

## **Chapter Five: Multiple Drafts Versus the Cartesian Theater**

### **1: THE POINT OF VIEW OF THE OBSERVER**

In this chapter we get a look at Dennett's model for the mental processes we call consciousness. He will claim that "some of the most perplexing paradoxes of consciousness arise because we cling too long to a good habit of thought, a habit that usually keeps us out of trouble." Just like the pleasure boater who charts his safe course by aiming at a buoy, and then distractedly runs right into it.

There is no question that where there is a consciousness mind there is also a point of view. The flash and then later bang of the fireworks is an example of this point of view. But what happens when we close in on the observer, and try to locate this "point of view" within the brain? The simple assumptions that worked so well on larger scales begin to breakdown on closer examination. Dennett maintains that there is no single point in the brain where the information funnels in, where "it all comes together" for "central processing." This has some far from obvious and quite counter-intuitive consequences.

Descartes thought hard about this problem, but he decided that the brain DID have a center: the pineal gland. He proposed that for a person to be conscious of something, traffic from the senses had to arrive at this location, where it then "caused a special- indeed magical- transaction to occur between the person's material brain and immaterial mind". And although today no one seriously believes there is a physical center to the brain, a Cartesian bottleneck of dualism, even sophisticated materialists often forget that once Descartes's ghostly *res cogitans* is discarded, there is no longer a role for a centralized gateway, or indeed for any functional center to the brain. The pineal gland nor any other physical feature, is not the fax machine to the soul, nor is it the Oval Office of the brain. There is no observer gland inside the brain.

Although we may discard the anatomical baggage of Descartes dualism, it is harder to discard the idea of a "central processing" area of the mind itself. Dennett calls the idea that there is a crucial finish line or boundary somewhere in the brain, marking a place where the order of stimuli arrival equals the "order of presentation", a sort of theater of "Cartesian materialism." The Cartesian Theater is a metaphor that seems like an innocent extrapolation of the familiar and undeniable fact that, for everyday and macroscopic time intervals, we can indeed order events into the two categories "not yet observed" and "already observed".

But for very short intervals of time, the "point of view" of the observer must be spread over a large volume in the observer's brain, so the observer's own subjective sense of sequence must be determined by something other than "order of arrival". As Dennett says: "If A beats B to one finish line, but B beats A to another finish line, which result fixes subjective sequence in consciousness?" (pp. 107-108)

We need to stop thinking of the brain as if it had a single functional summit or central point. To do so we need a new model and the Multiple Drafts Model is one possible candidate. Another is William Calvin's 1989 model of "Scenario Spinning". Next Dennett introduces the Multiple Drafts Model in Part 2.

### **2: INTRODUCING THE MULTIPLE DRAFTS MODEL**

According to the Multiple Drafts model, all varieties of perception- indeed, all varieties of thought or mental activity- are accomplished in the brain by parallel, multitrack processes of interpretation and elaboration of sensory inputs. Information entering the nervous system is under continuous "editorial revision. For instance, since your head moves a bit and your eyes move a lot, the images in your retinas swim about constantly, rather like the images of home movies taken by people who can't keep the camera from jiggling. But that is not how it seems to us.

. . . These editorial processes occur over large fractions of a second, during which time various additions, incorporations, emendations, and overwritings of content can occur, in various orders. . . This much is recognized by virtually all theories of perception, but now we are poised for the novel feature of the Multiple Drafts model: Feature detections or discriminations only have to be made once. That is, once a particular "observation" of some feature has been made, by a specialized, localized portion of the brain, the information content thus fixed does not have to be sent somewhere else to be rediscriminated by some "master" discriminator.

. . . These spatially and temporally distributed content-fixations in the brain are precisely locatable in both space and time, but their onsets do not mark the onset of consciousness of their content. It is therefore always an open question whether any particular content thus discriminated will eventually appear as an element of conscious experience, and it is a confusion. . . to ask when it becomes conscious. These distributed content-discriminations yield, over the course of time, something rather like a narrative stream or sequence, which can be thought of as subject to continual editing by many processes distributed around in the brain, and continuing indefinitely into the future.”. (pp. 111 – 113)

#### Later Additions

In the years following publication of the book, progress in cognitive neuroscience and criticism of his model from scientists and philosophers compelled Dennett to make some amendments. In an excellent essay called “Are we explaining consciousness yet?” which closed out an equally excellent volume of papers called *The Cognitive Neuroscience of Consciousness*, Dennett discussed the gathering consensus around the global neuronal workspace model of conscious and in light of that model introduced the concept of “fame in the brain”.

The global workspace model can be interpreted as a more neurologically grounded and less metaphorical version of the Multiple Drafts model, so it’s worth taking a look at. Dennett quotes from a summary of the model by Stanislas Dehaene and Lionel Naccache (same volume) and offers a series of elaborations and friendly amendments. First I’ll provide the summary, then Dennett’s elaborations. The numbers in bold point to the sentences Dennett comments on later.

The summary:

“At any given time, many modular (1) cerebral networks are active in parallel and process information in an unconscious manner. An information (2) becomes conscious, however, if the neural population that represents it is mobilized by top-down (3) attentional amplification into a brain-scale state of coherent activity that involves many neurons distributed throughout the brain. The long distance connectivity of these "workplace neurons" can, when they are active for a minimal duration (4), make the information available to a variety of processes including perceptual categorization, long-term memorization, evaluation, and intentional action. We postulate that this global availability of information through the workplace is (5) what we subjectively experience as a conscious state.”

Dennett’s remarks:

(1) Modularity comes in degrees and kinds; what is being stressed here is only that these are specialist networks with limited powers of information processing.

(2) There is no standard term for an event in the brain that carries information or content on some topic (e.g., information about color at a retinal location, information about a phoneme heard, information about the familiarity or novelty of other information currently being carried, etc.). Whenever some specialist network or smaller structure makes a discrimination, fixes some element of content, "an information" in their sense comes into existence. "Signal," "content-fixation," (Dennett, 1991), "micro-taking," (Dennett and Kinsbourne, 1992) "wordless narrative" (Damasio 1999), and "representation" (Jack and Shallice) are among the near-synonyms in use.

(3) We should be careful not to take the term "top-down" too literally. Since there is no single organizational summit to the brain, it means only that such attentional amplification is not just modulated "bottom-up" by features internal to the processing stream in which it rides, but also by sideways influences, from competitive, cooperative, collateral activities whose emergent net result is what we may lump together and call top-down influence. In an arena of opponent processes (as in a democracy) the "top" is distributed, not localized. Nevertheless, among the various competitive processes, there are important bifurcations or thresholds that can lead to strikingly different sequels, and it is these differences that best account for our pretheoretical intuitions about the difference between conscious and unconscious events in the mind. If we are careful, we can use "top-down" as an innocent allusion, exploiting a vivid fossil trace of a discarded Cartesian theory to mark the real differences that that theory misdescribed.

(4) How long must this minimal duration be? Long enough to make the information available to a variety of processes-that's all. One should resist the temptation to imagine some other effect that needs to build up over time,

because . . .

(5) The proposed consensual thesis is not that this global availability causes some further effect or a different sort altogether—igniting the glow of conscious qualia, gaining entrance to the Cartesian Theater, or something like that—but that it is, all by itself, a conscious state. This is the hardest part of the thesis to understand and embrace. In fact, some who favor the rest of the consensus balk at this point and want to suppose that global availability must somehow kindle some special effect over and above the merely computational or functional competences such that global availability ensures. Those who harbor this hunch are surrendering just when victory is at hand, I will argue, for these "merely functional" competences are the very competences that consciousness was supposed to enable.

The "global availability" of information in the workspace model can be understood as a sort of competition for cerebral "clout" or "fame" as networks of neurons representing the circulating "drafts" have increasing or decreasing mobilization of resources in the brain.

When processes compete for ongoing control of the body, the one with the greatest clout dominates the scene until a process with even greater clout displaces it. In some oligarchies, perhaps, the only way to have clout is to be known by the King, dispenser of all powers and privileges. Our brains are more democratic, indeed somewhat anarchic. In the brain there is no King, no Official Viewer of the State Television Program, no Cartesian Theater, but there are still plenty of quite sharp differences in political clout exercised by contents over time. In Dehaene and Naccache's terms, this political difference is achieved by "reverberation" in a "sustained amplification loop", while the losing competitors soon fade into oblivion, unable to recruit enough specialist attention to achieve self-sustaining reverberation."

A complete theory of consciousness must explain two things—how "fame" is achieved, and what happens in its wake. But, though "One may postulate activity in one neural structure or another as the necessary and sufficient condition for consciousness, . . . one must then take on the burden of the explaining why that activity ensures the political power of the events it involves—and this means taking a good hard look at how the relevant differences in competence might be enabled by changes in status in the brain."

Keeping these modifications to the original theory in mind, we move on to specific examples.

Motion picture and television rely on creating apparent motion from a rapid succession of "still" pictures, and since the dawn of the motion picture age, psychologists have studied this phenomenon. The simplest case just has two dots blinking on and off, and if the two dots are separated by as much as 4 degrees and briefly blinked, they will appear to be a single dot moving. A great applet to play with this is here (requires JAVA):

<http://epsych.msstate.edu/descripti...ion/apmove.html>

[Edit: The original links no longer work, but the following one is active at this time (10-14-2007)]

[http://www.philosophy.uncc.edu/faculty/phi/Phi\\_Color2.html](http://www.philosophy.uncc.edu/faculty/phi/Phi_Color2.html)

Try slowing the display down and increasing the distance between the dots until they appear to be just stationary and blinking as opposed to moving. Previous models trouble explaining this, especially a simple modification where two different colored blinking dots appear to be a single dot moving back and forth— but apparently changing color in mid-trajectory! See this URL for an example:

[http://www.uncc.edu/colleges/arts\\_a...Phenomenon.html](http://www.uncc.edu/colleges/arts_a...Phenomenon.html)

For example, even in single shot trials, subjects would see a red dot moving, turn green while it was moving and then a fraction of a second later, a green dot stopped. But nothing actually moved, only a red dot was shown and then slightly to one side a green dot a little later. Here's an example of random Phi motion and color change:

<http://www.cs.tufts.edu/~kreisman/phi/index.html>

[Edit: the following link explains the difference between phi and beta effects (10-14-2007)]

<http://www2.psych.purdue.edu/Magniphi/>

How could the brain, if the second dot was already in our consciousness, go back and change the color of the dot, seemingly in mid-flight? If it seems to you, gentle reader, that (although you experienced red, then red turning green and then green), the conscious experience of the whole event must surely have been delayed until after the green spot was (unconsciously?) perceived. If that seems reasonable then you are still locked in the Cartesian Theater.

### **3. ORWELLIAN AND STALINESQUE REVISIONS**

Two obvious solutions have been proposed to account for these experimental observations, one assumes a revision of perceptions and the other assumes a revision of our memories. Dennett calls these the Stalinesque or Orwellian explanations, but after a detailed analysis (which I will leave to those of you that are interested), he finally notes that both of these revisionist points of view depend on the subject having a single time sequenced point of view, which as we have already noted, is based on the idea that these experiences are presented to something (like a Cartesian theater).

In either case both models can account for the observations. One says they are the results of unconsciously discriminated contents, while the other says they are the result of consciously discriminated but forgotten contents. In either case, subjects should be unable to tell the difference between misbegotten experiences and immediately misremembered experiences. The two explanations just locate the mythical "place that it all happens" differently.

The point of the Multiple Drafts model is that there is no need to locate the "place where it all happens". If one wants to settle on some moment of processing as the moment of consciousness, it has to be arbitrary. There is no functional differences between that process and any prior stages and revisions that might be called unconscious or preconscious.

### **4. THE THEATER OF CONSCIOUSNESS REVISITED**

Dennett has argued that because of spatiotemporal "smearing" of what we call our "experience" in the brain, all the evidence there is or even could be, fails to distinguish between the perceptual revision (Stalinesque) or memory revision (Orwellian) mechanisms of consciousness (for example the Phi experiment). Therefore Dennett argues that there is no difference. But has Dennett pulled a fast one on us and left out the possibility that there might be facts of the matter that are unreachable by science, even if we include Dennett's "heterophenomenological" 3rd person reports of our experiences?

As some have argued, the obvious fact of our immediate conscious experience, consist of such unreachable facts. It might seem that while arguing against dualism (let's get rid of that spook stuff out of here), Dennett has "spirited away" something that Descartes was actually right about; a functional place where the items of phenomenology ... are "projected".

As Nelson Goodman (1978) argues about Phi perception, that whether the process involves a delay of perception or misremembered perception: "each of the intervening places along a path between the two flashes is filled in... with one of the flashed colors rather than with successive intermediate colors." Dennett argues that there is another possibility: maybe the brain doesn't actually have to go to the trouble of "filling in" anything. Because there is nobody in your head "looking".

As stated before, once a discrimination has been made, it does not have to be made again; the brain just adjusts to the conclusion that is drawn, making the new interpretation of the information available for the modulation of subsequent behavior and experience. Why should the brain bother to "produce" the "intervening places" in any case? Why can't the brain just conclude that there was intervening motion and insert that retrospective conclusion into the processing stream?

Dennett calls these speculations to a halt by asking whether "anyone" is viewing all this "filling in". If the Cartesian Theater is not viable, then so is the idea of an audience of the Cartesian Theater for all this 'filling in' and viewing of the "processing stream". Dennett agrees with Goodman that retrospectively the brain creates the content (the

judgment) that there was intervening motion, and then this content is made available to govern activity and leave it's mark on memory. But the Multiple Drafts Model goes one step further to claim that the brain does not go to the trouble of "filling in" the blanks. The judgment has already been made so the brain can get on with other tasks.

Others (Kolers), have taken these experiments to indicate a much more radical metaphysics. That an experience from one time is "projected" to another time. This metaphysics is similar to Thomas Hobbes and other early thinkers on consciousness, because it seemed so obvious: after light from an object struck the eye and produced there a kind of "motion" that led something to "rebound" back out into the world. After all, Hobbes claimed, don't we "see" the colors out there? Out there on the surfaces of objects?

Similarly, when you stub your toe, signals travel up to your brain's pain centers which then "projects" the pain back down to your toe where it belongs. After, that's where you 'feel" the pain, isn't it? As recently as the 1950's this idea was still taken seriously enough to provoke a British psychologist to write an article demolishing it.

Now the "projection" in time that some have argued to explain the Phi experiment, is not a projection into physical space. Even when we hear a stereo recording "coming from a place in between the two speakers in a room, we know that the sound isn't really coming from between the speakers, it just seems that way to us." Some might say, the guitar isn't being projected into physical space, it's being projected into say, "phenomenological space."

So where is this "phenomenological space"? Is it a physical place inside the brain? Maybe it's just a "logical space" inside the brain. Not a real space, but a kind of "experiential space". The point being that the representation of the experience is not the experience in the brain. Representations of space does not use "space in the brain", and the representations of time does not use "time in the brain"- at least at the short time scales at which our perceptions function.

Why do people cling to an "experiential" projection in the brain as actually existing as "spatial" or "temporal" properties in our brains? Because that's how the world outside our brains is. It's intuitive. Dennett's point is that these "spatial" or "temporal" things do not need to be represented by space or time in our brains. Pain doesn't need to be represented by "pain in the brain" and color doesn't need to be represented by "color in the brain."

The astronomer's rule applies to consciousness, "If you didn't write it down, it didn't happen". The Multiple Drafts Model would say that "writing it down" in memory is criterial for consciousness. There is no reality of conscious experience independent of the effects of these various vehicles of content and discrimination on subsequent actions (and of course on memory). Dennett asks: "Consider if something happened in my perception, but it left its trace on me for only "a millionth of a second," as in the Ariel Dorfman epigram. Whatever could it mean to say that I was, however briefly and ineffectually, conscious of it? If there were a privileged Cartesian Theater somewhere, at least it could mean that the film was jolly well shown there even if no one remembers seeing it. (So there!)

"The Cartesian Theater may be a comforting image because it preserves the reality/appearance distinction at the heart of human subjectivity, but as well as being scientifically unmotivated, this is metaphysically dubious because it creates the bizarre category of the objectively subjective- the way things actually, objectively seem to you even if they don't seem that way to you!"(p. 132)

The Multiple Drafts model could be considered a sort of first-person operationalism, for it denies the possibility in principle of consciousness of a stimulus, in the absence of the subject's belief in that consciousness.

Opposition to this type of operationalism usually appeals to facts beyond the possibility of observation, but the objection backfires: "Just because you can't tell, by any preferred way, whether or not you were conscious of X, that doesn't mean you weren't. Maybe you were conscious of X but just can't find any evidence for it!" Dennett asks: Does anyone, upon reflection really want to claim this? Putative facts about consciousness that swim out of reach of both "outside" and "inside" observers are strange facts indeed.

Some people presume that intuitively we can observe ourselves judging things, like the Phi experiment, as a result of those things seeming to be. No one however, has ever observed any such thing "in their phenomenology" because such a fact about causation would be unobservable (as Hume noted long ago). In other words: do you judge that the red spot moved right and changed color because it seemed to you to do so, or does it seem to you to have moved

because that is your judgment?

These kinds of Cartesian Theater ways of thinking about our perceptions are popular because it's a place where "seemings" can happen in addition to the "judging". But postulating a "real seeming" in addition to the judging or "talking" expressed in the subject's report is multiplying entities beyond necessity. It's even multiplying entities beyond possibility because where else could this "seeming" be taking place unless in a Cartesian Theater? There is not only no audience in brain, there is no "presentation of seeming" either.

## 5. THE MULTIPLE DRAFTS MODEL IN ACTION

Dennett says that: "Visual stimuli evoke trains of events in the cortex that gradually yield discriminations of greater and greater specificity. At different times and different places, various "decisions" or "judgments" are made; more literally, parts of the brain are caused to go into states that discriminate different features, e.g., first mere onset of stimulus, then location, then shape, later color (in a different pathway), later still (apparent) motion, and eventually object recognition. These localized discriminative states transmit effects to other places, contributing to further discriminations, and so forth." (p. 134)

So of course we must ask: Where does it all come together for consciousness?

"The answer is: Nowhere. Some of these distributed contentful states soon die out, leaving no further traces. Others do leave traces, on subsequent verbal reports of experience and memory, on "semantic readiness" and other varieties of perceptual set, on emotional state, behavioral proclivities, and so forth. Some of these effects — for instance, influences on subsequent verbal reports — are at least symptomatic of consciousness. But there is no one place in the brain through which all these causal trains must pass in order to deposit their content 'in consciousness.'" (p. 135)

Dennett goes on to say:

As soon as any such discrimination has been accomplished, it becomes available for eliciting some behavior, for instance a button-push (or a smile, or a comment), or for modulating some internal informational state. For instance, a discrimination of a picture of a dog might create a "perceptual set" — making it temporarily easier to see dogs (or even just animals) in other pictures — or it might activate a particular semantic domain, making it temporarily more likely that you read the word "bark" as a sound, not a covering for tree trunks. As we already noted, this multitrack process occurs over hundreds of milli-seconds, during which time various additions, incorporations, emendations, and overwritings of content can occur, in various orders. These yield, over the course of time, something rather like a narrative stream or sequence, 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 and 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 — and some will make no contribution at all — 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. (p. 135)

Dennett goes on to explain what information we might obtain by "probing" this flow of various threads:

Probing this stream at various intervals produces different effects, precipitating different narratives — and these are narratives: single versions of a portion of "the stream of consciousness." If one delays the probe too long, the result is apt to be no narrative left at all. If one probes "too early," one may gather data on how early a particular discrimination is achieved in the stream, but at the cost of disrupting the normal progression of the stream.(pp. 135-136)

"Just what we are conscious of within any particular time duration is not defined independently of the probes we use to precipitate a narrative about that period. Since these narratives are under continual revision, there is no single

narrative that counts as the canonical version, the "first edition" in which are laid down, for all time, the events that happened in the stream of consciousness of the subject, all deviations from which must be corruptions of the text. But any narrative (or narrative fragment) that does get precipitated provides a "time line," a subjective sequence of events from the point of view of an observer, that may then be compared with other time lines, in particular with the objective sequence of events occurring in the brain of that observer. (pp. 135-136)

Dennett sums up the chapter with some observations that philosophers have discussed before:

You have probably experienced the phenomenon of driving for miles while engrossed in conversation (or in silent soliloquy) and then discovering that you have utterly no memory of the road, the traffic, your car-driving activities. It is as if someone else had been driving. Many theorists (myself included, I admit) have cherished this as a favorite case of "unconscious perception and intelligent action." But were you really unconscious of all those passing cars, stop lights, bends in the road at the time? You were paying attention to other things, but surely if you had been probed about what you had just seen at various moments on the drive, you would have had at least some sketchy details to report. The "unconscious driving" phenomenon is better seen as a case of rolling consciousness with swift memory loss." (p. 137)

Some other examples from Dennett of conscious and not conscious:

Are you constantly conscious of the clock ticking? If it suddenly stops, you notice this, and you can say right away what it is that has stopped; the ticks "you weren't conscious of" up to the moment they stopped and "would never have been conscious of" if they hadn't stopped are now clearly in your consciousness. An even more striking case is the phenomenon of being able to count, retrospectively in experience memory, the chimes of the clock which you only noticed was striking after four or five chimes. But how could you so clearly remember hearing something you hadn't been conscious of in the first place? The question betrays a commitment to the Cartesian model; there are no fixed facts about the stream of consciousness independent of particular probes.