Daniel Dennett.

I will be focussing on Dennett's architectonic of the Multiple Drafts model of consciousness as he calls it with a view toward ultimately unravelling the meaning of its

claim to mesh seamlessly with William James' theses on the stream of thought. Also, my treatment of this central model for Dennett will form the foundation for the comparison of machine based and biologically based models of mind in the following chapters.

Therefore, a fourth chapter will take the Multiple Drafts model as a springboard for a deeper analysis of the Dennettean imputation of a mirror-like translatability between serial and parallel architectures of consciousness, and the more profound neo-homuncular theory of consciousness upon which this is founded.

Furthermore, the fourth chapter will see the opening of the question of verbal reports as ultimate touchstone of consciousness as envisioned by Dennett's Heterophenomenological method. We will begin to see in chapter four how this test of consciousness flies directly in the face of the Jamesean theses on psychology and in particular his formulation of the "stream of thought".

In chapter five I will delve wholeheartedly into the Jamesean theory of the stream of thought, and will further flesh out how this formulation has an almost diametrically opposed take on the role of 'verbal reports' in the economy of the mind-body-thought dynamism to that of the Multiple Drafts and (neo)Homuncularism. That is to say it will be seen that the flowing connectivity of the stream of thought is the ultimate datum if not *hypokaimenon* of the lived psychological experience for William James, and any such thing as verbal artefacts are secondary or tertiary congelations or detritus formed in the passing of this primordial fact. This however is quite clearly a contrapositive to the Dennettean primacy of verbal artefacts and gestural pantomime in the economy of the Heterophenomenological method.

A sixth chapter will return to Dennett's neo-homuncularism, and begin to contrast this with what Antonio Damasio sees as the necessary and definitional wholism evidenced in all living conscious beings. Damasio contrasts the example of the Boeing 777 and its "engineered concern" for survival with our brains/minds and their "global concern for the integrity of our entire living real estate, every nook and cranny of it" (LS 128).

And this will lead into a seventh chapter with its a treatment of what Damasio calls the "'body-mindedness' of the mind". Even more powerfully, Damasio claims that primacy must be placed upon the body in this dyad—whereby the mind only exists "because there is a body to furnish it with contents" (LS 206). Now this is clearly a far cry from the picture of mind painted by Dennett where the perhaps aleatory Legos-like jamming together of "stupid" humunculi might or might not ultimately result in an agent of whose consciousness we can only be sure at the end of the day via a perusal of verbal reports. For Dennett these are generated in the form of "press releases" resulting from some Hobbesian struggle amongst the always fragmentary, detached, and unconnected "narratives" which are constantly racing around in the brain, vying for "fame". So a contrast of these two models of mind will constitute the eighth and last chapter. Where Dennett sums up his position (in "Are We Explaining Consciousness Yet?", August 2000) in a recent article for the journal *Cognition*: "The basic idea is that consciousness is more like fame than television"—it will be one of the main aims of this work to show that consciousness is really more like a Shakespearean play than either Hollywood, or television.

CHAPTER TWO: WILLIAM JAMES; THOUGHT AND THE BODY

Dr. William James, a professor of medicine before becoming a renowned American philosopher—aptly begins his two volume *magnum opus: The Principles of Psychology* with an in-depth analysis of the consciousness structure of that most common of anatomical laboratory ‘animals’, that most dissectable of creatures—the frog.

Beginning with a most basic sketch of the correlations of various functions to the differing ‘centers’ of this relatively simple vertebrate’s nervous system, James will build up from this simplistic bodily/neural/mental model a completely new way of looking at human consciousness—whose famous emblematic has come to be known in popular culture as the “stream of consciousness” though we need to be careful to distinguish that which this phrase has come to mean from the quite different though equally interesting “stream of thought” which Joyce borrowed and molded to his own ends.

So, to begin with, James relates a chain of experiments wherein various “sections” are made in ascending order—starting with a section between the top of the spinal cord and the medulla oblongata, “thereby cutting off the brain from all connection with the rest of the body”¹. While mostly completely locomotively and vocally inert—just lying on its back if flipped over, etc.—interestingly the frog still reacts quite vigorously to any sufficient irritation, for example rubbing its chest with both fore paws.

What James here is setting up is a chain of correlations between complexity of apparent choice in organism—environment interactivity and the ‘level’ of brain centers

¹ See Fig. 1 this page.

connected seamlessly to the whole organism. As James will sum up in concluding this first level of neural—bodily attachment (that consisting solely of the spinal cord in its connections to the body):

The most striking character of all these movements [those of the frog whose nervous system has been sectioned between the base of the medulla oblongata and the top of the spinal cord], after their teleological appropriateness, is their precision. They vary, in sensitive frogs and with a proper amount of irritation, so little as almost to resemble in their machine-like regularity the performances of a jumping-jack, whose legs must twitch whenever you pull the string. The spinal cord of the frog thus contains arrangements of cells and fibres fitted to convert skin irritations into movements of defence. (PP 15)

The next experiment involves a section between the optic lobes and the medulla oblongata and the cerebellum—the latter remaining thus attached to the spinal cord. The results of this section are much freer and interactive forms of locomotive response to the environment. Jumping, swimming and swallowing are added to the catalog of actions undertaken. The frog will turn itself right side up if thrown on its back, and will even engage in manoeuvres to compensate for random incident whole-body disturbances such as being spun around in a shallow dish suspended in a bowl of water. But the buck stops at the tilted board off which he or she will helplessly slide all the way up until the optic thalami are preserved in interconnection to all the lower complexes.

“When, finally, a frog’s cerebral hemispheres alone are cut off by a section between them and the [optical] thalami which preserves the latter, an unpractised observer would not at first suspect anything abnormal about the animal.” (PP17)

That is, for all intents and purposes the frog *reacts* just like any frog that had not been tampered with whatsoever. But therein lies the rub—the frog does not seem to be capable

of any actions which go beyond the purview of mere reaction to incident stimulation of direct or environmental nature, reacting to touch or swimming when placed in fluid etc.

Interestingly, he manifests neither signs of hunger nor of fear. As James puts it:

“In a word, he is an extremely complex machine whose actions, so far as they go, tend to self-preservation; but still a *machine*, in this sense—that it seems to contain no incalculable element. By applying the right sensory stimulus to him we are almost as certain of getting a fixed response as an organist is of hearing a certain tone when he pulls out a certain stop.” (PP 17)

However, all this changes according to James when we observe the activities of the frog with hemispheres intact. For James, the crucial difference is that we observe a notable admixture of spontaneity in the longer and more complex routines of locomotion in which the frog now engages. It is as if the frog is now motivated by some (external to the present situation) idea, and reactions to the same stimuli as used with the ‘sectioned’ frogs vary their forms—such the frog can no longer be seen to be merely a “jumping-jack” or machinic puppet reacting always predictably to any incident prodding.

So why the somewhat curious and lengthy discussion of frog neurology and consciousness, if you will, at the very beginning of a two tome work on psychology—presumably human psychology at that?!

I would say that it forms a generative account of sorts, working from a very basic neural schematic as backdrop to a consciousness formation. And perhaps along with the analyses of the neurology of dogs and even human babies that ensues, James is trying to make the strong statement that his psychology is going to be based very much in the physical, the biological and hence the scientific and empirically demonstrable as over against any one that will start with abstract precepts and work ‘backwards’ as it were

toward putative ‘empirical’ demonstration in the form of real world instances which happen to link up with the metaphysics presumed.

And perhaps this is a not unwise stratagem given the great potential that the charge of unfounded a priori be levelled against the theorization in chapter nine of the famous “stream of thought”. What I submit James wants to be absolutely sure of is that no-one mistakes this model for one pulled out of thin air in the great Kantian tradition, but rather instead understand that it is one based very much upon observable facts and “the phenomena themselves” as he iterates in chapter eleven. It is this very attention toward the phenomena themselves and the empirical basis thereof upon which his psychology is founded which leads me to label James a proto-phenomenologist (it should also be noted that James was completing the *Principles of Psychology* while Husserl was still grappling with the Kantian consequences of his *Logical Investigations*) which would also clearly theoretically mesh with his later status as one of the prime movers of the pragmatism movement in American philosophy.

However in the case of the frog experiments, there is clearly more than just this procedural point being made. To foreshadow my point—let me suffice it to say that James is laying out an argument by opposites for the intrinsic necessity (for the proper operation of consciousness) of the lower centers of the neural system and indeed the constituency of the organism as a whole, by showing first the clear insufficiency for a full, free consciousness of the interconnectivity of just the lower centers of the brain and the body. What will be needed most crucially ‘at first’ is the connection of the ‘cerebral hemispheres’ to the rest of the whole—these being the “seat of the memory” and that neural zone wherein the organism comes in virtual contact with things and situations and

modalities of reacting not necessarily ready to hand in the strictly present physically environing situation.

But what will become clear ‘later’ in the chain of argumentation, is the fact that just as insufficient for a full free consciousness would be even the seamless interconnection of all the centers of the brain without its interconnection to the spinal cord and all the bodily peripherals to which this connects. This will return again to the question of the cerebral hemispheres being the seat of memory—though it will become clear that even if they are the seat of the memory, where the memory truly *lives* is in the body and the body’s own sense that it owns itself and only itself and that its self is the same self that it has always had.

It is from this moment that James kicks off his indelible hypostatization that it is the very warmth and intimacy *felt* of the same body, always the same body, and all the time—even when this is not ‘consciously’, rationally even, cognized that forms the very kernel and core of the *sense* of self which further in true proto-phenomenological and a pre-pragmatic sense can be seen to be as close as we might come to a true definition of the ‘self’ in any case. As James puts it in “The Stream of Thought”:

Our own bodily position, attitude, condition is one of the things of which some awareness, however inattentive, invariably accompanies the knowledge of whatever else we know—We think, and as we think, we feel our bodily selves as the seat of the thinking. If the thinking be *our* thinking, it must be suffused through all its parts with that peculiar warmth and intimacy that make it come as ours. Whether the warmth and intimacy be anything more than the feeling of the same old body always there is a matter for the next chapter to decide. (PP 241)

And indeed whether it is anything more than the *feeling* of this same old body always there is quite duly decided in the ‘next chapter’ of *The Principles of Psychology*: “The

Consciousness of Self’. What we will see unfolding in this quite radical chain of reasoning of James’, at least for the turn of the last century, is the idea that the center, core, or very “kernel” in James’ words itself of the self, the seat of subjectivity, is the body itself, perhaps in a sense turning the old Cartesian hierarchy on its head.

As James notes in a key footnote mid-way through this chapter [X, Ten] (“The Consciousness of Self”):

I and *thou*, I and *it*, are distinctions exactly on a par with these,—distinctions possible in an exclusively *objective* field of knowledge, the ‘I’ meaning for the Thought nothing but the bodily life which it momentarily feels. The sense of my bodily existence, however obscurely recognized as such, *may* then be the absolute original of my conscious selfhood, the fundamental perception that *I am*. All appropriations *may* be made to it, *by* a Thought not at the moment immediately cognized by itself. [Thought capitalized in the original] (PP 341)

Again, the primacy is here being laid upon the bodily—and it is to that which the thoughts themselves are stitched. The ever present existence of the body, its constant sensuous enframing, for James is “the absolute original of my conscious selfhood”. Far from knowing that we are, that we exist, because at least we can be certain that we are thinking, and of tantamount importance that we are in essence thinking things—with James this notion is turned on its head, and we are only certain that we are, and that we have a self, a self that endures over time, because we can feel at the root of all our thoughts a bodily warmth presence and intimacy as the anchor of all.

Or as Bergson says in a slightly different context: the body is that which gives the mind ballast and poise. Yet clearly James goes further than attributing a centering influence on mind of the bodily. For James, the sensorial and bodily is the *essential* root and anchor to which all other phenomena, including all mentation, are attached. Without

this anchor, James implies in *The Principles of Psychology* citing the case of a young man who completely loses all sense of self when he closes the one eye with which he can see² (being completely anaesthetic except for the sight of one eye and the hearing of one ear), all personal consciousness disappears.

Now, I will submit that along with “personal” consciousness goes consciousness itself as well. And perhaps this is a fairly self-evident point to most. However, it will be a central theme of contention throughout this work, as will become more and more clear as we delve in the following sections into the theses of one of the most famous of the proponents of the homuncular, or neural automata, models of mind. That is, those holding to the homuncular model will say that there is consciousness and there is personal consciousness, and not only do the two not have to be coincident, but that the one (“consciousness”, eviscerated of the flimsy category “personal”) is more than enough for us to construct empirically correct models of mind. And in fact, as soon as we introduce anything smacking of the purely “personal” any claims to scientific propriety in our description of the consciousness experience must necessarily go out the door.

But to return to our treatment of the bodily component, at the very least, of consciousness—I would like to address the issue of identity, as this is a crucial underpinning of what will become a full-fledged theory of the stream of thought for James. However, once again, we will see that for the homuncularists, anything so vague as ‘personal identity’ is going to be one of the somewhat fictional categories of selfhood that will only exist after the fact, after the stream of thought or consciousness has been frozen into rigid bits of ‘solid meaning’ such as words or publicly observable deeds.

² The boy said: “Wenn ich nicht sehen kann, da BIN ich gar nicht.”

Seen in the most positive of lights, this might be read as a kind of neo-Nietzschean constructivism and perspectivism. Seen in a more critical light, this might be read as a sort of attempt to replace subjects and their subjectivities with “little robots” (automata, neural nets, homunculi) and the publicly observable data they eventually spit out in the form of verbal artefacts and gestural behaviour.

But I will save my full on Jamesian critique of the use of verbal reports as ultimate touchstone for consciousness, and a deeper discussion of how this flies directly in the face of the Jamesian stream of thought for chapter 5 below. For now, I would simply like to shore up the importance of the bodily warmth and intimacy felt behind every living thought as anchorage for any further treatment of the operation of thought in the human and a comparison of this with the operation of thought in robots and other entities whose consciousness is controlled solely through the agency of “little robots” or automata, neural firings, homunculi.

From the same section of the “The Consciousness of Self” chapter quoted just above from *The Principles of Psychology*, James drives home the aforementioned point that the core of any higher thought structure and the abgrund from which any representations might spring is precisely this always sensed, however vaguely, intimate and warm kernel of bodily existence. He will even submit, as we will see, that even if a thinking process is not directly aware of itself at any given moment, in other words not presently engaged in any kind of self-reflexive activity, nonetheless there will subsist beneath it all a warm and intimate sense of the same old bodily self always there grounding whatever might eventually spring to mind.

Speaking of any present stream of thought, with specific regard to its current anchorages, James says:

Its appropriations are therefore less to itself than to the most intimately felt part of its present Object, the body, and the central adjustments, which accompany the act of thinking, in the head. These are the real nucleus of our personal identity, and it is their actual existence, realized as solid present fact, which makes us say ‘as sure as I exist, those past facts were a part of myself.’ They are the kernel to which the represented parts of the Self are assimilated, accreted, and knit on; and even were Thought entirely unconscious of itself in the act of thinking, these ‘warm’ parts of its present object would be a firm basis on which the consciousness of personal identity would rest. (PP 341)

Now this talk of the stream of thought’s present “Object” with a capital “O” will be seen to be rather reminiscent of the architectonics of mind charted out by Antonio Damasio, which we will revisit in chapter 7 in more detail. Damasio, taking his departure from Spinoza’s: “The Object of the Idea constituting the human Mind is the Body” (Proposition 13, Part II, *The Ethics*) will discuss at length the “body-mindedness” of the mind, and the concomitant necessity for the smooth operation of consciousness of a distinct sense of self grounded in the constant formation of “body images” in the mind.

Without these, Damasio claims, the organism cannot perform the massive integration of information crucial for survival in the timely fashion required. In fact, with the temporary suspension of the “self-component” of consciousness is entailed a complete disruption of our life-management functions, which will return us “to a state of dependence comparable to that of a toddler.” (LS 207) This happens, Damasio claims, especially clearly when in cases such as akinetic mutism—much like the case of the boy with only the sight of one eye and the hearing of one ear mentioned by James above.

And so for James, the ‘central adjustments’ in the head, and that most intimately felt present object, the body, are the “real nucleus” of our identities. And further, these are the solid present facts, whose existence, and role in the composition of ourselves and our selves is as sure, as certain, as our very existence itself. Moreover, these form the very *kernel* to which all secondary re-presentations are attached or as James says accreted, assimilated, even “knit on”. But even more crucially—and we will see where this will fly directly in the face of Dennettian materio-homuncularism and verificationist (Dennett will call it “teleo-functionalism”) verbal reportism in the next three chapters—this obtains for James even in cases where the present stream of thought is not entirely conscious of itself in the act of thinking. Still, James proclaims, these “warm” parts of its present object, this object being for James as for Spinoza the body, will form the “firm basis” upon which all later personal consciousness, and identity, will come to rest.

So let us delve a little further into this fascinating and recurring theme of “warmth” and “intimacy” which James imputes to the stream of thought, even when unaware of itself, that is, even at moments when self-reflexiveness is at an ebb.

As James states in “The Consciousness of Self”:

We spoke a moment since of warmth and intimacy. This leads us to the answer sought. For, whatever the thought we are criticizing may think about its present self, that self comes to its acquaintance, or is actually felt, with warmth and intimacy. Of course, that is the case with the bodily part of it; we feel the whole cubic mass of our body all the while, it gives an unceasing sense of personal existence. Equally do we feel the inner ‘nucleus of the spiritual self,’ either in the shape of yon faint psychological adjustments, or (adopting the universal psychological belief), in that of the pure activity of our thought taking place as such. Our remoter spiritual, material, and social selves, so far as they are realized, come also with a glow and a warmth; for the thought of them invariably brings some degree of organic emotion in the shape of quickened heart-beats, oppressed breathing, or some other alteration. (PP 333)

And so here again we have the “whole cubic mass of our body” unceasingly making us aware that we exist, that we are alive, and as per the paragraphs above, that our identity is ours alone, and fixed. We will get into more detail as to the surety of ownership of our thoughts just below, and in specific reference to the stream of thought in chapter 7, where we revisit the bodily, and then begin the journey of examining the role of words in the Jamesian stream of thought and alternately in the Dennettian Multiple Drafts model of mind and its vouchsafe in his Heterophenomenological method.

But for now, suffice it to say that the bodily is not only something James wants to brush aside, but indeed that which lies at the core and acts as the primary anchor to any later consciousnesses and sense of self.

Obviously this warmth and intimacy of which James speaks is naturally going to be something only living, organic thinking agents are going to be capable of feeling, exhibiting, and having. Therefore, at this point it would seem to constitute a bit of a cheap shot with regard to the debate concerning living, organically wholistic thinking agents over against artificial thinking agents such as AI systems and the various modern robots that implement such in attempts to simulate embodiment situations. And it is in no wise my intention to effect such. I am merely setting up the conditions based upon which the original theory concerning the “stream of thought” as it envisioned by William James was constructed, and merely intend for this to serve as a groundwork for later treatment of the stream of thought as it will come into conflict with verbal reports as touchstone of consciousness, and finally as background for the very up-to-date theorization of the role of body images in the operation of mind undertaken by Antonio Damasio.

But to return to the long quote above for just a moment, I believe it is illustrative as to how central this very bodily notion of warm and intimacy as a concomitant to any stream of thought is for James that it will accompany, in the form of “organic emotion” even those more ‘distant selves’, even the spiritual and social selves come with a “glow and a warmth”. And what this is implying is of course the tenet that there is a recursive architecture at play whereby even “the pure activity of our thought taking place as such”, read mind in operation, can effect a global change in the over-all body state.

Therefore, without a fully circular and wholistic economy obtaining between the so-called pure activity of thought operating on its own (whose purity is hereby of course always tinged with the bodily thanks to this very wholism) and the bodily substrate which underwrites as it were all activity of thought with its warm and intimate grounding, there is indeed no such thing as a stream of thought in the Jamesean sense.

In fact, James will go on to submit that it will be the very clinging aroma, the “animal warmth” which attaches itself to thoughts that are or were particularly ours, and therefore thoughts which play a large part in the constitution of what it is to *be* us, which will be his ultimate touchstone in the game of deciding what counts as consciousness. In this case James is placing especial emphasis on the ‘ownership’ of the thoughts or stream of thought, though this obviously reduces to consciousness itself. What is perhaps of some interest in this regard however is the fact that taken from an ‘ownership’ perspective, the Dennettian neo-homuncular and Multiple Drafts scenarios do not at all reduce personal streams of thought to the stream of thought itself or consciousness necessarily self-evident. Quite far from discussing any “animal warmth” or felt intimacy lying at the back and the base of any consciousness experience—that which safeguards

any stream of thought or consciousness experience's entry into the realm of really being a consciousness experience is the "globally accessible" verbal artefact which can be reviewed by all, and said to be a consciousness experience. This situation even obtains in the case of a rather personal consciousness experience, say one undergone while sitting alone in our rocking chair looking out the window—as we will see below in Dennett's discussion of his Heterophenomenological method (for deciding consciousness).

I will end this chapter with the words of James himself, once again from *The Principles of Psychology*, where with a faint tinge of Sergio Leone he paints a rather picaresque tableau of how we come to know that our thoughts are our own, and that we are indeed imbued with consciousness—not dreaming others' dreams and not sleeping in others' bodies:

Any distant self which fulfils this condition will be thought with such warmth and intimacy. But which distant selves do fulfil the condition, when represented? Obviously those, and only those, which fulfilled it when they were alive. Then we shall imagine with the animal warmth upon them, to them may possibly cling the aroma, the echo of the thinking taken in the act. And by a natural consequence, we shall assimilate them to each other and to the warm and intimate self we now feel within us as we think, and separate them as a collection from whatever selves have not this mark, much as out of a herd of cattle let loose for the winter on some wide Western prairie the owner picks out and sorts together when the time for the round-up comes in the spring, all the beasts upon which he finds his own particular brand. (PP 334)

CHAPTER THREE: DANIEL DENNETT AND THE MULTIPLE DRAFTS

On the “human” side of these cyborg, homuncular, debates I will focus on Daniel Dennett’s grand theoretical construct offered in the pages of *Consciousness Explained*, his *magnum opus* (to date) on consciousness. This is what Dennett has called the Multiple Drafts theory, and it attempts at one stroke to brush aside any residues of Cartesianism still haunting modern materialism, and at the same time to resolve some rather quirky if not paradoxical experimental findings on the nature of the temporality of consciousness and its correlations to particular perceptual and reflex reactions to the outer world. And it is here in the formulation of the Multiple Drafts theory that we can see what I am going to be calling “cyborg temporality” encroaching on perhaps not just our theoretical understanding of the temporalities of human consciousness but perhaps as well the actual really lived temporalities thereof.

In a word, the Multiple Drafts theory, as John Searle has claimed in another context³, while putatively aimed at getting rid of the Cartesian homunculus and the ultimate spatial approximation of consciousness in the grey matter—instead could be seen as really aiming at getting rid of our standard notion of “consciousness” itself and only by-the-way serving as a critique of the old-fashioned Cartesianism which would dare to pin-point the exact *place* where consciousness occurs in the human mind. What is effected is a rather complicated sleight-of-hand whereby a dogged Cartesian materialism is replaced by a hyper-materialism; only one which is glossed over thanks to the auspices

³ John Searle, *The Mystery of Consciousness*, p. 101.

of the Multiple Drafts which in its very slipperiness promises to get us past the old *homunculus* but quite self-admittedly only ends up multiplying and shrinking the poor old beast *ad infinitum*. And perhaps its just this *infinitum* that confusingly seems to promise the smooth curve which strings together all the underlying points in a mistaken calculus of metaphoricity whereby it's forgotten that all Dennett is giving us is the points ("bits of dumb code", unconscious "sub-processes" etc.) and pretends to give us the 'curve' of the stream of thought only through the blurriness of the Multiple Drafts model. But as Searle points out, the insidious plot behind all these complicated metaphoricity is really aimed toward getting rid of the phenomenality of the curve, of the stream of consciousness itself which is ultimately effected a posteriori thanks to the Multiple Drafts' sister theory of "Heterophenomenology" which in somewhat putting the cart before the horse says that the only way to come to know of the stream of thought really is via the verbal (whether uttered or unuttered) report which follows upon whatever neural processes have been occurring in the Multiple Drafts scenario. And hence through a simple *elenchus* since we cannot in any way "know" of this stream of consciousness, we don't have any right to say it exists. As seems to be often the case with Dennett, he therefore seems to get to have his cake *and* eat it when it comes to the employment of the concept of the stream of thought or consciousness.

Now, it should be clear by this point that for my central theme of the transcendence of human conscious temporality and conversely the temporalities of human transcendence an unambiguous understanding of the not only phenomenal reality of the stream of thought as an organic unity (for this could easily fall prey to a Dennettean critique based upon the *aporia* of delusional and pathological cases as well as

standard illusions to which we all fall prey from time to time⁴) but also the essential ontological (in the spirit of, again, Searle's 'subjective ontology') metaphysical, scientific and empirically provable reality of the stream of thought will become necessary. Hence the goal of this essay is not so much to say that the Multiple Drafts model is completely wrong—but much more so to recuperate the stream of thought from its interstices, and if this turns out to be impossible without dismantling the Multiple Drafts, then so be it.

Another way of looking at the analyses of this chapter is as an investigation into the micro-scale of the temporalities of the stream of thought, and one angle of attack on its wholism by cyborg thought and the machine like models of consciousness as espoused by the modern technophilic priesthood.

So, after hopefully having recuperated the stream of thought in its wholism and integrity from the Multiple Drafts (and its ilk) in this section, I will go on in the next section to a treatment of the macro-scale of the temporalities of the stream of thought in consciousness whereby I will focus on the converse of that importation of cyborg thinking into human thought that we will have seen in this section—that converse being the imputation of strict humanity to cyborg thought itself, and in particular in one of its star avatars to date: the COG robot project underway at the MIT laboratories in Rochester, New York.

This will introduce us to a discussion of the emotive and feeling structures underlying 'higher' more rational forms of thought evidenced in the human creature, naturally bringing us back to the doorstep of Antonio Damasio who, in his ground-breaking *The*

⁴ A perfect example of which being that exposed by the "color phi experiment" upon which I will be dwelling at length below for obvious parenthetical reasons.

Feeling of What Happens (1999) formulates a model of consciousness which can be seen as a seamless collation of the Bergsonian imagistic metaphysics of Action with the Jamesian theory of the stream of thought.

Dennett time and again prides himself on being one of the few to go into the “microdetails” of the consciousness experience as it occurs on “microscopic” [sic] time scales. Further, it is in the zone of these very experiments which attempt to dissect the temporality of consciousness that Dennett admits will come any “scientific” challenge to his theory (of the Multiple Drafts)—and hence have I chosen to focus on a furtherance of Dennett’s exploration of the classic “color *phi*” experiment with respect to how well the Multiple Drafts theory handles the data, and with an eye toward potential alternatives which may offer even better explanations of the data. So on this front, I will begin with a basic outline of the Multiple Drafts theory.

But before I get involved in the meat of the discussion of the “human” side of the ‘cyborg debates’ and the rest of this chapter, let me briefly foreshadow where we will eventually be going in the next chapter—the “digital” side of the cyborg debate as, as I’ve mentioned above, this is a very Janus-faced issue as enunciated by its computer and cognitive science proponents. In chapter four I will be arguing that while advances in neural network modeling and ever increasing computational capacities will enable the effectuation of better and better *simulations* of human conscious activities—there will always be certain essential components of the human consciousness experience that will not be precisely *reproducible* in the digital media in no matter what advanced avatars might be visible on the horizon. These essential components are the following: (a) pain and its concomitant—the survival instinct, (b) the organic interconnection and

interdependence (on the cellular level) of the human being, and (c) the homogenous interpenetration of the ‘stream of consciousness’ as experienced in the human—which I take to be intrinsically related to, and an evolution of (a) and (b).

So, to begin with, allow me to introduce the Multiple Drafts theory as presented by Dennett in chapters five and six of *Consciousness Explained*. As I mention in the introduction, one of the main goals of elucidating this theory for Dennett is to assure himself and us that the final nail has indeed been driven into the coffin of Cartesianism. According to Dennett old-fashioned, ‘barefoot’ Cartesianism is all but dead to any serious theoretical and scientifically minded debate. This is the form of Cartesianism which says that there are two distinctly separate “substances” in the universe—on the one hand the physical, that which obeys the laws of physics and is scientifically observable thereof and therefore, and on the other the mental—which operates in its own autonomous zone which has nothing to do with the ‘material world’ and hence needn’t operate according to any of its rules nor necessarily be observable in any scientific sense. However, aside from the lingering crack-pots who might still believe in good old-fashioned Cartesianism—there also remain, even in the reasonable and scientifically minded community of theorists those who (whether they know it or not) still hold on to a sort of transmogrified, or at least modernized version of Cartesianism—and this school of thought is what Dennett labels Cartesian materialism.

In Cartesian materialism, Dennett claims, the theorist who may delude himself into thinking that he has avoided the Cartesian trap—finds himself right back in it, though at least in this case in the improved position of being trapped entirely on the material side of the debate, safely ensconced in a sound materialism. And how this

Cartesian materialism plays itself out on the level of brain processes and their theorization is in the wrong-headed lingering doctrine that yes, even though we have admitted that all there is to conscious occurrences in the human brain is that of physically describable processes occurring in the physical material of the brain—however, and this is the part that rubs Dennett wrong—there must be therefore a specific *place* in the brain where this consciousness process is happening, for otherwise, if it be at no place—then how can we say that what is taking place is taking place strictly on the material plane? Thus for Dennett, anyone beholden to this theory is merely playing a shell game with Cartesianism—and replacing the ineffable “mental” and its divine driver the “soul” (with its go-between the pineal gland) with an equally ineffable “Cartesian Theater” which, though at least theoretically ‘materially’ situated in the brain—is just as mythical as the classic Cartesian format, and just as ineffective at offering real explanatory powers for what is going on in human consciousness.

So, over against this Cartesian Theater paradigm of consciousness to which many modern theorists of mind have fallen prey—Dennett offers the Multiple Drafts model. Starting at the very beginning (chapter five, section two) of his elucidation of how this theory works—Dennett presents one of these experimental (apparent) paradoxes whereby the timeline of consciousness and the timeline of the physicist do not seem to mesh up. Rather than cave in to the paradoxicality of the data and either (of course not!) give up on a strict materialistic position, or allow the addition of hypothesized (and according to the very Occam-like Dennett rather unnecessary) sub-faculties of the consciousness process which Dennett dubs Orwellian and Stalinesque style revisionism (more on this directly below)—Dennett says that the foundational problem in the interpretation of these

experimental data has been the implicit but never expressed as such Cartesian materialism which has made the theorists interpreting these experiments rely upon the conception of (and hence they always look for) a Cartesian Theater in which the consciousness events must be occurring, and further—a sort of Cartesian Time (though Dennett never calls it such) at which or at least within which a consciousness event must take place. Wrong, wrong, wrong!—says Dennett. There is no Cartesian Theater, no privileged space within which consciousness occurs; there is no Cartesian Time (frame) within which envelope the consciousness event must occur. Rather than seek for these privileged times and places—we should realize that at base all that is going on is a “sort of multi-track process” whereby certain “editorial functions” are being undergone, “revised”, and then either submitted or not submitted as the ‘final draft’—the ultimate mark, for Dennett, of consciousness. However, since there are none but these “multi-track” processes constantly occurring in the brain—even this ‘final draft’ is a bit of a chimera, that is, until nailed down finally in the form of a (for Dennett it would appear always *verbal*) report. But I have gotten a bit ahead of myself, so let me follow Dennett’s lead and start from the beginning: the *phi* phenomenon, first scientifically and systematically elucidated by Max Wertheimer in 1912.

In its most basic form, the *phi* phenomenon involves the illusion of apparently smooth, homogenous motions created in the mind of the perceiver by what are in actual fact a rapid succession of momentary flashing or “still” images: something which we are all well aware of in its application in the cinema, and at a higher rate of succession, on the television. Taking the phenomenon to its barest form—Wertheimer simply lit two small spots at about four degrees of angular separation, and studied the effects on the subjects

of his experiments. What people invariably saw was a single spot fluidly moving back and forth between two extremes—which extremes were in fact the only truly existent spots. Then, more interestingly, in 1976, spurred on by philosopher Nelson Goodman’s query as to what might happen if two differently colored spots were moved, the psychologists Paul Kohlers and Michael von Grunau illuminated a red and a green spot of light for 150 milliseconds each, with a 50 millisecond interval between illuminations. What did the subjects experience? They saw a red spot begin to move to the right, and then abruptly, in mid-trajectory, turn into a green spot which proceeded to complete the four degrees of angular motion.

So the conundrum that clearly arises is: how is it that the mind seems to “know” ahead of time that a green spot is going to succeed the red spot—or, put the other way around—how can the stimulus of the red spot be seemingly effecting the conscious experience of a traveling spot even before that stimulus can putatively have reached the brain and traveled far enough along the consciousness producing process to affect the “movie in the mind” when it necessarily must be succeeding the primary stimulus by at least the 50 millisecond interval mentioned above. Well, says Dennett—the whole problem with this conundrum is not that a seeming temporal paradox must be overcome or explained away, but that there is no temporal paradox to begin with—only bad framing assumptions. And these are, first and foremost, that there must be some central (even if not spatially central, at least logically central) seat of awareness-production in the brain that would count as a ‘finish line’ the arrival at which of various mental processings would constitute the final conscious experience. Further, and as outgrowths of this central misbegotten tenet—theorists are forced into a double bind whereby certain ancillary sub-

processes *must* be going on in the brain such that the experimental data can all hang together.

And this is where Dennett brandishes his Occam's razor and says that these processes, which he calls *Orwellian* and *Stalinesque* revision—not only need not be hypothesized as a necessary corrective for the experimental data, but that the very processes themselves are mere will-o'-the-wisps arising from the general mists laying about the doctrine of the Cartesian Theater. Allow me to go into more detail concerning this Orwellian and Stalinesque revision as Dennett describes them.

What Dennett is claiming is that if there is indeed some such zone, some black box called consciousness however constructed with respect to its boundaries in time and space—as long as its borders are such that an item in the mind, a process in the brain, is either “outside” of or “inside” of it—then we are left with the false dilemma of describing the seemingly paradoxical data of the color *phi* experiment by the one or the other of these Orwellian or Stalinesque revisions. In short, the Orwellian revision would work something like: the red dot registers on the consciousness zone, a little later the green dot registers on the consciousness zone—and then, for some reason of efficacy one would imagine, the two dots are *doctored up in the “memory”* such that what is ultimately experienced as the stream of consciousness is a red dot moving right, drastically turning into a green dot and then fluidly shifting over to the final resting place which is the actual position of the green dot in physical space.

Alternatively, in Dennett's Stalinesque revisionism—some sort of ‘delay loop’ is hypothesized before the window or zone of consciousness and betwixt that and the stimulus, during which the alterations to the actual input data are made—and that is

where the two dots are melded into one seamless flight. Now, to step back and look at the big picture, what Dennett is saying is that it is much less unlikely a scenario that there is in fact no such zone or window (with definite physical boundaries, with definite temporal boundaries) of consciousness—than that such bizarre alterations are occurring either before or after any such zone. And the elegant solution to this problem according to Dennett is the Multiple Drafts theory. According to his 1995 recapitulation of the multiple drafts theory in an essay titled “Is Perception the ‘Leading Edge’ of Memory?”:

Now which theory [Orwellian/Stalinesque] might be the truth? The Multiple Drafts theory declares that neither one is the truth. The truth is that the brain is quite capable of putting retrospective content elements into it’s [sic] narrative stream. It can decide there’s a circle on the left and it’s red and there’s a circle on the right and it’s green, and that there must have been change in between them. But this natural but mistaken conclusion [i.e. that there is actual motion between the spots] is “predated”: it is given a “postmark” which places it at an earlier time in the sequence in your own stream of consciousness. Now this is an idea that many people find extremely hard to accept because it suggest to them that there must be some sort of backwards causation in time “ (PLEM 75)

Now, insofar as Dennett is ridding our models of consciousness of any determinate physical Cartesian space, theater *in* which consciousness must or must not occur—we are very happy for the insight. Insofar as there are “multi-track processes” occurring at distributed times and distributed places around the brain that all combine to create one whole “stream of consciousness” we are also very pleased with the model. Where things get sticky, however, is when Dennett attempts to add the concepts of

“predating” and “postmarking” to the temporal aspect of this multi-track process. And to further cloud the waters—in defense of this very “postmarking” Dennett [in *very* characteristic fashion] relies upon a most unfortunate metaphor—that of the refracting effects of the periscope upon vision, this being clearly a very *spatial* phenomenon. And what is not objectionable as such is not the spatiality of the metaphor as such—but rather the fact that Dennett claims that all of the parameters of the periscopic situation equally well apply to the temporal refractive situation in the brain, in a seamless parallelism [and we will find many more of these seemingly seamless parallelisms below!]. So, let me go into a little more detail, as indeed Dennett would always encourage according to his own rhetoric, and see if we can’t unravel what I see as the distinct non-parallelism of the spatial periscopic metaphor to the temporal refractive situation which occurs in the translation between the time of the physicist and the temporality of the stream of consciousness. To begin, the framing of Dennett himself:

In an exactly parallel way, phenomena such as the *phi* phenomenon . . . show that the brain can create what we might call temporal periscopes, curious occasions when time itself is apparently bent by the way the brain deals with the events falling on it. (PLEM 77)

Now, while it is quite clear why the image of an elephant in the mind, in the brain should not and could not exactly correspond to the spatiality of the elephant itself—and that the spatial perspective effected by the use of a periscope could so warp sense of place—there seems to me to be something of a category shift involved in discussing Dennett’s “temporal periscopes”. And the problematics involved in such a paradigm shift veiled in the cloak of complete parallelism appear to me to be of two orders: A) Dennett’s

insistence on the discrimination of “time represented” from “time of representing” (with its alternate formulation “that a judgment of simultaneity need not itself be simultaneously arrived at or rendered; it might mature over a long period of time . . . “, i.e. do the architectonics of *representing* over against *represented* operate under the same logical paradigms in *both* spatial and temporal relations? And B) Dennett’s alternative use and non-use of the “stream of consciousness” figure to bolster his own proper argument when and only when it matches his own ends.

As we shall see, this second consideration (part B above) will read back through and dictate the outcome of the first (part A above)—in that it will embroil Dennett in an infinite shell game consisting of constant shifts between the figure of a “stream of consciousness”, and its more ‘scientific’ counterpart which he seems to label for lack of more strict terminology or perhaps conceptualization: “subsequent processes in the brain”, or later on, “center of narrative gravity”.

So, to begin with, let me unravel Dennett’s ‘temporal periscope’ effect, and try to explicate how it fits into his idea that most theorists of mind ‘have not understood the necessity’ of distinguishing between “time of representing” and “time represented”. As Dennett has it, we need to make a sharp distinction between the “when” of the occurrence of the processing event of any given cortical sub-process (such as the detection of motion, recognition of a certain color, luminosity, direction of travel etc.) and the “when” of its ultimate “temporal content” in what will be the ‘stream of consciousness’. Now, what he means by temporal content is that the time of an event as (ultimately) *represented* in what will have been the outcome of the combined cortical processes—need not be a direct reflection of when its *representing* originally occurred within the

pell-mell of ongoing, multifarious cortical sub-processes, but rather a reflection of the “when” which is assigned to it by the (seemingly more ‘intelligent’—as over against the, as he calls them time and again, relatively “stupid” sub-processes) aggregate of subsequent sub-processes which ultimately come to be the ‘stream of consciousness’, though for Dennett—only once they have been translated into some kind of “verbal” report or other⁵. In short—the time at which any certain brain processing-event occurs need not be the same in relation to other surrounding, relevant processings that will ultimately obtain as it seems to the person in which these processes are occurring.

And the trick here, as I mention above—is that there really isn’t any such thing occurring in the brain as what is traditionally thought of as the “stream of consciousness”. Otherwise, as we will see below when I dive back into the color *phi* phenomenon experiment in order to clarify all this abstraction—there would be distinct problems for the attempt to hold together a notion of “stream of consciousness”, materialism, and the results of the color *phi* experiment at one time. So, instead of a “real” stream of consciousness really occurring in the brain at every moment—as most other theorists would posit if bothering to use the term, for Dennett there is merely the illusion of such which arises only when a subject attempts to formulate in semantic terms to themselves or to others what they may have recently or distantly undergone. Hence does Dennett dodge any seeming temporal paradoxicality with respect to the stream of consciousness and/or any causal wackiness that might be going on when certain bits of stimuli/information seem to be showing up in the ‘stream of consciousness’ basically before they should—were it to be an ongoing phenomenon tied to the present *and* the

⁵ For Dennett, a “verbal” report can be any articulation with demonstrable and relatively unambiguously understandable semantic content.

time of the physicist *and* the time of all ongoing cortical sub-processes, as it must were it to be really occurring without any strange delay loops or memory tamperings. So, once again, Dennett's 'eloquent' solution: get rid of it! However, and as John Searle is very apt to point out, he retains the terminology! (but, as Searle is also unabashed to claim, perhaps only to further confuse a matter that at heart remains confused!)

Before returning to the color *phi* experiment itself and its specific data for clarification of these issues, and hence my own hypotheses concerning what might actually be going on—I would like to take a short detour into the question of Dennett's employment of metaphors as a purported way of 'illustrating' and 'shoring up' theories. What I'm going to submit is that this 'illustration' motif claimed by Dennett throughout his work is really just a clever way of disguising the fact that there really is no neurophysiological or any other kind of proof for the theory he claims. Now, this is just good old fashioned cognitive scientific methodology provided there can be demonstrated a logical necessity that embedded brain processes *must* be operating just like such-and-such. However, in most cases—and particularly in the case of the Multiple Drafts theory, this does not seem to be the case at all. As is quite quizzically demonstrated by Dennett's more than adequate elucidation of the competing hypotheses. So, let me go into a little detail on just a few of the fusillade of metaphors that Dennett parades forth in defense of his most "rigorous" and "difficult" theory presented in the pages of *Consciousness Explained*

What matters for the brain is not necessarily when individual representing events happen in various parts of the brain (as long as they happen in time to control the things that need controlling!) but their *temporal content*. That is, what matters is that the brain can proceed to control events "under

the assumption that A happened before B” whether or not the information that A has happened enters the relevant system of the brain and gets recognized as such before the information that B has happened. (Recall the commander in chief in Calcutta: First he was informed of the battle, then he was informed of the truce, but since he can extract from this information that the truce came first, he can act accordingly.) (CE149)

And here we have the major thesis of temporal content being the prime mover of temporal awareness in the brain, as over against the temporalities of “individual representing events” (the physical timing of sub-processes in the cortex) essentially couched between two main metaphors—That of the Battle of New Orleans on January 8, 1815 [and here one may presume that the exact dating and naming can only add to the scientific plausibility of the move!] and after the metaphor of the starting gun at a sailboat race.

The battle from the War of 1812 is utilized to illustrate the fact that an imaginary commander of British forces in colonial India, unlike the poor commander in Belgium who sent his forces to their demise at the Battle of New Orleans fifteen days after the truce was signed—thanks to the postmarking of the letters received relaying the news of the end of the War and the battle can make the proper discriminations of order of events. Specifically, Dennett says that even if news of the battle arrives on day 35 (after the signing of the treaty) and news of the treaty on day 40 (due to some aleatorily hypothesized slow overland route!)—thanks to the convention of postmarking of dates on letters, the commander in Calcutta can still make out which real world event occurred before which. This, most importantly, for Dennett illustrates the vehicle/content distinction in communication theory. Again, the *time represented* (by the postmark/*content*) and the *time of representing* (when the letter arrives/*vehicle*) need not

be the same. Now, Dennett readily admits that no such exact ‘postmarking mechanism’ appears to be available to “the brain’s communicators”. Quizzically however, “the general principle of the content/vehicle distinction is relevant to the information-processing models of the brain . . .” (CE 147)

And the second metaphor in this most rigorous of chapters (again, chapter 6 of *Consciousness Explained*) is that of the aforementioned starting gun at the sailboat race. Much like in the War of 1812 metaphor, what Dennett is essentially trying to illustrate is that its not so important for the observer, for the brain, what the exact “when” of an input is, as what the “when” turns ultimately out to be for final perusal once it has been *represented*: once again the *representing* ceding place to the *represented* of time. Specifically in the sailboat race: although one may clearly ‘see’ a sailboat cross the starting line ‘before’ the starting gun has sounded: thanks to the differential in transmission times of light waves and sound waves—one can never really be sure whether it did or did not illegally cross the starting line ‘too early’ until some privileged judge can properly make the discrimination, properly making the calculations and then “leisurely” conveying this information to the interested participants. For Dennett this judge is the committee boat, or, “some circuit in the brain” (and he really does seem to *equivocate* the two):

So timing of some representings matters *until* a discrimination such as left-to-right (or over the line early) has been made, but once it is made, locally, by some circuit in the cortex (or some observer on the committee boat), the content of the judgment can be sent, in a temporally sloppy way, anywhere in the brain where this information might be put to use. (CE 150)

Now, beyond the generally slipshod methodology of shoring up operational models of brain/mind with metaphors, and only metaphors—we must admit that Dennett doesn't really rely on this 'postmarking' theory to make the temporality of his Multiple Drafts model work—rather he claims there is a “less foolproof but biologically more plausible way”, this which he calls: “*content-sensitive settling*. A useful analogy would be” And, here we go again! Next we get a metaphor of the synchronization of filmic images with their related sound tracks in film editing. Again:

Moreover, the temporal order of the stages of the process of putting the pieces into registration is independent of the content of the product; the editor can organize scene three before organizing scene two, and in principle could even do the entire job running the segments “in reverse”. (CE152)

The important apparent difference from and addition to the sailboat race and War of 1812 metaphors here is that there is no longer need of any 'postmarking' phenomenon, as the contents of the signals, the information itself will dictate how to properly correlate the different channels of inputs. And, voila!, here we finally have the explication of how this dirty work is actually accomplished in the brain, as presumably:

Quite “stupid” processes can do similar jiggling and settling in the brain. For instance, the computation of depth in random-dot stereograms is spatial problem for which we can readily envisage temporal analogues⁶. (CE152)

⁶ Now these “stupid” processes do not only figure here in the question of temporal registration in the brain—but as well in a much more broad manner with regard to the function of the brain and consciousness in general. Particularly, as we will claim in the second part of this essay—these “stupid” processes will provide a segue for Dennett's ultimate claim that the parallel wiring of a bunch of Mac-II's will provide the sufficient platform for his Cog robot to eventually attain a human-like consciousness.

Now, as I have mentioned above—it is not the over-all concept of Multiple Drafts with which I have no truck, rather, it is the fact that it would appear to throw the baby out with the bath water: instead of just getting rid of the Cartesian Theater (which project I wholeheartedly embrace), at the same stroke Dennett would get rid of the stream of consciousness itself. All that he leaves us with is a vestige of such, a phantom which is reconstructed in retrospect using either postmarking or content-settling of subsequent (always subsequent, which seems to pose a bit of a problem here) cortical sub-processes. But once the temporal non-correspondence of the stream of consciousness to physical processes in the brain has been established (as in the color *phi* experiment), once “temporal kinks” occur, as Dennett calls them, in the logical mapping of the stream of our thought onto the measurable time-line of cortical processing-events—then we are left in the explanatory dilemma of either adhering to some form of anti-materialist conjecture that there must be some kind of disjunction between the physical processing going on in the brain and the over-all process as we can “see” it in the stream of thought, or, there is some kind of shell game going on in the brain whereby the brain is duping itself in effect by either altering inputs, delaying them somehow (Dennett’s famous Stalinesque revision) or alternatively editing the outputs of processing events before we ‘ourselves’ become aware of them (Dennett’s Orwellian revision).

Moreover, Dennett submits a third alternative on the non-anti-materialist horn of the dilemma: perhaps it is just that the multi-tracked plurality of brain-processing events allows certain processing conclusions (*contents*, for Dennett) to just kind of shunt around in the brain until they are needed by a later sub-process, or are perhaps utilized by a concatenation of later sub-processes in a conglomeration of processes, a meta-process

whereby quizzical seeming temporal kinks are smoothed over in the ultimate *content settling* that must occur for the brain to navigate us through our world and our decisions.

All of which seems well and good—until we realize, as Dennett must have, that if this be the case—and certain sub-processing are constantly sort of buzzing around on their own until utilized by later processes in the creation of a provisional ‘final’ version of what’s going on—then something’s got to give: either the temporality of the physical processes (which would lead to Orwellian/Stalinesque revision), or, the temporality of the stream of consciousness (which is what the Multiple Drafts model ultimately implies). Further, Dennett realizes that once one tampers with the so-called temporality of the stream of consciousness, one might as well get rid of the entire concept itself—for a stream of consciousness temporally kinked within itself is at least a paradoxical-seeming entity, if not some form of “wind-egg”.

So, what does Dennett do faced with this conundrum? Well, he simply sweeps this temporal problem with regard to the stream of consciousness under the rug as it were. Rather than simply desist in the use of the term “stream of consciousness”, which would seem to be the intellectually honest maneuver—he keeps it as boon companion which he befriends and insults by turns at it befits his needs of the moment. At present, in chapter six of *Consciousness Explained* that is—he seems to need it as one of his cognitive science black boxes, kept on hand to make the arcane mathematics of the Multiple Drafts work out. That is, he needs some grey zone in which he can shuffle around the multi-track, “stupid” cortical sub-processes going on such that they need not, for a while at least, map directly onto anything but themselves and their own internal, as it were, debate.

And as we will eventually see—how Dennett ultimately effects this shuffling of contents and processes around in his multi-track Multiple Drafts model is by making the Janus-faced double move of saying yes, sure there “is” such a “thing” as a “stream of consciousness”—only, this thing is no thing until it finally *appears*, and how else can it *appear* than through our reports, our, better yet, verbal reports which finally give it form. For, in a rather Platonic mood—he implies, without distinct form, without definite appearance, how can we rightly say that something *is* after all!

Now , I have wasted far too much of our time on the nitty-gritty of Dennett’s Multiple Drafts theory. And my only defense in this is to remind the reader of the metaphysical gauntlet number five thrown by Dennett at the very beginning of chapter six:

Several of the notorious experiments we will examine have been heralded by *some* distinguished experts as the refutation of the sort of conservative materialistic theory I am presenting, so if there is to be a scientific challenge to my explanation of consciousness, this is the battleground that has been chosen by the opposition. (CE 140)

That having been said, I, constituting a party of the opposition—have done little but background, frame, and explicate Dennett’s position on the specifics of the temporality of consciousness—and in this instance its two theoretical forkings into the temporality of measurable cortical sub-processes and the temporality of the stream of consciousness (however variously defined and taken up).

Therefore, quite clearly, and quite necessarily—I would like to set out some arguments for the opposition. These arguments will follow three main paths: A) the classical Searlean critique of the methodological apparatus which ultimately backs up

Dennett's employment of the multiple drafts theory: what Dennett calls Heterophenomenology; B) What happens when the Multiple Drafts and its logical progression into a general theory of consciousness (for Dennett the 'von Neumannesque Virtual Joycean Machine) get actually implemented in the real world: in robotics in general, and in particular here in Dennett's and MIT's Cog project; and finally C) my own alternative to the specifically reductive (as Searle puts it) materialism as evidenced by Dennett and other lingering clingers to the Strong AI doctrine as well as the complete interpenetration of nature and culture, the machine like and organic, natural and artificial etc....

CHAPTER FOUR: HOMUNCULARISM AND VERBAL REPORTS

So, to begin this discussion of neo-homuncularism and Dennett's adherence to the necessity of verbal reports in verifying human-like consciousness—allow me to frame one of the more marginal to this essay's concerns, yet quite central to popular, public debate on these issues theorist's (John Searle) attack on Dennett's peculiar form of homuncularism and in particular his utilization of the Multiple Drafts theory of how neurophysiological processes produce cognition as well as the methodology championed by Dennett as the only true way to access not only the consciousness of others, but indeed one's own! This Dennett dubs Heterophenomenology—and since it is here that Searle aims his first lance, I will begin with this curious construction of Dennett's. Quite simply, according to Dennett, to rely solely upon the protestations of any participant (for the scientist, more generally—any person) as to the true nature of his or her conscious experience would be to relegate our data to that most fallible and flimsy zone which the scientist calls, usually in a slightly derogatory tone, the 'subjective'. And the subjective, as we all know, is fraught with all those weaknesses of which the good scientist is so wary: it is prone to error in short, to illusion, to misrepresentation, which ultimately may well and often does lead to erroneous data—the bane of the experimental scientist! So clearly, Dennett wants to avoid this situation. However, he is very clear as well that he does not wish to throw the baby out with the bath water. That is, he does not by any means wish to completely discredit the fact that at least people 'think' they are consciously experiencing real phenomena in their mind: in short, he does not wish to say there is no such thing as phenomenology.

And for this rather tricky seeming conundrum Dennett has what he finds to be an eloquent and somewhat radical solution: the aforementioned Heterophenomenology. With Heterophenomenology, putatively, we get the best of both worlds—we get to maintain our scientific objectivity, paramount for Dennett, and as well we can keep our phenomenology (at least according to Dennett—though it’s precisely this point which Searle is to question, as we will see). The way this works for Dennett is that we (the listener, the experimenter) record the statements of the subject (of any given conscious experience), giving them the benefit of the doubt for being honest—and attempt to report as accurately as possible what their experience truly was, and then—and this is the crucial part, we check this report against what we know to be “the facts”. The color *phi* experiment analyzed above would serve as an example: we believe that the subject really thinks he saw a red spot drift right and then drastically turn green and continue to drift till it remained stationary for a time four degrees of arc to the right of the red spot. However, and this is how we cut through the illusory haze constitutive of ‘subjective’ subjectivity—we the listener, observer, and scientist “know better”, know that these are quite simply not the facts. The facts are that there were two stationary spots of light lit up with a 50 millisecond time lag in between.

And so far so good for both Dennett and Searle as well I would presume. It’s what comes next for Dennett that Searle has no truck with. For Dennett, tied as always to scientific facts concerning that which is going on in the brain—in this instance the particular sub-processes which must be going about discriminating features of the input stimulus in real time—the overall picture that emerges is one not of apparently paradoxical linkages between the time of the conscious experience and the physicist’s

timeline which underlies such. No, what emerges in the Heterophenomenological analysis for Dennett is a bunch of disparate sub-processes (which he will later curiously wish to dub *homunculi!*) buzzing along with their own interpretations of the inputs, only ultimately to make a holistic picture or story out of all the variegated data *sometime later*, always with later sub-processes somehow combining the informational conclusions of earlier sub-processes' activity, sometime later but always, as non-contradictory science would dictate—before any communication of this wholistic picture could take place.

Thus it is, if one were to be critical of this Multiple Drafts theory *cum* Heterophenomenological methodology—that in what I earlier called Dennett's shell game of consciousness, Dennett sweeps any “real” stream of thought under the rug of the some hundreds of milliseconds required to “prepare” a speech act and the ensuing hundreds and hundreds needed for any enunciatory act (each syllable taking some two-hundred milliseconds to effect). So it is that Dennett maintains the vocabulary of “stream of consciousness” while at the same time eviscerating it of its normally intended content. The major metaphor which Dennett uses to either illustrate or prove (it's not ultimately clear to this reader which one it is) the Heterophenomenological method is that of reading a fictional text (in particular—*Sherlock Holmes* stories)—which if I take him correctly constitutes a world just like that of a subject's Heterophenomenological world. Then taking particular statements made by the subject as if they were things say proposed to obtain in Holmes' London though not explicitly stated in any of the texts—it is the job of the Heterophenomenological observer/literature critic to figure out just whether this *really could* obtain in Holmes' London.

And the obvious underlying implication here is that consciousness itself is nothing but a fictitious text itself—to be read with caution and trepidation when translated to one through another’s words—which by the way seems to be the only way ultimately we ever gain access to it, even in the case of our *own* consciousness!

As Dennett notes in his final summation of his theses on Heterophenomenology, in chapter 12, “Qualia Disqualified”, section 6, “Getting Back on My Rocker” of *Consciousness Explained*:

Since I have encouraged us all to be heterophenomenologists, I can hardly exempt myself, and I ought to be as content to be the subject as well as the practitioner, so here goes: I apply my own theory to myself. As heterophenomenologists, our task is to take this text, interpret it, and then relate the objects of the resulting heterophenomenological world of Dennett to the events going on in Dennett’s brain at the time As it is, the text portrays a mere portion (and no doubt an idealized portion) of the contents of the author’s consciousness. We must be careful, however, not to suppose that the “parts left out” in the given text were all “actually present” in something we might call the author’s stream of consciousness. We must not make the mistake of supposing that there are facts—unrecoverable but actual facts—about just which contents were conscious and which were not at the time. (CE 407)

And, as I say above—I believe that this pretty much sums up the situation for Dennett. And what is so stupendous about this last of Dennett’s formulations of the Heterophenomenological method is that he is claiming that it is being done by himself upon himself. What better proof could one ask for? If, for example, some such “brute fact” as an occurrent stream of consciousness content were indeed potentially “recoverable”—who other than its own one-time owner could perform such a feat? To the contrary, claims Dennett—there is no such thing [as such an occurrent stream of

consciousness]— it is merely the report which survives. Anything else is merely an apparition based upon a shade which never existed in the first place. I feel I need to repeat this: here we have Dennett himself, in discussing his own past, his own passed “stream of consciousness” claiming not in effect, not virtually, not after the fact—but *essentially* and *really*—that there never was any such thing—only and nothing but what was ultimately to become a verbal artifact.

Thus it is that with the explanatory mechanism of the Multiple Drafts theory and the investigatory hermeneutic of Heterophenomenology—any such thing as has traditionally been understood as consciousness (having its own internally, intrinsically integral existence—evidenced in what are most generally called *qualia*, the *phenomena* of phenomenology) has been, according to myself and Searle, completely ushered off the theoretical stage. For Searle this is an egregious, no *the* egregious, error committed by Dennett in his theses on consciousness. For it is precisely these *qualia*, these *data* as he calls them—that must be explained by any adequate theory of consciousness. And this is due to the, what Searle calls, Subjective Ontology of consciousness which means that when it comes to consciousness, in a word, what you see is what you get.

Searle’s counter example is that of a sunset—where he claims that though what we think we see is a yellow orb slowly becoming more and more bloody, more orange till it peaks into a deep red and sometimes, very rarely, appears to jet forth a tiny flame of green phosphorous off its upper limb—science clearly shows that this is a mere apparition caused by particles in the atmosphere bending the light rays ever more as more atmosphere must be traversed. However, with consciousness, claims Searle (and I would agree with him here)—the case is rather different, for, as the term Subjective Ontology

should suggest, with consciousness the *appearance* is in fact the *reality*. What he means is simply that if something appears to you to be in your consciousness experience, well then *a fortiori* it very well is! Now this doesn't mean that it is necessarily an accurate translation of incident stimuli nor accurate portrayal of the surrounding world, but, at the very least (echoing in a minor way Descartes here) it is indeed what it seems to be, and what it seems to be is what it is! Now, this position seems to be simply that mandated by not just 'common sense' and it is just this sort of 'common sense' postulate and hermeneutic that Dennett is more than happy to joust with, if not at times scoff at. And in a sense he appears to do just that in the pages of *Consciousness Explained*.

Below I offer Searle's injunction for what must be included in any worthwhile theory of consciousness, as well as his ultimate analysis of Dennett's methodology as it appears in the pages of Searle's *The Mystery of Consciousness*:

The subjective feelings are the *data* that a theory of consciousness has to explain, and the account of the neural pathways that I sketched is a partial *theory* to account for the data. The behavioral dispositions are not part of the conscious experience, but are caused by it. (MC 99)

Now this is in direct loggerheads with Dennett's hypotheses in *Consciousness Explained*. For Dennett, it is precisely "behavioral dispositions" that underlie the whole mechanism of consciousness. However, for Dennett, it is far too weak a claim to stop at saying that behavioral dispositions *underlie* consciousness—for at the end of the day, that is basically all that we are left with after Dr. Dennett has gleaned from the mind with his Occam's razor just what makes it tick, and nothing more. Everything else, including all

qualia, are what Searle characterizes as a “mistaken judgment” on behalf of the subject according to Dennett.

Which inkling will lead Searle to his ultimate analysis of the Dennettean model of mind (as implicitly, again, arising from the Heterophenomenological paradigm of interpretation):

It [mind] is all a matter of third-person phenomena: stimulus inputs, discriminative states (p. 372 ff.), and reactive dispositions. The feature that makes all these hang together is that our brains are a type of computer and consciousness is a certain sort of software, a “virtual machine” in our brain. (MC 100)

Now, such that the reader doesn't think that Searle is exaggerating or mis-attributing any theoretical constructs to Dennett that Dennett didn't actually intend, allow me to dip back into *Consciousness Explained* for a moment, to give you the theory in Dennett's own words:

Human consciousness is itself a huge collection of memes (or more exactly, meme-effects in brains) that can best be understood as the operation of a “vonNeumannesque” virtual machine *implemented* in the *parallel architecture* of a brain that was not designed for any such activities. [emphasis in the original] (CE 210)

Now, it is my intention to show how by a very clever double move, Dennett takes what he believes he has proven as to the non-existence of the stream of consciousness via the Multiple Drafts theory in conjunction with what this leads him to: that, as we see above, *all* that human consciousness is the implementation of serial “vonNeumannesque” computational logic (on the “software” level, if you will)—and in their conjunction finds

not only that what we humans are is some kind of inefficiently embodied computer (ergo: robot), but that conversely, and by the extension of current technologies to the near future—what robots are (or are about to become) . . . is us!

To quote Dennett from a rather flattering December 9, 1998 interview in *The Atlantic Online*, conducted by Harvey Blume—where we have Dennett waxing what Hubert Dreyfus might have warned based upon the pages of *What Computers Still Can't Do* could be seen as a little *over-optimistic*: “The best reason for believing that robots might some day become conscious is that we human beings are conscious, and we are a sort of robot ourselves.” (AO 2)

Now, so far so good; so far so demure . . . but in true Dennettean fashion this is but a facade to set one up for the sucker punch to come in a few paragraphs [here riffing on his most recent *Darwin's Dangerous Idea*]:

You write small pieces of dumb code that work with other simple pieces of code in order to produce systems of greater complexity, which in turn interact with other complex systems in order to give higher degrees of functionality, and so on, until you wind up with a program that is smart—or, at least, smart enough to do something that needs doing. Finally, you get operating systems, you get an Internet—or, depending on your raw materials and the time allotted, yet get DNA, mammals, and self-awareness. (AO 3)

So here we have an argument concerning design which mirrors the logic of the one which says since we're robots—why robots are us!, only going one level deeper one step further and claiming that since all that the creation of organic beings entails is the laborious concatenation of “dumb code” added to more “dumb code” added to more

“dumb code” presumably the lab technician lacks none of the abilities, only the time to create what we know as life. However, rather than taking the ‘traditional route’ and claiming that what they are doing in the lab with Cog the cognitive robot is really special because of x, y, and z and hence he’ll probably be mimicking if not embodying a human like consciousness perhaps sometime soon—Dennett takes the opposite tack (in fact has been undertaking the opposite tack in full force for the past fifteen years ever since the publication of *Consciousness Explained*)—and wants to say that what you call life and what you call consciousness is actually really nothing special at all, it’s all quite machine like at heart, and hence our reproduction of it is nothing really special at all either. All it is the addition of one stupid encoded sub-system, set of instructions, added to another and another etc. (ergo the congregate of *homunculi* mentioned above).

Hence, to clarify what the last part of this chapter will attempt: on the one hand I will like to investigate the theses of Dennett on the correlation between his ‘cognitive robot’ Cog and human consciousness—and see whether claims for their identity or more recently near identity are borne out in actual fact in the comparison of two particular common traits claimed by Dennett—the experience of pain (and hence ‘meaning’ in a sense) and the existence or non-existence of a stream of consciousness. I will begin with the latter first, as this has been the issue with which the essay has dealt most centrally up to this point.

I will begin with a summary of Dennett’s treatment of the stream of consciousness as he has it appear in the human consciousness experience—then see how this apparently seamlessly translates into a theory of robot consciousness being just (like) that of humans. Once we have consciousness as translated into the robot form—it will be

my interest to see whether this is actually re-convertible back into human consciousness, or whether due to the exigencies of robot-creation Dennett in effect begins to slide back somewhat from his fully staunch Strong AI, cyborg friendly position. After we have analyzed these questions, as I mention above, I will end with a rumination on pain: pain in the robot and pain in the human and whether the twain shall ever meet.

But enough—to wind up Dennett’s theses on time consciousness, allow me to return for a moment to the article from which I drew at the beginning of chapter 5: Dennett’s “Is Perception the ‘Leading Edge’ of Memory?”, where he ultimately concludes that yes there is a sort of extendedness, a certain blurriness to the temporality of consciousness. What we have to be very careful about, however—is the question as to what mechanism, and what reality actually underlie this blurriness for Dennett, which as I think we shall see will lean back upon his notions of Heterophenomenology and the Multiple Drafts model of mind. So, his concluding paragraph from this article putatively discussing memory and perception:

. . . since Descartes was wrong about there being a point in the brain “where it all comes together”—if you ask yourself the question, “When did *I* become conscious of a particular event?” that question can have only a vague answer, not a precise answer. It could have a precise answer only if we could locate *you* at some point in your brain. Since the transmission of information around in the brain is relatively slow, the dating of events in consciousness—the dating *for you*—has to be smeared over maybe as much as 200 msec, a fifth of a second. (PLEM 78)

Now, with this ‘smearing’ it would begin to look like Dennett might be drifting toward some kind of theory which reifies the stream of consciousness at the end of the day after all. But as we will see this is to be far from the case. For what is really being blurred in

the Dennettean scenario is not so much the ‘when’ in the “when did I become conscious of a particular event?” but rather the ‘I’ in the interrogative.

That is, note the exact phraseology employed by Dennett: “the dating of events in consciousness . . . has to be smeared . . .”, which makes it clear that what is fundamental is not some agency which needs must blurry some external events-times such that it maintain its subjectivity but rather the primordial facts for Dennett are the *events* themselves, which due to the chaotic multiple drafts scenario needs must be simply blurred—mis-taken in their being taken up by the brain, by consciousness. To clarify this point, let’s return one last time to Dennett’s final summation in his own words of, by his own lights, the most difficult and technical section of *Consciousness Explained*, that on the Multiple Drafts theory. I have decided to quote this passage at length to preclude any doubts that I am “putting words in Dennett’s mouth” as it were in my characterization of this theory of consciousness:

As we have already noted, this multitrack process [hence the Multiple Drafts] occurs over hundreds of milliseconds, during which various additions, incorporations, emendations, and over-writings of content can occur, in various orders. These yield, over the course of time, something *rather like* a narrative stream or sequence [italics in original], which can be thought of as subject to continual editing by many processes distributed around in the brain, and continuing indefinitely into the future. Contents arise, get revised, contribute to the interpretation of other contents or to the modulation of behavior (verbal or otherwise), and in the process leave their traces in memory, which then eventually decay or get incorporated into or overwritten by later contents, wholly or in part. This skein of contents is only rather like a narrative because of its multiplicity; at any point in time there are multiple drafts of narrative fragments at various stages of editing in various places in the brain. While some of the contents in these drafts will make their brief contributions and fade without further effect—others will persist to play a variety of roles in the further modulation of internal state and behavior and a few will even persist to the

point of making their presence known through press releases issued in the form of verbal behavior. (CE 135)

So, to look forward to the next chapter, I would like to ruminate on just what this theory of “press releases” in the guise of “verbal behavior” might imply for our consciousness as a whole, and in particular how it might impinge on the form and nature of our streams of thought. Silently.

CHAPTER FIVE: THE JAMESEAN STREAM OF THOUGHT; CRITIQUING VERBAL REPORTS

Now that we have a fairly good idea of where Dennett is trying to go with his neo-homuncular model of mind and the verbal reports which act as vouchsafe for any full-blown acts of consciousness, facts of consciousness—we are ready to return to the purely Jamesean notion of the stream of thought. I believe we can detect therein, especially within his theses concerning the correlation of the stream of thought to the verbal artifacts which are produced by those exhibiting such, a doctrine in good keeping with the Bergsonian notion that language is merely one of means by which we fix the always mobile reality subtending our necessarily rigidified social interactions. But a more essential reality lies beneath this solidified, reified, and congealed exterior which we present to each other as *zoon politikein*. Underneath this frozen veneer of language lies a seething, streaming, writhing flow of thoughts and ideas and images which never end, and are never truly disjointed in the manner in which words and photographs and serial von Neumannesque computer processors are.

Therefore, by way of reminding the reader of the nondiscontinuity of the bodily, sensorial substrate attributed to the stream of thought by James in his formulation of this essential concept in *The Principles of Psychology*—allow me to give one more rendition of the Jamesean version of this foundational principle:

Resemblance among the parts of a continuum of feelings (especially bodily feelings) experienced along with things widely different in all other regards, thus constitutes the real and verifiable ‘personal identity’ which we feel. There is no other identity than this in the ‘stream’ of subjective consciousness which we described in the last chapter. (PP 336)

Here we have another clear-cut argument for the continua of bodily feeling, a bodily substrate, as the primordial fact of our personhood, our personal identity, and foundation for our stream of thought.

No longer in the Cartesian space of a *res cogitans* relying upon a pineal interface to the body—with James we have a wholistic continuum of bodily sensations, emotions feelings and thoughts which arise from such and yet are never out of contact with these very crucial components of the mind-body-thought complex.

Now it is not my supposition that the Dennettian homuncular and Multiple Drafts models deny the bodily substrate in itself in the least bit. However, it *is* my position that when the ultimate arbiter of consciousness becomes a verbal report and the primordial bedrock of the stream of thought becomes a digitalized synapse which is on or off, open or closed, perhaps even ‘zero’ or ‘one’ according to some particularly technophilic adherents of Dennett, then that which is fluent in living thought and that which makes the unique, irreducible, and unmistakable experience of consciousness just what it is and not some print-out generated by a machine directed by an algorithm created in turn by conscious beings gets lost in the mix.

And it is just this fluency of the stream of thought which will be seen to be the thread which attached an erstwhile fragmentary and always disjointed self which would

be all that we were left with were we to be left with nothing but verbal reports and possibly zombie-like gestures, along with their concomitant potentially digitally generated substrata.

As James says in *The Principles of Psychology*:

The transition between the thought of one object and the thought of another is no more a break in the *thought* than a joint in a bamboo is a break in the wood. It is a part of the *consciousness* as much as the joint is a part of the *bamboo*. (PP 240)

Therefore James believes in essence that there is never a break in the stream of thought—and that it is just this fact that makes it *ours* and us by turns *its* unwavering progenitor.

Yet this is quite obviously speaking only to our interior and inner lives, the exterior self, the self we present to society and the self we express in language and in our public acts is an entirely different creature—when viewed through a certain optic.

And it is just this divergence of points of view with which I am primarily concerned here. For at heart, I do not believe that Dennett or other neo-homuncularists do not believe there is something such as an occurrent stream of thought which is phenomenally evident to every onlooker in the movie of life. Rather, the question of emphasis becomes so tantamount that at the end of the day one is left with two models of mind.

In the one, the Dennettean neo-homuncular one, the movie in the brain is so fraught with difficulties, so apt to be deluded and to delude itself, that it basically cannot be trusted to give us a bedrock for the consciousness thing itself. Therefore we must look for such in more trustworthy phenomena, such as the firing of neurons or the words which we speak and write.

For others, such as Damasio, Bergson, and William James, the movie is of paramount importance—and of paramount reality *and* veracity. They too are interested in how this movie comes to be, what biological processes underlie it, and to what ends, in a more philosophical vein, it might ultimately be pointing...

However, those of the latter bent of mind are not willing to throw the baby of the stream of thought, of the movie in the brain, out with the bathwater, even if it is rather grey and murky—out with the bathwater of complications we are faced with when we try to use the tools, and only the tools available to us, which modern science has to offer.

But I will have more to say on this in the last two chapters. For now, my primary goal is to elucidate how unimportant and fragile for James at the very least and surely for Bergson as well, linguistic production is, in the overall game of pinpointing just what it might be to describe what “consciousness” itself could be.

For starters, both Bergson and James point to the very social pragmatism and practical dictates which force our entry into the linguistic realm. Our use of language is of a very practical nature, and for Bergson always already aimed at some ultimate Action.

And to summarize much of Bergson’s ruminations on the state of modern metaphysics and the capitulation to science as the ultimate descriptor of existence *soi-meme*, it is through our confusion of the rigidified forms of communication and even the

thought which has absorbed this rigidity with the underlying realities that has led to such quizzical post-modern positions such as that one which holds that linguistic artifacts are final and sole arbiters of consciousness.

And why might this be? Well for Bergson at least, it is all rather simple: “The reason is that our outer and, so to speak, social life is more practically important to us than our inner and individual existence.” (TFW 130)

Hence it will be with the question of language in general, and in particular with the need for verbal reports to shore up any certainty of psychological events occurring and hence consciousness either existing or not existing as the case may be.

Language is obviously a social construct we who exhibit it use to better organize ourselves as a whole and better fit ourselves as individuals into the whole. And insofar as this dynamic is of paramount importance for our survival as a species and as individuals, it is not surprising that we may be inclined to place a great amount of importance upon the artifactual productions of language itself over against language in the process of being produced.

And in William James’ discussion of language in the chapter “The Stream of Thought” from *The Principles of Psychology*, he pays particular attention to just this fluidity of language in the process of being produced. In discussing the connectivity of thought-processes in the brain, and hence thoughts, at a time well before the MRI had been invented (*The Principles of Psychology* was originally published in 1890) and many years before Husserl created his famous schematic of the interaction of protentional and retentive memories, James talks about the vivid excitation of “brain-tracts” and their subsequent dying vibrations. Much simpler *esquisse* than Husserl’s seeming elevation of

a cantilever bridge, James' diagrammatic representation of these vivid excitations and dying vibrations, always coloring each other and bleeding one into the other, the dying vibration of *a* never being completely disconnected from the vivid excitation *c*—looked like a backsliding tenor clef with no stave and only three notes.

But beyond these proto-Husserlian diagrammatics of the excitation and dying embers of various brain-tracts involved in the process of thinking, James likes to focus on what happens when and if we attempt to take a cross-section of this ever evolving thought process, that is if we attempt to arrest it in the midst of its activity, or take a snapshot view of it in Bergsonian terms. As he says quite concisely “if our purpose is nimble enough and we do arrest it, it ceases forthwith to be itself.” That is, this operation is akin to that of trying to seize a spinning top to “catch its motion”, or turn up the gas (these days it would be turn up the dimmer I suppose) quickly enough to see the darkness. That is, again echoing Bergson, all of these attempts to arrest thought in the process of thinking, which is also tantamount to translating thought in the process of being thought *directly* into language—all boil down to the same type of fallacy as those exhibited in the paradoxes of motion immortally laid out by Zeno of Elea in the sixth century B.C.E. In other words, simply put: where *is* the arrow, which is in flight, at any *instant* of its flight? Which involves the deeper paradox, delineated over and again in innumerable venues by Bergson, of the attempt to build motion out of immobilities.

And quite interestingly, it is just this paradoxical attempt to freeze thought in its tracks, in its endless flowing, that James imputes to those “doubting psychologists” who will surely challenge anyone to “*produce* these psychoses” (psychoses in the sense of

mental state, not the clinical modern psychiatric conception) that is to produce *concrete* evidence of such, who could contend for their existence.

However, remarks James:

As a snowflake crystal caught in the warm hand is no longer a crystal but a drop, so, instead of catching the feeling of relation moving to its term, we find we have caught some substantive thing, usually the last word we were pronouncing, statically taken, and with its function, tendency, and particular meaning in the sentence quite evaporated. (PP 244)

There is the simple grammatical point that the meaning of the whole sentence is lost as we stop on any particular word. However, this is not James' main point here. What he is trying to point out is the "baleful" results we always come up with when we try to catch ourselves in the act of thinking, with the very thoughts the only tools we have to do the catching. To use another metaphor to symbolize this attempt to immobilize the mobile, and create mobility out of immobiles—Bergson cites in a slightly different context the "disappointment" of the child who tries to grab hold of a wisp of smoke by clapping her hands together.

In fact, Bergson will go on to posit in effect two different main kinds of 'self'—a social self, and a more fundamental self underlying the one who breaks itself and its thoughts into little pieces the better to get along with others and act in a world which prefers stabilized separated entities, especially when it comes to such things as acting together in a congruent and wholistic manner such as performing the Nutcracker Suite during the holidays or invading Iraq in the springtime.

As Bergson notes in his earliest work, *Time and Free Will*:

Consciousness, goaded by an insatiable desire to separate, substitutes the symbol for the reality, or perceives the reality only through the symbol. As the self thus refracted, and thereby broken to pieces, is much better adapted to the requirements of social life in general and language in particular, consciousness prefers it, and gradually loses sight of the fundamental self. (TFW 128)

Here we can see that according to Bergson it is the exigencies of social life which encourage and eventually demand this fragmentation of the self and substitution of symbol for reality which is so important to the individual organism. And after habitually and repetitively doing this for some period of time, such as occurs when one has reached adulthood in the case of humans, the more mobile and fluid fundamental self underlying all these symbolizations gets occulted by the very symbols which arose in the first place to aid this organism in its adaptation to the world. Hence the very tool of adaptation becomes itself a veil thrown over not just the surrounding world but across the eyes of the self doing the adapting. And in the latter stages of this progression, the subject immersed in its own language begins to perceive the reality only through the symbol, eventually substituting the symbols designed originally to help in understanding and dealing with reality for the reality itself.

And stepping back for a moment into the Jamesean stream of thought, we have another rather poetic ‘symbolization’ of the ongoing progress of thought in its own internal evolution:

As we take, in fact, a general view of the wonderful stream of our consciousness, what strikes us first is this different *pace* of its parts. Like a bird’s life, it seems to be made of an alternation of flights and perchings. (PP 244)

If the fundamental self's life is intertwined with a streaming consciousness such that like in the course of the life of a bird there are alternations of "flights and perchings" then according to the tenets held by Bergson above, it would seem that the perchings have come to roost in the place where once there were flights. Which is to say that for doubting psychologists such as Daniel Dennett and for those who have forgotten that the veil of language was invented in the first place to better adapt as whole and evolving organisms to the open and always environing world the symbol has indeed become substitute for the reality, and it is only through the gauzy veil of the symbolic that we can catch glimmers of the underlying mobile reality. One might be reminded in this regard of Julia Kristeva's famous semiotic/symbolic distinction from *The Revolution In Poetic Language* but this would take us far afield from our narrow focus on the robotic or nonrobotic, homuncular or nonhomuncular basis of human consciousness.

To return to the bird's life for a moment, it is telling that in a letter to William James dated 6th January 1903 Bergson mentions that his conception of psychological life would seem to be entirely compatible with the one delineated in *The Principles of Psychology* save for one detail, if detail we can call it: "I see *places of flight* in the *resting-places* themselves, rendered apparently immobile by the fixed gaze of consciousness." (HBKW 357)

The places of flight being again the transitive relational and transitory modes of thought while the "resting-places" or perchings are more substantive and concrete points of solidification—Bergson echoes here the difficulties foreseen by James above when 'put to the test' by certain "doubting psychologists" the theorist who might contend for the

existence of a prefatic stream of thought in itself will face when trying to ‘prove’ via concrete evidence that the “psychosis” is really *there* after all. But Bergson would go one further than James, and submits that even the perchings or “resting-places” themselves are more mobile than we are normally ready to accept. Here the “fixed gaze of consciousness” acting in the role of cross-section of the thought in progression as James saw it, which cross-section would always be bound to disappoint as the mobility of the thought in process itself always loses something when being confined to the frozen and rigid boundaries dictated by the nature of its translation into stable language, into words, in short.

Or as Bergson rather more eloquently puts the same general notion in *Time and Free Will*:

In short, the word with well-defined outlines, the rough and ready word, which stores up the stable, common, and consequently impersonal element in the impressions of mankind, over-whelms or at least covers over the delicate and fugitive impressions of our individual consciousness. To maintain the struggle on equal terms, the latter ought to express themselves in precise words; but these words, as soon as they were formed, would turn against the sensation which gave birth to them, and, invented to show that the sensation is unstable, they would impose on it their own stability. (TFW 132)

Thus for Bergson words are storehouses of the stable, commonly held placeholders in the ongoing game of somehow exteriorizing our interior life such that we can tell our lover just where it is on our back we want them to scratch, or the doctor just what part of our anatomy hurts and how!, or our pilots just what village in Iraq we want them to bomb this morning! But due to their necessary commonality they thereby are no longer so personal,

and hence the “delicate and fugitive impressions of our individual consciousness” get covered-over and in some cases over-whelmed by these very placeholders themselves.

And this is just what appears to have happened to those “doubting psychologists” such as Dennett who would require stable placeholders in the stead of the delicate and fugitive impressions which form the flora and fauna of our internal lives. Rather ironically, here we can see Bergson bemoaning the intransitory nature of the words, the placeholders we use to translate and thereby cover-over our internal impressions. Far from wanting (stable) placeholders to translate the (concrete) internal impressions created in our minds and perhaps brains, as does Dennett, Bergson is wistful that the words we create to translate our fluxional interior lives by their very nature needs must stabilize and freeze into commonly understood and hence somewhat trammled meanings. All of the fluidity of the interior stream of thought is thence sublimated into rigid placeholders whose stability imposed from without thereafter make of an essentially numinous quality and feeling of tendency toward an uncertain end, which is the nature of the stream of thought in its very flowing, into something frozen into way-stations imposed in a certain sense from without due to the dictates of utilizing a vocabulary born in the commonalities of a society’s shared linguistic structures. But this operation only works when effected from *without*, even if the *without* in this case is a thinker reifying their own stream of thought after the fact and reporting to themselves on its course having already been taken and thereafter frozen into stabilized meanings, such as is the case with Dennett’s famous utilization of the Heterophenomenological Method on himself. What gets lost in the mix here is the feeling of tendency, a crucial aspect of the stream of thought according to

James. As he says in “The Stream of Thought”: (For the doubting psychologist—) “The tendency is a *psychical* zero; only its *results* are felt.”

Once again reminding us of Bergson’s main idea in *Time and Free Will* that we can only spatialize, chart-out, map distinctly a thought in its flowing, a decision made under the auspices of free will after the fact of having flowed, after the decision has been made—and the recursive operation effected whereby the stream is frozen and the decision mapped-out in a coordinate system delineating the branches of possibility as viewed from the perspective of hindsight. This happens in instances such as the ‘use’ of the Heterophenomenological Method applied to oneself or others, but this operation only ‘works’ in retrospect as it were, after the stream of thought has run its course, and gets frozen and reified into publicly (“globally accessible”) verifiable coinage, where it can be “cashed-out” as the Anglo-American philosophers like to say.

But under this paradigm, all of the vagaries and uniqueness of each individual’s stream of thought and interior conscious life get smoothed over and flattened by the paving machine of language. But it was one of James’ main goals in his psychology to rescue some of the vagueness and specificity of the interior life of the stream of thought from their overly brutal concretization on the behalf of those “doubting psychologists”. As he mentions in *The Principles*, “It is, in short, the re-instatement of the vague to its proper place in our mental life which I am so anxious to press on the attention.” (PP254)

I will leave off this topic for the moment with another more poetic passage of James’ restating the same basic idea:

The traditional psychology talks like one who would say a river consists of nothing but pailsful, spoonsful, quartsful, barrelsful, and other moulded forms of water. Even were the pails and pots all actually standing in the

stream, still between them the free water would continue to flow. It is just this free water of consciousness that psychologists resolutely overlook. Every definite image in the mind is steeped and dyed in the free water that flows round it. (PP 255)

CHAPTER SIX: MORE HOMUNCULARISM, ENTER WHOLISM

Now, to return to the “press releases” mentioned in chapter four—in a recent attempt to further flesh out his Multiple Drafts model, Dennett, in his 2000 “Are we Explaining Consciousness Yet?”, submits that the Multiple Drafts did not provide “a sufficiently imagination-friendly antidote to the Cartesian imagery” (AWECY 3) we have all been lumped with. Therefore, he has come up with the new, improved “guiding metaphor” which he calls “fame in the brain” or “cerebral celebrity” [Dennett, 1994, 1996, 1998] which basically means that what counts ultimately as consciousness is the congregate of those familiar “stupid” sub-processes in the brain which actually ‘make it’ to the ‘stage’, who garner enough “clout” among the other sub-processes to actually gain “control of the body”.

In so far as the Multiple Drafts model is discussing “control of the body” I can see no problems. The problems arise when Dennett further attempts to use this model as a weapon against “phenomenality” itself. Which “phenomenality”, defined by his straw man interlocutor Block as “experience” itself, Dennett basically wants to do away with.

As Dennett sums up the section in the essay:

But then what is left of the claim that phenomenality is experience? What is *experiential* (as contrasted with what?) about a discrimination that is not globally accessible? As the convolutions of Block’s odyssey reveal there is always the simpler hypothesis to fend off: there is *potential* fame in the

brain (analogous to the dispositional status of poor Jim, the novelist) and then there is fame in the brain, and these two categories suffice to handle the variety of phenomena we encounter. Fame in the brain is enough. (AWECY 4)

What Dennett is driving at with this talk of ‘potential’ fame in the brain being meaningless is once again reminiscent of the tenets of the Heterophenomenological method, that is to say that in lieu of a demonstrable physical action (hence the “barefoot behaviorist” pole of his theory of consciousness) or an unambiguous “globally accessible” articulation, all we have is potential fame, potential consciousness. While hiding somewhat behind the blind of the turgidity of the phrase “global accessibility”—let us not let ourselves be fooled: what Dennett is claiming, again, in 2000—though with a new, improved phraseology, is just that which he imputed to “consciousness” in *Consciousness Explained*. That is to say that without an unambiguous verbal report, or a limb moving in some scientifically detectable manner—there *IS NO CONSCIOUSNESS*. Or, as Dennett puts it rather more eloquently in this essay—at the very beginning of the subsection under discussion titled “Competition for Clout”: “The basic idea is that consciousness is more like fame than television . . .” (AWECY 3)

And now that I’m fairly sure that everyone is pretty well convinced that the horse of ‘what does Dennett mean by consciousness?’ isn’t going to flinch any more when kicked—allow me to try to tie all this in, as promised above, with what Dennett thinks is going on with robots, and what he thinks with respect to whether they will ever be conscious like us humans. I hate to say it, but we must return one more time to the “stupid” brain sub-processes which are from time to time raising one another upon each other’s shoulders to the status of “fame” and hence producing consciousness—as it is

precisely through the clarity of Dennett's characterization of this pell-mell that he is able to attach consciousness to machines, and at that in a much simpler manner than had heretofore been achieved.

Ironically I think we can see this most clearly when we see Dennett fending off the phantasmatically imagined counter-argument to his brilliant "fame in the brain" which might ask: 'Who is doing this elevating of "demons", of *homunculi* to the status of "most famous one"? Doesn't this necessitate some *agency* at least, albeit an inferior, *sub-agency* if you will? And aren't we simply caught in an infinite regress of agencies—even though the bigger agency is resting on a smaller agency and so-on until we wind up with such an infinitesimal *micro-agency* that we might as well not call it *agency*—would these not be partaking of agency itself all the same—and hence would we not really be getting rid of it altogether? Thus even though divided into a thousand or even a thousand bits—don't we still wind up back with some form of Cartesian *ego*?, albeit diffracted a million times over.' (Imaginary Interlocutor)

Well, as usual, Dennett's pithy solution is so eloquent that we might imagine it dividing the edge of Occam's razor in twain!:

The looming infinite regress can be stopped the way such threats are often happily stopped, not by abandoning the basic idea, but by *softening* it. As long as your *homunculi* are more *stupid* and *ignorant* than the intelligent agent they compose, the nesting of homunculi within homunculi can be finite, bottoming out, eventually, with *agents* so *unimpressive* that they can be replaced by machines. (AWECY 3)

And again, what I will submit is that this is all well and fine for a consciousness based upon the hypotheses of Dennett as elucidated in the Multiple Drafts model and as

discovered via the Heterophenomenological method. Especially if we are mostly concerned with ‘intelligence’ capacities (read processing abilities) and the spatiality of where what is happening in the brain. What remains problematic, I am going to insist—is the experience of consciousness as it is really experienced. And yes, I take it as a fundamental postulate that consciousness is really experienced, Block-straw-men notwithstanding. Following Searle’s lead, I will claim that the rough experiential data of these experienced conscious events must be explained by any complete theory of consciousness.

So, as I’ve said above, having beaten quite dead the horse on the side of the equation which is Dennett’s theory of consciousness in humans—lets take a peek at what Dennett has been doing in the machine world, with an eye toward hopefully balancing both sides of the equation into one clear picture of why, randomly inserted coefficients notwithstanding⁷ something is missing in both accounts!

So, on to the thick(et) of Dennett’s robot-world!: and in particular his recent essay “The Practical Requirements for Making a Conscious Robot” wherein he deflects the second of “a nested series of reasons someone might advance for the impossibility of a conscious robot”:

(2) Robots are inorganic (by definition), and consciousness can exist only in an organic brain.

Why might this be? Instead of just hooting this view off stage as an embarrassing throwback to old-fashioned vitalism we might pause to note that there is a respectable, if not very interesting way of defending this claim. Vitalism is deservedly dead; as biochemistry has shown in matchless detail, the powers of organic compounds are themselves all mechanistically reducible and *hence* mechanistically reproducible at one scale or another in alternative physical media; but it is conceivable—if

⁷ For example, the 200 msec needed to prepare a “verbal” report.

unlikely—that the sheer speed and compactness of bio-chemically engineered processes in the brain are in fact un-reproducible in other media. [author's italics] (PRMCR 1)

And it is precisely at this “hence” which I have italicized that I have problems here. I don’t see how it necessarily follows that just because all the organic compounds in a living being are “mechanistically reducible”, that is analyzable—that they are therefore and mysteriously because of this “reproducible”. Now, I’m not saying, again in keeping with Searle—that I believe it is logically *impossible* to reproduce these systems—all I’m saying is that the claim that since something is analyzable, thanks to this it is *reproducible*, is *not* true. We very well may be able to reproduce whole organisms in the future, and indeed in-depth analysis is the first step of this process. Further, we have already seen great progress in the field of prosthetics in even the last few years—whose benefit to humanity is beyond question and a wonderful boon. I would just sound a cautionary note—in the spirit of *What Computers Can’t Do*—that Dennett should be careful about jumping the gun, as seems to be his wont, because as Dreyfus has shown in painstaking detail, the over-optimistic chest-pounding and predicting which is a tradition in the AI technophilic community can backfire in more ways than one!

But aside from these generalities, let’s look at the particularity of how Dennett has begun to *reproduce* a human in the form of the Cog robot in his collaboration with the MIT team who is working to create the “cognitive automaton”. Of tantamount interest to this article is of course how the ‘mind’ is proposed to be reproduced in the Cog robot, bodily limb and even organ reproduction at this point being much further advanced and much less controversial (robotic arms and even heart valves have been around for years—but brains, no!) Furthermore, the whole essay up to this point has served just to

give a theoretical platform, from the perspective of Dennett—for how it is that what is going on in Cog is just what is going on in humans (though in reverse, of course, as we have been seeing Dennett describing what’s going on in humans first—which, amazingly!, happens to be what goes on in robots!—but here *I* jump the gun!)

So back to “The Practical Requirements for Making a Conscious Robot”:

In other words you can think of Cog’s brain as roughly equivalent to sixty-four Mac-II’s yoked in a custom parallel architecture. Each node is itself a multiprocessor, and instead of running Mac software, they all run a special version of parallel Lisp developed by Rodney Brooks, and called, simply, L. (PRMCR 5)

So here we have the basic physical structure as it were of Cog’s brain. Now, let’s see just why this gives it all it needs, according to Dennett, to attain human consciousness. As we saw above the Multiple Drafts model of what’s going on in consciousness elaborated itself in Dennett’s thought into the “vonNeumannesque Virtual Machine” in the pages of *Consciousness Explained*. But it has also been maintained up to the present—the virtual machine being a model not only of what’s going on in humans, but of just *all there is* as well. Once more, allow me to give you the words of Dennett himself:

Since any computing machine at all can be imitated by a virtual machine on a von Neumann machine, it follows that if the brain is a massive parallel processing machine⁸, it too can be perfectly imitated by a von Neumann machine. (CE 216)

⁸ It is here that Dreyfus directs his main assault against Strong AI, and while completely agreeing with Dreyfus’ theses on the question of whether *all* that is going on in consciousness is computation (to which he answers a resounding “No!”) I obviously do not have the space here to pursue this particular line of attack on Strong AI.

So here we have it, the human brain can be “perfectly imitated” by the von Neumann machine—ergo, Dennett has set himself up ten years in advance for saying, ‘well, look at my (irrefutable) *Consciousness Explained* (Dennett 1991): Cog, with his sixty-four Mac-II’s has “perfectly imitated” the human mind!’ What’s more though, he does himself one better a few pages later in *Consciousness Explained*, where he shows that not only can any vonNeumannesque machine *imitate* the human mind (I guess that would give him mere Weak AI?) but that this vonNeumannesque machine is actually *superior* to the inefficient stuff nature has given us humans:

Now we are ready to turn this standard idea upside-down. Just as you can simulate a parallel brain on a serial von Neumann machine, you can also, in principle, simulate (something like) a von Neumann machine on parallel hardware, and that is just what I am suggesting: Conscious human minds are more-or-less serial virtual machines implemented—inefficiently—on the parallel hardware that evolution has provided us. (CE 218)

Now there’s a whole argument that once Dennett introduces the concept of *virtual*—why then he can get away with just about any claim he wants, because he can always retort: ‘well, I was talking about a *virtual X*, not a *real X*!!’ but leaving this sort of epistemological loophole aside, let’s see just what he is doing. What he appears to be saying is that thanks to the cleanliness of its architecture and presumably the speed of its processing power—the computer, once it has been correctly wired to properly mimic the parallel architectonics of the human mind—will then be not only in all respects similar, but importantly, from an efficiency standpoint, *superior* to the human mind.

Were we to concede even this point, that the human brain is simply an unnecessarily confusingly wired *virtual* machine which would be much more efficient

were it wired up the way . . . oh, surprise!—the way Cog’s “brain” is wired—we would still have all the arguments advanced by Dreyfus and company with respect to the in fact non-computational nature of consciousness in the human. And these arguments are mostly centered around the question of whether or not any Strong AI scenario could be actually “situated” (in the strong Heideggerian sense) in the way a real, living human being always-already is. However, as I mention just above I don’t have the room here to go into these arguments in any depth, so I would even concede that point—let’s say okay, computers (and computers on wheels, hence ‘situated’ according to the robot-situationists) are *situated* in a world and hence can have their symbols grounded, etc. [Ergo: I will simply give in to Dennett’s brow-beating “I submit that Cog moots the problem of symbol grounding . . .”(PRMCR 5) from the “Practical Requirements” essay] What I will not give into, however, is the question of anything theoretically “mattering” to any present artificial intelligence, even one putatively “embodied” such as any current robot such as Cog.

For I will claim that no robot constructed so far, nor any that we can foresee constructing with current technology, is embodied in the least bit the way we humans are, or even the way clams or houseflies are. This will be the final section of this essay, and it is my hope to tie the argument back into the question of temporality in consciousness and what I am claiming is its concomitant—the *qualia* which are the fauna and flora of our mindscapes.

Further, I believe this is a perfect example of the actual physical instantiation in the real world of a scientist’s scientific hypotheses coming around to bite his or her postulates in the rear as it were. For while it may be somewhat easier to sweep under the

rug or play shell games with such things as qualia and/or the specific temporality of consciousness when it is a question of *interpreting* the actions or statements of a being already assumed to be conscious, it is an entirely different thing if it can be shown that lacking X, consciousness cannot arise in the first place! That is to say, I am claiming that lacking qualia, and in particular a real *embodied* pain complex, Cog will never be able to learn in the way humans do—and since, given the current artificial neural network architecture being employed in robotics *learning* is going to be crucial to the fomenting of true ‘adult-like’ consciousness: the lack of this sort of qualia forms a particularly formidable road-block to the development of human like consciousness in present robots.

Let me modify this slightly however by saying that not all of the Cog team has such a black and white view of the situation, not all necessarily believe that *both* situatedness and embodiment have *already* been achieved in robotics. In particular Rodney Brooks (mentioned above, and interestingly the one actually designing the software on which Cog runs) is a proponent of what has been called “Nouvelle AI”, a school of thought which “. . . is based on the physical grounding hypothesis. This hypothesis states that to build a system that is intelligent it is necessary to have its representations grounded in the physical world To build a system based on the physical grounding hypothesis it is necessary to connect it to the world via a set of sensors and actuators.” (CB 376) Furthermore this school of thought is strongly committed to “the two cornerstones of the new approach to Artificial Intelligence, situatedness and embodiment”. (CB 376)

So far so good as far as Dennett is concerned—sensors and actuators, of course! But the question, if you will, is not so much one of what the robot *should* have in order to

be generally situated and embodied with respect to general theoretical background conditions. The question when it comes to Dennett's takes on these issues is whether Cog *already* has sufficiently fulfilled these conditions.

For as we will see, while Dennett will not deny these requirements for intelligence in any robot—he will take what seems a rather curious tack in submitting that these conditions have *already* all been met! Once again, to quote from the “Practical Requirements” essay:

Another claim that has often been advanced, most carefully by Haugeland (1985) is that nothing could properly “matter” to an artificial intelligence, and mattering (it is claimed) is crucial to consciousness. Haugeland restricted his claims to traditional GOFAI systems, and left robots out of consideration. Would he concede that something could matter to Cog? The question, presumably, is how to weigh the import of the quite deliberate decision by Cog's creators to make Cog as much as possible responsible for its own welfare. Cog will be equipped with some “innate” but not at all arbitrary preferences, and hence provided of necessity with the concomitant capacity to be “bothered” by the thwarting of those preferences, and “pleased” by the furthering of the ends it was innately designed to seek. Some may want to retort: “This is not *real* pleasure or pain, but merely a simulacrum.” Perhaps, but on what grounds will they defend this claim? Cog may be said to have quite crude, simplistic, one-dimensional pleasure and pain, cartoon pleasure and pain if you like, but then the same might be said of the pleasure and pain of simpler organisms—clams or houseflies for instance. Most, if not all of the burden of proof is shifted by Cog, in my estimation. The reasons for saying that something *does* matter to Cog are not arbitrary; they are exactly parallel to the reasons we give for saying that things matter to us and to other creatures. (PRMCR 6)

And here again we see the ‘metaphorics as proof’ propensity of Dennett's which seems to run throughout his work. By using the modifier “exactly”, when he claims parallelism between the “mattering” for an organic, living organism and a robot—he seems to be implying that they are basically the *same thing*. Which is essentially the import of the

whole preceding quote. However, importantly, he does leave himself once again the potential loophole, if pushed, of saying ‘well, I said *parallel* after all, not the *same*!!’ But what is being claimed with respect to pain (a crucial form of “mattering” of course) is exactly an identity relation. And it is with this that I will have no truck. The pain of a clam or a housefly is precisely not like that of a robot in that the clam or housefly, for one, actually *feels* the pain—it is *in* the body as it is *in* the mind (albeit primitive mind)—in a way in which it just never will be for a processor connected via wires and sensors to a mechanical chassis with motors etc. The feeling of pain that is in the clam, housefly, and human is different from that in the robot precisely in the sense that it originates, no matter where in the body, from an area that is *integrally connected* to the *whole organism* via cellular and nervous direct communication. And this communication and solidarity is important precisely to the extent that it serves to promote the *survival* of the *whole* organism—which, by definition, being one integral *whole* cannot lose any one *part* or suffer *damage* to any one *part* without *suffering* thereby an *insult* to the *whole being*. And a robot simply will never have a unified, integral whole being in the way an organic living individual does—and certainly *especially* in its present form of processors wired to a mobile chassis with sensors and actuators.

CHAPTER SEVEN: ANTONIO DAMASIO AND ORGANIC WHOLISM

So to re-enter the question of survival and wholism in an integrated and unified entity, though this time discussing not a robot like Robbie from *Lost in Space* but rather a more realistic and yet not for that reason any less ‘futuristic’ robot in the form of a Boeing 777, I offer the theses of Antonio Damasio on the difference between the way the components of a robot such as a 777 hang together over against how the components of a living organism such as a human being interact, rely upon each other, and yet at the same time encompass their own ipseity and life-story, life-risk. That is, the components of a robot don’t need each other in the same way the living cells of a living organism need each other, and the components of a robot cannot die, as a living cell can die.

As Damasio says in *Looking for Spinoza*:

Some of the components of the aircraft are “animated”—slats and flaps, rudder, air brakes, undercarriage—but none of those components is “alive” in the biological sense. None of those components is made of cells whose integrity depends on the delivery of oxygen and nutrients to *each* of them. On the contrary, *every* elementary part of our organism, *every* cell in the body, is not just animated but living. Even more dramatically, every cell is an individual living organism—an individual creature with a birth date, life cycle, and likely death date. (LS 127)

Here we can see a stark contrast between simple “animation” and that which is truly living. The simply animated does not need nutrition, the simply animated does not need

oxygen to survive. In fact, the term “to survive” has no meaning for the animated, but unliving. The further question of course will be: is that which is made up of unliving components itself living? There are those who believe that since we can now break down all living, organic, carbon-based molecules into their inorganic ‘components’ that hence we have already arrived at the threshold where we can therefore synthesize, reverse engineer if you will, organic-based molecules that are living. However, it is my understanding that there is a great gulf between synthesizing a large carbonic molecule and synthesizing life and something which is living. We have learned to culture blastomeres to create various organs, and we have learned to split cells at very early stages such that copies or ‘clones’ emerge—but neither of these operations as far as I can tell is actually creating life *ex nihilo*. All of this hubbub surrounding the thesis/hope that says since we have ‘completed’ the human genome project and analyzed the last little organic molecule down to its inert, dead, components—why we have basically discovered the secret of life—reminds me eerily of the mediaeval alchemist’s dream of converting base metals into gold and discovering the universal elixir, or elixir of life. Have we not just traded in the dusty dungeon-bound vials and philtres for gleaming florescent-lighted test-tubes and Petrie dishes? The obviously fantastic rabbis’ Shellyean dream of a golem for the not so obviously improbable computer programmers’ Kurzweilean dream of artificial neural nets animating the Cmdr. Datas of the future to give us the best offspring of all—copies of ourselves?

But to return to more mundane matters, and Damasio’s discussion of the Boeing 777:

But the physical matter of the aircraft is not alive, its parts are not made of living cells possessed of a genetic inheritance, a biologic destiny, and a life risk. And even if one were to argue that the plane has an “engineered concern” for its survival, which allows it to preempt the wrong maneuver of a distracted pilot, the blatant difference is inescapable. The plane’s integrated cockpit computers have a concern for the execution of its flying function. Our brains and minds have a global concern for the integrity of our entire living real estate, every nook and cranny of it, and underneath it all, every nook and cranny has a local, automated concern with itself. (LS 128)

Here we can finally see Damasio’s wholism coming to the fore. While programmers can engineer an external ‘concern’ into the computers running a Boeing 777, this concern for ‘survival’ will always remain extrinsic to the being of the Boeing 777—as there is no globality inherent to the structure of such a being. It has no integrity in the way a living conscious being does.

As Damasio says, our thinking components—our “brains and minds” have an intimate and vital concern for the entirety of the “real estate” that comprises our beings. And this is due in no small measure to the fact that there is a web of fluids and of nerves and perhaps even other media of communication which we have yet to discover (though I realize this very phrase must come as dour recidivism and pessimistic contrariness to those who feel that on the kitchen table of the cosmos, we have basically all but completed the puzzle of life—such that we can go on to more “wondrous” questions: the hot coco I suppose for those like Dennett who can now turn to more nifty issues such as will the robot I download my consciousness into be male or female? Will it look more like Brad Pitt or Angelina Jolie? And what video games should I bury with myself in the Egyptian funereal crypt *cum* robot laboratory of the future for my postmodern after-life

with myself?) connecting “every nook and cranny” of our living whole organism with each other.

It is this web-like structure of the fluids and the nerves and synapses that insures the wholism of which I have been speaking. Without this wholism, and without this democratically shared life risk (though like all good democracies I should add the caveat that some parts of this animal are more equal than others—suffering gross insult to optical organs or crucial sensory apparatuses is not equally weighted to other less vital areas of potential loss) we would have to rely upon extrinsic algorithms to dictate as it were that which is important and that which is not important for our survival as unities in the cosmos. This touches on the “engineered concern” above of course mentioned by Damasio in the case of the Boeing 777. Yet I believe it applies equally as well to any creature whose parts do not equally, or somewhat equally, share the life risk that goes along with being alive, as opposed to being dead—and hence risking nothing.

And furthermore, along with this life risk, can be seen a certain subjectivity if you will, even on the smallest level of the living organism. This is where wholism most clearly comes into conflict with neohomuncularism and the theory that living conscious beings are simply complex concatenations of uncaring automata that somehow come up with caring structures once they combine into the super-structure level of “fame in the brain” and what looks to the innocent onlooker and bystander like consciousness and a conscious being (with zombiehood always lurking in the shadows, as numerous famous base-line rallies between Dennett and Searle have well pointed out!).

Or as Damasio puts the same point in more scientific, physiological, and objective terms:

Notice something quite curious and also chronically overlooked: the nerve sensors that convey the requisite information to the brain and the nerve nuclei and nerve sheaths that map the information inside of it are living cells themselves, subject to the same life risk of other cells, and in need of comparable homeostatic regulation. These nerve cells are not impartial bystanders. They are not innocent conveyances or blank slates or mirrors waiting for something to reflect. (LS 129)

Even and especially our nervous system is alive. Therefore sharing a life risk with the other cells, these nerves are not impartial to the fate of every other cell in the conjunct which consists of the living organism.

Yet, not how similar “innocent conveyance”, “blank slate” and especially “mirror” is to the stupid homunculi nested within more and more stupid homunculi till they bottom out in some neohomuncular structure that is composed of homunculi so stupid they may as well be an opened or closed “gate” in a microchip—which clearly is exactly where Dennett is going with his homuncular crowd clamoring for attention, fame, and consciousness.

Therefore, I see a strict divide at the most basic structural and component level between the conscious being which is alive, and hence composed of living cells throughout (which reads equally both ways to my lights: which is composed of living cells throughout and hence alive) and the conscious being which is composed of stupid, dead, inert neohomunculi, at least in the brain and mind.

One way this divide plays out on the larger scale, or “macro-scale” as Dennett would put it in more scientific terms—is in that highly contested and argued over area known to us conscious beings as “feeling”. And of course it is in the area of feelings and

emotions, and their impact on if not necessary substratum of any ratiocinative process in conscious beings for which Damasio's work has been seen to be most ground-breaking and withal what he is most famous for in the larger community of thinkers.

And so it is perhaps not surprising that he concludes his treatment of the Boeing 777 in *Looking for Spinoza* with a comparison of feeling structures in the Superliner and feeling structures in the living conscious being.

Confronting those who would grant consciousness in a human manner to Boeing 777's and by extension any robot driven by modern highly complex ANN's Damasio ultimately winds up with:

My answer to the last naysayer is that the 777 is unable to feel anything like human feelings because, among many other reasons, it does not have an equivalent interior life to be managed, let alone portrayed. (LS 130)

And there we have it: Damasio comes down clearly on the side of the divide which says that neither COG nor the Boeing 777 have feelings, which would include pain of course, since they have no interior life.

Now, one might ask whether a clam has an interior life, or a housefly for example. I guess this is a whole field of analytic philosophy, to which question I believe the answer is usually "the question is moot". Moot for the clam, moot for the housefly, moot for the bat even. Moot because we will never know the answer to that question, because we can never ask the clam, the housefly, the bat how it feels, let alone understand any potential response.

Which would happily throw us back into the world whose divine arbiter and ultimate judge of consciousness would be the Heterophenomenological Method. Which, not surprisingly, only works when communication can be assured. And that means, for hard-liners like Dennett—when the communication is unambiguous, and verbal.

So what about the poor bat, clam, housefly? Well, as Dennett claims in a recent article from 2006, just because these little fellows aren't conscious like us sophisticated and superior humans—why that doesn't mean there are any moral/ethical implications as to how we should treat them. No, that's a totally different question. I suppose that question is dealt with by Dr. Jones, the Kantian, two offices down the hall on the left.

In any case, I would submit that whether or not we can ever prove that a clam feels pain or is conscious (though I have to admit the desperate ruckus made by steaming lobsters might lead one to think they are not feeling nothing—and I believe a recent 'scientific' study pretty much concluded that lobsters *do* in fact feel pain, but perhaps they are several levels higher in the great chain of being and so don't really apply to Dennett's clam-cartoon-pain argument?) the question of mattering is not nevertheless trumped. That is, a clam's supposed cartoon pleasure and pain has very little to do with how we might begin to view ourselves and in fact all living creatures if the 'fact' at heart were that they were made up of a bunch of dead components. Dennett clearly implies that it does not and would not matter. I disagree. Sometimes people mention the case of Commander Data from the *Star Trek* series. As in: you care about whether Data lives or dies, right? And why couldn't we build a robot like that in two hundred years time? Well, maybe so. But no matter how seamless the facsimile, I believe that such a robot would still be either dead or alive. Unless, that is, in two hundred years time we have discovered

the secret of life itself. But I don't think that Dennett is arguing for this case. Rather, I think he's simply arguing that life itself is not life itself. That is, we living things are not really alive in the way unsophisticated thinkers have thought in the past. As in, there is no categorical difference between the living and the dead. The living are simply shambling mounds of dead inert stuff that thanks to evolution happen to hang together in ways that we have heretofore described as alive. But we are really dead, at the core, we just "act" as though we are alive—and so under a "teleofunctionalist" optic, we can call "alive" if we still need this pre-Dennettean-Copernican-Revolution-in-consciousness crutch to help us get along in unsophisticated world peopled with troubadour love and vital forces which separate the quick from the dead.

To close this chapter on Damasio I would like to return to *Looking For Spinoza* for a moment, and see if we can't detect one of the many reasons Damasio decided to center his third work dealing with the emotive bases of ratiocinative thought around the thought of Spinoza. For Damasio, one of Spinoza's key insights was the *conatus* (a classical Latin term) which symbolized each individual organism's constant struggle to survive, and in Damasio's interpretation, even each organism's ongoing effort to maximize its efficiency in its navigation through the world and hence for lack of a better term its happiness, if not joy. As we will see shortly, for Damasio this conatus function operates not on solely the 'macro-scale' of the entire organism or individual, but equally on the 'micro-scale' of the cells which compose said organism, individual. And if this cellular *conatus* if you will is such a crucial structure for Damasio, then I just don't see how servo-motors and microchips wired together, no matter how elegantly and

mimickingly they approximate pain and hence the survival instinct—can combine to form an over-all *conatus*-driven entity like a living conscious being?

Once again, I leave you with the words of Damasio himself:

But now we can go deeper and discover a finer origin underneath that level of description: the many cells that make those body parts and exist both as individual organisms with their own *conatus* and as cooperative members of the regimented society we call the human body, held together by the organism's own *conatus*. (LS 131)

CHAPTER EIGHT: WHOLISM, PAIN, AND THE ENVIRONMENT

Now I am not claiming in this essay that robots in the future may not be able to approximate and copy the outward flourishes of a human-like consciousness status. However at the present moment they are far from being even near to fulfilling this requirement. And while Dennett appears to be staunchly propounding that the “mattering” question in the form of pain and pleasure has already been answered by COG—he does give himself the beginning of another loophole in another part of “The Practical Requirements for Making a Conscious Robot” where he discusses, as I’ve mentioned above—the ‘legality’ of grafting organic material onto any future Galatea which may warrant such with respect to questions of ‘efficiency’:

If making a conscious robot were conceived of as a sort of sporting event—like the America’s Cup if somebody were to invent some sort of cheap artificial neural network fabric that could usefully be spliced into various tight corners in a robot’s control system, the embarrassing fact that this fabric was made of organic molecules would not and should not dissuade serious roboticists from using it—and simply taking on the burden of explaining to the uninitiated why this did not constitute “cheating” in any important sense. (PRMCR 1-2)

I have found that it is usually a good marker that something at least somewhat tricky is going on in Dennett’s arguments when a long involved metaphor appears right where you would expect the meat of the argument—and this particular instance does not appear any

different to me. I don't see who's accusing Dennett of "cheating" here, and at this point Dennett doesn't even have a straw man in whose mouth to place the argument.

However, and more interesting than these facets, is what I see as an attempt to open the theoretical possibility of 'splicing' or 'grafting' as I've chosen to call it, of various organic components onto the "metal, silicon chips, glass, plastic, rubber and such" (PRMCR 2) which heretofore comprise the modern robot. And what I am submitting is that this is in fact a question of Dennett's foreseeing the possibility that his pain argument is not going to stand the test of time. Perhaps the parallelism imputed to Cog with respect to "mattering", the fact that programmers may always have to "design" in the "innate" predispositions and "preferences" which comprise the basis for Cog's pain, might all add up to the fact that Cog will never have a sense of survival and wholism which even the most 'stupid' of clams inherits the day it's born on some dumb rock.

And maybe he foresees as well that it will be via synthesizing a body of some sort that ultimately those such as COG may one day achieve the wholism, the organic solidarity, and hence the sense of survival which is crucial not only for self-preservation, but consciousness itself⁹. As the bio-robotics researchers Sharkey and Ziemke claim in their illuminating 1998 article in *Connection Science*, "A Consideration of the Biological and Psychological Foundations of Autonomous Robotics", a robot will not even be able to learn, something Dennett has admitted will be crucial for Cog were he to attain adult

⁹ I jump ahead in this footnote to another passage on robot embodiment written by the team about to be discussed in the body of the text, cited here for reasons of space-conservation: "Thus robotics faces a fundamental dilemma: a robot that is built and equipped with a body the conventional way will always be heteronomous in the sense that its body and its environmental embedding are not its own, but designed; thus its behaviour, much like a wishful attribution, derives meaning from the intentions of its designers or observers." (CB 388)

human-like consciousness, unless it somehow is able to attain a body much more like ours than anything any present robot possesses. As they say in section three of the essay, “Robot Learning and Evolution”:

For our purposes, the most important difference between robot reinforcement learning and animal learning is that animals have living bodies and have behavior completeness. The body is important if reinforcement learning is to have realism. The robot does not experience pleasure or pain as does the animal; it experiences only modifications to its ANN weights or to its program code during learning. Thus the psychological meaning of the behavior is lost by the absence of bodily needs and changes. (CB 381)

While this flies directly in the face of Dennett’s proclamations with regard to COG, it becomes more poignant in section four of the essay “Minds, Bodies, Robots” where Sharkey and Ziemke claim that it’s the very lack of integral solidarity in the robot’s body that will lead to the lack of anything “mattering” for the robot, quite reminiscent here of Damasio’s proclamations in chapter seven concerning the Boeing 777:

Without an integrated body it [the robot] does not experience pleasure or pain in reinforcement learning; there are only weight changes or program changes. The actual putative ‘experience’ of a robot undergoing reinforcement learning is the same for both reward and punishment. The organism, on the other hand, is driven by its bodily aversions and needs. (CB 385)

And with an eye toward the future, and perhaps the counter-arguments from those like Dennett who would claim that ANN’s, sensors, and actuators are all the ingredients that are needed for consciousness—Sharkey and Ziemke conclude:

Certainly we can equip robots with sensors and effectors, and some hold that this provides them with a rudimentary form of Umwelt (cf. Brooks, 1986a, 1991b). We know, however, very little about how to design coherent artificial organisms, which have the solidarity of living systems and are one with their world. The key problem for the development of a truly intelligent robot is how to synthesize such a coherent self. (CB 388)

If I haven't already made it clear—it is not so much my position that robots will never be like us, or maybe even “just” like us, even with respect to their consciousnesses. However, I have zeroed-in on the “mattering” and “pain” questions as they relate to robots and humans respectively because however asymptotically we may see the outward flourishes of the robots we create approach being just like us, until we discover the secret of life, their parts will not be alive. Further, not in an America's Cup way of grafting biological bits into an otherwise mechanical creature, but in a rather more tricky way—I see this neohomuncularism and using verbal reports as ultimate touchstone of consciousness as a kind of cheating—not to get past the consciousness question, as Searle submits, but rather more fundamentally to get past the LIFE question.

To bracket for the moment the question of whether or not the neohomuncular model is ultimately correct, true, whether the heterophenomenological method is indeed the best if not only way of testing for consciousness and figuring out what form it needs must have—if at the end of the day all we have are really stupid little automata (the Dennettean homunculi nested within themselves) and the words that describe our thoughts (the verbal reports)—then the next logical step to take is to say that these (living, cellular and neuronal) automata can be replaced by microchips, these verbal reports can be spat out by any concatenation of ‘sub-processes’ which combine in a

clever enough manner to be clever like us. But to step back from the back and forth debate—which can not be decided conclusively for now anyway—as to whether it really is little cellular automata that combine in a multiple drafts kind of way to create consciousness, we don't have the medical science know-how to determine the question now, we don't have the computer science know-how to copy these automata right now even if we wanted to—we can see that on a deeper level what has been jumped over, and effectively elided, is life itself. Life no longer matters in this picture. And the claim is that it doesn't matter that life doesn't matter—as long as that which is not life acts just like that which is life, does not wash.

Ever since the Enlightenment, there has been a certain strain in human thought that seems to be goaded on by a giddiness inspired by our *knowing so much more* than we ever have about the cosmos and ourselves that wants to pretty much explain *everything*, and *now!* With the tools present at hand. We can analyze the chemical composition of all basic components of living beings, we can describe the human genome, therefore we know life itself and can reproduce it.

I am not coming from some deistic perspective, I don't care what or who goes into making life what it is, but at the very least I think the categorical difference between the living and the dead, the truly alive and the merely animate, still has some 'work' to do for us, to put it in American Pragmatist terms. If the coffins coming back from the Taiwan War in 2015 draped in American flags contain robot soldiers, I don't think I would feel the same as I would if they contained the son of the guy who owns the gas station down the street.

BIBLIOGRAPHY

Bergson, Henri. *Creative Evolution*. Translated by Arthur Mitchell. New York: Dover Publications, 1998.

———. *Matter and Memory*. Translated by W. Scott Palmer and Nancy Margaret Paul. New York: Zone Books, 1991.

———. *Time and Free Will*. Translated by F.L. Pogson. London: George Allen & Company, 1913. Cited as TFW.

Blume, Harvey. "The Digital Philosopher." In *Digital Culture*, September 12 1998; available from www.theatlantic.com/unbound/digicult/dc981209.htm. Cited as AO.

Borrett, Donald, Sean Kelly, and Hon Kwan. "Phenomenology, Dynamical Neural Networks and Brain Function." *Philosophical Psychology* 13, no. 2 (June 2000): 213-29.

Damasio, Antonio. *Descartes' Error*. New York: Harcourt, 1995.

———. *The Feeling of What Happens*. New York: Harcourt, 1999.

———. *Looking For Spinoza*. New York: Harcourt, 2003. Cited as LS.

Dennett, Daniel. "Are We Explaining Consciousness Yet?" Final Draft for *Cognition*, August 27, 2000. Cited as AWECY.

———. *Consciousness Explained*. Boston: Little, Brown and Company, 1991. Cited as CE.

———. “Is Perception the ‘Leading Edge’ of Memory?” In A. Spadafora (ed.) *Iride: Luoghi della memoria e dell’oblio*, anno. VIII, n. 14, April 1995, pp. 59-78 [Italian]. Cited as PLEM.

———. “The Practical Requirements for Making a Conscious Robot.” From Daniel C. Dennett’s personal website. Also available at www.ecs.soton.ac.uk/~harnad/Papers/PY104/dennet.rob.html. Cited as PRMCR.

———. “Review of *The Rediscovery of the Mind*, by John Searle.” *Journal of Philosophy* 60, no. 4 (April 1993): 193-205.

Dreyfus, Hubert L. *What Computer Still Can’t Do: A Critique of Artificial Reason*. Cambridge: MIT Press, 1999.

James, William. *Principles of Psychology*. Vol. 1. New York: Holt, 1890. Cited as PP.

Merleau-Ponty, Maurice. *Phenomenology of Perception*. Translated by Colin Smith. New York: Routledge, 1996.

Mullarkey, John, ed. *The New Bergson*. Manchester and New York: Manchester University Press, 1999.

Neisser, Joseph Ulric. “On the Use and Abuse of Dasein in Cognitive Science.” *Monist* 82, no. 2 (April 1999): 347-62.

Searle, John R. *The Mystery of Consciousness*. The New York Review of Books: New York, 1997. Cited as MC.

Sharkey, Noel E., and Tom Ziemke. “A Consideration of the Biological and Psychological Foundations of Autonomous Robotics.” *Connection Science* 10, no. 3/4 (December 1998): 361-92. Cited as CB.