

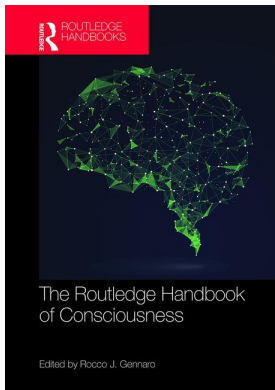
This article was downloaded by: 10.3.97.143

On: 02 Dec 2023

Access details: *subscription number*

Publisher: *Routledge*

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: 5 Howick Place, London SW1P 1WG, UK



The Routledge Handbook Of Consciousness

Rocco J. Gennaro

The Multiple Drafts Model

Publication details

<https://www.routledgehandbooks.com/doi/10.4324/9781315676982-12>

Francis Fallon, Andrew Brook

Published online on: 26 Mar 2018

How to cite :- Francis Fallon, Andrew Brook. 26 Mar 2018, *The Multiple Drafts Model from: The Routledge Handbook Of Consciousness* Routledge

Accessed on: 02 Dec 2023

<https://www.routledgehandbooks.com/doi/10.4324/9781315676982-12>

PLEASE SCROLL DOWN FOR DOCUMENT

Full terms and conditions of use: <https://www.routledgehandbooks.com/legal-notices/terms>

This Document PDF may be used for research, teaching and private study purposes. Any substantial or systematic reproductions, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The publisher shall not be liable for an loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

11

THE MULTIPLE DRAFTS MODEL

Francis Fallon and Andrew Brook

The phrase “Multiple Drafts Model” (MDM) refers to Daniel Dennett’s alternative to a Cartesian model of mind and, in many cases, serves as a synecdoche for Dennett’s general theory of consciousness. According to Dennett, clinging to the Cartesian conception of mind involves unwarranted assumptions and precludes a properly naturalistic understanding of mind, including experience. Providing an alternative, on the other hand, opens up the possibility of genuinely explaining consciousness.

Dennett makes a very strong, but also counterintuitive, case, which accounts for its divided reception. Largely thanks to Dennett, Cartesianism, in the sense relevant here, has become a charge most philosophers and scientists would hasten to avoid. This acceptance of Dennett’s negative case has not translated into a general acceptance of his positive theory (although the latter has been influential upon many). A clear explanation of the basic principles should show that giving up certain tempting and familiar beliefs is of a piece with demystifying consciousness.

1 The Cartesian Model

Descartes famously – or perhaps infamously – defended substance dualism, the claim that the mental and the physical belong to different realms of existence. (The mental is directly known and not extended in space; the physical is not directly known and has extension.) With some exceptions, more modern theories of mind reject this dualism, claiming that both the mental and the non-mental belong to the same realm of existence (monism). Of these, many propose some version of naturalism. According to naturalism, the mental exists but in principle finds explanation by reference to the natural world.

Throughout his long career, Dennett has advocated explaining mental phenomena as continuous with the workings of the natural, physical world. In other words, he endorses naturalistic principles. Still, even within a broadly naturalistic paradigm, some discussions of consciousness share certain features with Descartes’ account. Descartes believed that mechanical events in the brain were unconscious until they passed through the pineal gland, the “turnstile” of consciousness. The idea that consciousness will involve disparate, non-conscious elements all “coming together” in one place (not necessarily the pineal gland) and at one time has appeal. One might hold such a position without committing to any dualism. So, Dennett terms this “Cartesian materialism.”

Why should non-conscious mental events have to unite in one time and place in the brain in order to rise to consciousness? A popular metaphor for experience depicts it as a play or movie unfolding in the brain. This implies an internal viewer, who is watching the show. Dennett describes this as the Cartesian Theater, complete with an audience. Such an audience would be a homunculus, i.e., an agent within the experiencing person. This metaphor does not offer genuine explanation. If you experience because you take various visual and auditory percepts into your brain, and they remain unconscious until they unite in an inner theater and are received by a homunculus audience, the question remains: What would allow this homunculus to have one unitary experience of the various elements that have just debuted “on stage”? The only recourse would involve a further regress, where these pass once again into the homunculus’s own “brain,” where there exists a further theater and audience, *ad infinitum*.

Now, a dualist might want to insist on a special meeting place for the mental and physical substances,¹ and so would have a motivation for positing a Cartesian Theater, but is there a motivation for the monistic naturalist to posit any such place? Intuitively, it feels like consciousness is unitary. All in one moment we see a cloud pass across the sky, and hear a flock of birds take flight from nearby trees. We focus on one bird for an instant, simultaneously taking in its outline, the backdrop of the sky, and the sound of its cawing. Moreover, it seems that experience proceeds in one single stream – as a storm approaches, one experiences a lightning bolt across the sky, followed by a crash of thunder overwhelming a car alarm, after which comes a cascade of rain soon joined by a gust of wind, and then in the very next moment a combination of some or all of these. At any point, it seems, we are experiencing certain elements at once, and these points together make up our stream of consciousness.

We take this single, unified stream of experiences to be very rich: to include, e.g., detailed vision out to the edges of our visual field, the sound of many individual raindrops pelting the ground, etc. Even if we forget almost immediately where the clouds were, if the lightning bolt began at the western or eastern side of the sky, or if there were more than ten audible raindrops per second, there is a fact of the matter about just what we were experiencing at any given point.

Put briefly, then, intuition motivates even some naturalists to commit to a Cartesian Theater: consciousness seems like a unified stream of experiences proceeding past the “audience” within us. We have seen, though, that this move is non-explanatory, viciously regressive even. So, we have to give up our intuition (or give up on explaining consciousness).

The brain is headquarters, the place where the ultimate observer is,² but there is no reason to believe that the brain itself has any deeper headquarters, any inner sanctum, arrival at which is the necessary or sufficient condition for experience. In short, there is no observer inside the brain.

(Dennett 1991: 106)

Given our habitual comfort with Cartesian Theater metaphor, this may appear to remove a useful option, but now we know better, that the promise of explanation via inner movies and audiences is a false one. Instead, this latest step has removed a constraint upon explanation. What does explain consciousness will not have to conform to our (prior) intuitions (although it would help if it could explain their existence).

2 Multiple Drafts

Having let go of the requirement that mental events must pass through a central processing area in order to achieve consciousness implies that consciousness does not have to consist of a

seamless stream of unified experiences, even if it seems that way. Immediately, this appears to present a paradox. How could consciousness really be one way and yet seem another way? Isn't consciousness precisely in the seeming? Doesn't the subject have direct access to it, and so infallibility concerning it?

These questions cannot all receive satisfactory answers right away. Dennett knows this, noting in his central expression of MDM that making it a "vivid" and "believable alternative" to the Cartesian Theater "will be the hardest part of the book" (1991: 114). A temporary (and rather unsatisfactory) general answer might note that these questions all reflect deep intuitions, and recall, from the end of Section 1 just above, that explaining intuitions, even while not necessarily granting them authority, should in principle suffice. Only once we pull apart the mechanisms of various "seemings" can we assess their claim upon the reality of our mental lives.

To say that there does not have to be one single stream of consciousness is to say, in other words, that there does not have to be one single, authoritative narrative that makes up consciousness. The brain, in cooperation with senses, registers multiple stimuli, but does not need to re-process those registrations into a final copy for "publication." In Dennett's words:

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. In other words, discrimination does not lead to a re-*presentation* of the already discriminated feature for the benefit of the audience in the Cartesian Theater – for there is no Cartesian Theater.

(1991: 113)

This describes a disjointed process, in tension with our belief in a stream of consciousness. Indeed, "this stream 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" (Dennett and Akins 2008).

An example will help illustrate how MDM and the Cartesian Model differ in their implications for assessing experience (1991: 137–8; 2008). Intently reading in a study (perhaps having found shelter from the storm), you observe the person sitting across from you look up, and just then you become aware – seemingly for the first time – that the grandfather clock has been chiming. Before the other person looked up at it, this had not come to your attention. You then find yourself able to count, retrospectively, the (three) chimes before you had become aware (at the fourth chime).

What has happened? Were you conscious of the chimes all along, and then became "extra" aware of them? Were you unconsciously registering the chimes, and then called them forth once prompted by an environmental stimulus? Nothing at the level of introspection answers these questions definitively. Mechanisms in the brain will have registered the chimes, possibly in different ways, but why should an examination of these speak with authority to exactly when one became conscious, since introspection will be incapable of confirming one way or the other?

Only on a Cartesian model do these questions require answers, and so only on a Cartesian model does the apparent inability to settle them pose a problem. On MDM, because one single official draft does not proceed through time along a continuous line, there does not have to be a fact of the matter about these issues of timing. Were you to insist that there must be a fact of the matter, this would introduce the strange category of objective facts about your awareness, of which facts you yourself are unaware. Instead – and this is a crucial point – the privileged

status of consciousness is conferred *retroactively* upon (even very recent) memories when stimuli prompt us to attend to them.

This, then, is how MDM respects the powerful appearance of a single, “official” draft. Our conviction about the existence of a master narrative does not reflect its existence in the first instance, but is in the event a creation of (sometimes non-veridical) retrospective assembly of various perceptual fragments and associations. What at first may strike us as paradoxical becomes merely (but deeply) counterintuitive: there is no unitary stream of consciousness, but there are drafts whose availability for recall supply the material for the ad hoc manufacture (upon prompting) of linear narratives, and this regular capacity for spinning such yarns makes it *seem* – even in their absence – as though one linear stream of consciousness exists.

Maybe this is too quick: maybe there *was* a fact of the matter of your consciousness. Maybe you were aware of the chimes in real time, but forgot them almost as quickly. Alternatively, the initial chimes were registered unconsciously and then introduced into consciousness later, on a time-delay. Both these interpretations preserve the intuition that one single continuous draft of consciousness exists. The next section addresses this issue.

3 Orwellian and Stalinesque Streams of Consciousness

Dennett speaks to both of these possible interpretations directly. He maintains that while a mental event may bear description as conscious or non-conscious, “it is a confusion... to ask *when it becomes conscious*” (1991: 113). The argument claims not just the non-necessity of there being a fact of the matter concerning such precise timing, but the incoherence of requiring such facts.

It will help to address the terminological distinctions. Imagine the following case of false memory. You remember seeing a woman in a hat at yesterday’s party (even though there was no woman in a hat). If you had no initial experience of the woman in the hat, and then after the party you misremembered (or someone surgically implanted a false memory, for that matter), then the chronology is similar to the first interpretation of the chimes case. In both cases, something happens after the fact of our conscious experience to alter our memory of it. Dennett calls such instances Orwellian, because in Orwell’s *1984*, the Ministry of Truth rewrites history (1991: 117–18).

The other possibility works pre-emptively. For example, you saw other people in hats at this party, non-consciously perceived a woman without a hat, and quickly afterward, in your single authoritative draft of consciousness, fused hat-wearing with the experience of seeing the woman. Here the chronology is similar to the second interpretation of the chimes, because unconscious registrations are introduced into consciousness with some slight delay. Dennett terms such cases Stalinesque, after the staged trials that took place under Stalin (1991: 119).

Empirical evidence may support one or the other interpretation at a comparatively macro timescale. For instance, if you mentioned the hatlessness of the woman at the party yesterday, but today (mis)remember her as having worn a hat, this would suggest the Orwellian interpretation, where your consciousness was over-written. It seems natural to think that this should apply even at a micro timescale, which of course is what any theory of a single stream of consciousness expects.

Dennett uses a thought experiment to show that at the micro timescale, things change (1991: 117–19). This time, imagine that a woman with long hair runs past you. One second later, the memory of a woman with short hair and glasses contaminates the memory of the long-haired woman, and you believe that you had a visual experience of a long-haired woman with glasses. The Orwellian interpretation suggests itself: you experienced the woman without glasses run past, but then your brain “wrote over” that experience almost immediately. The Stalinesque will

work too: “your subterranean earlier memories of that woman with the eyeglasses could just as easily have contaminated your experience on *the upward path*,” so that the one authoritative stream of consciousness included only the experience of a woman with glasses running by. No way of determining the truth of the single stream of consciousness makes itself available here. Introspection is blind to the causal mechanisms at work, and unlike in the earlier example, where someone might remind you of having mentioned a hatless woman yesterday (thereby giving the Orwellian interpretation support), there is no further way to settle the matter, “leaving no fact of the matter about whether one is remembering mis-experiences or mis-remembering experiences” (1998: 135). There is nothing unsettling about this on MDM, because unlike Cartesian models it denies the existence of one “official” draft of consciousness.

Empirical experimentation bears out the point. Dennett discusses Kolers’ “color phi phenomenon.” In this experiment, subjects are shown a red dot (A) at one place on a screen, rapidly followed by a blank screen, and then a green dot (B) on another part of the screen. The experiences involve movement and change of a single spot: “Subjects report seeing the color of the moving spot switch in midtrajectory from red to green” (1991: 120). The Orwellian gloss on the Kolers experiment posits an accurate conscious experience, immediately obliterated and replaced by the midtrajectory shift report: AB, quickly forgotten, replaced with ACDB (where C and D are intermediary imagined spots), which gets reported. The Stalinesque interpretation posits something like a “slack loop of film,” allowing for editing and censoring, before consciousness takes place. This has the subject inserting CD preconsciously, so that the whole sequence of color conscious events is ACDB.

So here’s the rub: we have two different models of what happens in the color phi phenomenon.... [B]oth of them are consistent with whatever the subject says or thinks or remembers. Note that the inability to distinguish these two...does not just apply to the outside observers.

(1991: 122–3)

Whether cases like this phenomenon have Orwellian or Stalinesque origins would have to have an “answer if Cartesian materialism were true...even if we – and you – could not determine it retrospectively by any test” (1991: 119). On a model of consciousness where there is a strict, non-smearred sequence of events streaming past a conscious homunculus, or entering and exiting a stage in a Cartesian Theater, there would be a fact of the matter about the origins, on any time scale. We may, perhaps through neuroscientific progress, find answers. “But this is just where the reasons run out... [T]here is no behavioural reaction to a content that *couldn’t* be a merely unconscious reaction” (124). Focusing on one or another mental event of brain processing as the moment of consciousness “has to be arbitrary,” because:

[T]here are no functional differences that could motivate declaring all prior stages and revisions to be unconscious or preconsciously adjustments, and all subsequent emendations to the content (as revealed by recollection) to be post-experiential memory contaminations. The distinction lapses in close quarters.

(126)

The problem for the Cartesian model therefore runs deeper than an epistemological shortcoming awaiting empirical resolution: nothing can settle the question of the “true” stream of consciousness, because there isn’t one. A distinction in which the truth or falsity of the two sides of the distinction makes no difference is not a basis for an explanation of any kind.³

4 “Fame in the Brain” and Probes

In introducing MDM, Dennett describes it as “a first version of the replacement” for the image of mind suggested by the Cartesian model (1991: 111). Since then, he has not abandoned the principles of MDM, but he has augmented it with an alternative metaphor. The original MDM “did not provide... a sufficiently vivid and imagination-friendly antidote to the Cartesian imagery we have all grown up with, so... I have proposed what I consider to be a more useful guiding metaphor: ‘fame in the brain’ or ‘cerebral celebrity’” (2005: 136; see 1998, pp. 137–9, for an early treatment of this metaphor).

The Cartesian model encourages us to think of consciousness as a play (or movie) in the mind, viewed by an audience in a Cartesian Theater within the brain. The tempting notion of a single stream of consciousness fits this well: one single series of conscious states, much like the frames that make up a television show. MDM denies that otherwise unconscious contents travel to a central processing place, where each finds its place in a queue to form the stream of consciousness. Instead, unconscious contents compete with each other for “fame.” Not all people can be famous, so the process of becoming famous is competitive. Both fame and consciousness are “not precisely dateable” (Dennett and Akins 2008; for the classic Dennettian analysis of the implications of states of consciousness taking time to come into existence, see Dennett and Kinsbourne 1992). Section 3 above showed why this holds for consciousness, and gaining fame similarly defies exact chronology, even if it can be assessed at a comparatively macro timescale. Moreover, each “is only retrospectively determinable since it is constituted by its sequelae” (Dennett and Akins 2008).

Even if this metaphor does not encourage us to think of consciousness as a medium of representation, like television or theater, might it accidentally rely on a homunculus to decide “fame”? Understanding that the nature of the fame in question commits Dennett to no such fallacy requires returning to a “crucial point” noted in Section 2 above. The privileged status of consciousness is conferred *retroactively* upon (even very recent) memories when stimuli prompt us to attend to them.

Following Dennett, we have been citing instances where attention plays a role in the generation of consciousness. While this indicates an overlap with attentional theories of consciousness, Dennett does not seem to require attention *per se*. The crucial requirement for conferring consciousness is the involvement of one or more “probes”. A probe can be “whatever event in the brain happens to boost some aspect of the current content-fixations into prominence. *In the simplest case*, a probe is a new stimulus that draws attention...” (Dennett and Akins 2008, emphasis added). Because Dennett’s examples of probes involve attention, we will continue to feature it centrally.⁴ To return to the chimes case, when someone else looked up at the clock, this prompted you to consider the number of chimes – a case of probing mental contents. This drew attention to just-registered sounds. Without this attention, they would not have gained any “fame”; they would have registered as temporary micro-drafts, but without any probing would have remained unnoticed, never rising to prominence. In this context, it makes sense to quote more completely a passage cited in Section 2 above:

[A]t any point in time there are multiple drafts of narrative fragments at various stages of editing in various places in the brain.... Probing...produces different effects, producing different narratives – and these are narratives: single versions of a portion of ‘the stream of consciousness’.

(Dennett and Akins 2008)

Because of the probe, these partial drafts become available for further judgments, which may include the retroactive framing of these elements as part of a seamless stream of unitary experiences (which Dennett sometimes calls “retrospective coronation”).

Consciousness comes about when mental contents get noticed. Such notice, or fame, depends upon the actualization of available judgments. No re-presentation to an experiencing homunculus enters into the explanation, nor does it incorporate any reliance upon properties qualitatively distinct from discriminative judgments. “Consciousness, like fame, is not an *intrinsic* property, and not even just a *dispositional* property; it is a phenomenon that requires some actualization of the potential” (2005: 141). Only its prominence in cognition – and not a further special quality – makes a mental content conscious. “[T]his is not the prominence, the influence or clout, those contents would have had anyway in the absence of the probe” (Dennett and Akins 2008).

Section 3 explained that requiring an exact moment for consciousness misses an essential truth about experience, that no one definitive chronology of consciousness exists, because it is temporally “smeared” among multiple drafts. The preceding discussion of probes shows that certain (portions of) drafts win competitions for fame, get noticed, and earn judgment as fitting into one single stream. Consider the familiar question of whether you were conscious during your commute home. At first it might seem as though you were not, but upon trying, you find that you recall a number of details. Must you have been conscious of them all along? You certainly registered these in a way that disposed you, upon probing, to recall them. It also stands to reason that more temporally local probes would have resulted in at least as detailed recall. The question is ill-posed. Succinctly put, “A temporally punctate event need not make the transition from unconsciously discriminated to consciously experienced in a temporally punctate moment.” In other words,

We can expect to find, and time the onset of, necessary conditions for fame in the brain... but when sufficient conditions ripen slowly and uncertainly over longer periods of time, identifying these onsets of necessary conditions as the onset of consciousness is at best arbitrary and misleading.

(Dennett and Akins 2008)

The dispositions are necessary for entering what one takes to be the stream of consciousness, but are insufficient to count as consciousness without a subsequent boost in content-fixation (as in attention), exemplified by an ability to report these things (veridically or not) to yourself, upon probing, which probing may happen almost in real time, or at quite a delay.

5 The Ontology of Consciousness

Descartes’ dualism gives us the most obvious case of claiming different realms of existence for the mental and the physical. As noted, most philosophers and scientists reject dualism in favour of naturalism, but the question of how to explain the mental by reference to nature persists. In particular, the endurance of the consciousness debate stems from its seeming to *be* a different kind of thing from material or arrangements and functions of matter. Even among those who claim common allegiance to naturalism, then, the *ontology* of consciousness remains controversial.

Dennett, sensitive to this, introduces his MDM only after articulating a methodological approach he calls heterophenomenology (1991: 71–78). This approach maintains strict neutrality with respect to the ontological status of experiential (phenomenological) components. Recall the questions posed near the beginning of Section 2 above: How could consciousness really be one way and yet seem another way? Isn’t consciousness precisely in the seeming? Doesn’t the

subject have direct access to it, and so infallibility concerning it? Heterophenomenology begins by making no assumptions about the answers to these questions. It refuses to take for granted that the intuitive responses to these are correct, that intuitions are generally infallible or fallible, or even that these questions are posed unambiguously. The proper methodology is the most cautious: examining the empirical evidence and determining what conclusions it allows.

Returning to the Kolers phenomenon illustrates how one may begin neutrally and proceed to a defense of a particular ontology. As a matter of empirical fact, no changing spots (CD) exist in the middle of the screen. Strictly, then, the subject does not see such a spot, although she may sincerely insist upon having seen such spots. A tension exists, then, between the subject's reports and the empirical evidence. One way to attempt to resolve this, without discounting the subject's authority concerning her experience, maintains that her experience does in fact include referents for the spots CD: phenomenal units, dubbed "qualia." The term "qualia" is, by itself, ontologically neutral. Sometimes it simply serves as a placeholder, covering the various elements in experience, however they might receive characterization or explanation. More typically, however, "qualia" refers to inner, intrinsic, irreducible "bits" of consciousness. This characterization holds important implications: If the components of consciousness are inner, intrinsic, and irreducible, then they are impervious to explanation by reference to an objective, or "third-person," ontology. This rules out any standard scientific explanation of first-person, subjective experience.

Heterophenomenology might admit the logical possibility of such a position, but denies that there is reason to grant it truth. If the proponent of this robust understanding of qualia – Dennett terms such thinkers "qualophiles" – defends her claim on the grounds of its intuitive nature, this simply begs the question concerning the authority of our intuitions. It follows, then, that if an empirical, third-person explanation is available, and moreover can satisfactorily address our intuitions, we should prefer it.

On MDM, the individuated, qualic event "spot changing color in the middle of the screen" is not irreducible. That is – in principle at least – reference to mechanisms can account for the subject's conviction that she saw such a change in spots mid-screen. Mechanisms of perception, association, and memory all work in parallel in the subject's brain. The stimuli include only two spots (A and B), and we cannot assume inner, irreducible CD spots. The experiment itself requires the subject to attend, and therefore serves as a probe. Given these stimuli and the probe, the subject engages in a rapid retroactive synthesis of multiple parallel, non-conscious drafts. This gives rise to a non-veridical, although sincere, judgment that in the middle of the screen a spot changed from red to green (see Dennett [1988] for the classic treatment of the claim that we do not need a notion of ineffable, irreducible qualia; see also [1991: 369–411]).

This respects the subject's conviction about the changing spots. It really seems to her that they existed, in the place and order she reports. That this seeming consists in non-veridical judgment is no denial of that. She has infallibility about how it seems – which is to say that she has authority about what her judgments are – but her judgments themselves are fallible, and in this case are false. At the same time, the MDM explanation has not posited any special objects in its ontology that stand beyond the reach of a standard naturalistic vision. "Conscious experiences are real events occurring in the real time and space of the brain, and hence they are clockable and locatable *within the appropriate limits of precision for real phenomena of their type*" (1998: 135, emphases added). The appropriate limits preclude very fine-grained and irreducible, serial qualic events, such as spots C and D: "I am denying that there are [qualia]. But... I wholeheartedly agree that there seem to be qualia" (1991: 372).

Dennett routinely describes naturalism about the mental as requiring that each mental phenomenon receive explanation by reference to simpler mechanisms, ultimately bottoming out at the mechanical level of description. Excising irreducible qualia from MDM's ontology is

necessary to such an approach. “There is no reality of conscious experience independent of the effects of various vehicles of content on subsequent action (and hence, of course, on memory)” (1991: 132).⁵ Those cognitive events that influence action at least have a chance of disposing us to judge them as parts of our stream of consciousness; those that “die on the vine” (i.e. do not influence action) cannot. Probes generate prominence, determining which of the multiple drafts receive retrospective coronation as conscious.

No one particular homunculus decides what content is prominent (just as one person’s regard does not confer fame upon another). This role is discharged throughout the brain. Many subpersonal mechanisms underpin the judgments at the personal level that constitute our conviction of having a unified stream of consciousness with particular and seemingly irreducible or contents. Prominent mental content may exert influence upon a variety of actions; among these, the clearest demonstrations come in the form of verbal reports. This is not due to such reports’ infallibility (as we have seen, they are fallible in one, direct, sense), nor to verbalization’s residence in one privileged conscious arena (MDM has denied anything holding the place of a Cartesian Theater from the very beginning). Rather,

The personal level of explanation is defined by the limits of our abilities to respond to queries about what we are doing and why... A reported episode or nuance, current or recollected, has left the privacy of the subpersonal brain...

(Dennett and Akins 2008)

Just as the life of an organism is explained ultimately by reference to non-living parts, the person is explained by interplay at the subpersonal level; consciousness is explained by the functional roles of non-conscious mental content. Details of how such function itself finds explanation at the mechanical level is a matter of ongoing empirical research.

6 Situating MDM

The introductory passage of this chapter noted that “Multiple Drafts Model” can refer to Dennett’s overall theory of consciousness, and what followed linked MDM to the more recent “fame in the brain” metaphor, as well as to the methodological approach of heterophenomenology. Dennett has hewed closely to the core principles of MDM for decades, augmenting it without altering the fundamental arguments, and applying it with varied emphases to suit different contexts. Throughout, his arguments concerning consciousness have enjoyed a high profile: to give an indicative overview of the field, even a brief volume on consciousness would need to include a discussion of MDM. Because MDM challenges familiar assumptions about consciousness, and also because it fits a certain scientific worldview, it has generated an enormous body of literature – hundreds of papers’ and several books’ worth – both sympathetic and critical. Situating it in the broader discussion in a limited space will have to sacrifice precision for balance of coverage.

By now, it goes without saying that MDM stands at odds with dualism. Perhaps it should go without saying that it stands opposed to eliminative materialism, the position that consciousness strictly merits no ontological status. Dennett has eschewed this association all along, but it is still a matter of some controversy (Fallon, forthcoming). Very recently, for example, Dennett felt the need to offer clarification anew: “Consciousness exists, but just isn’t what some folks think it is” (2017: 223).

Those who read Dennett’s restrictions on the ontology of consciousness as too reductive accuse him of “explaining consciousness away.” Section 5 gave the reasons why denying

ontological status to irreducible qualia may not amount to denying that consciousness exists: that things (really) *seem* how – but don't necessarily *exist* in just the way that – they seem. This defense has failed to satisfy many. Among such critics are some of the most influential figures in philosophy of mind, and among their arguments are some of the most famous thought experiments in contemporary philosophy of any kind, themselves very durable and appearing in countless discussions in the literature. (Often the original versions predate MDM.) Uniting all of these is the conviction that Dennett's MDM "leaves something out."

Ned Block has consistently criticized Dennett's theory for being overly cognitive, failing to account for essentially non-cognitive experiences or elements of experience. He maintains a separation between phenomenal consciousness, a domain arguably coextensive with qualia, and access consciousness. Dennett's functional theory has the resources to treat the latter, but the former altogether eludes the explanatory net of MDM. Block (1990) presses his point through the inverted qualia thought experiment, which has several iterations. The basis of each is the intuition that you could see green wherever I see red, and the two of us could function in identical ways. Therefore, function does not exhaust phenomenal experience.

David Chalmers (1996) argues that nothing currently known to science about matter or its arrangement in the brain logically implies experience. We cannot tell why physical systems such as ours could not operate as they do, while remaining "in the dark," i.e. without generating the experience we enjoy. He makes use of a zombie thought experiment: we can imagine a complete physical and functional replica of a human being that has no interior life at all, so current physics and neuroscience cannot account for experience. He entertains the possibility of an augmented, future science that identifies fundamental experiential (or proto-experiential) properties in the physical world.

Another well-known thought experiment that casts doubt on physicalism, and so applies to Dennett, comes from Frank Jackson (1982): Mary is a scientist who has a complete knowledge of the objective facts about color – surface reflectance, visual cortices, conventions of naming, etc. She is confined to a black and white laboratory her entire life until, one day, she is released into the outside world and experiences color for the first time. She has learned something new, which was unavailable to her earlier, despite her expertise about the third-person facts. So, the physical facts do not suffice to explain subjective experience. Similarly, Thomas Nagel (1974) argues that knowing third-person facts about a bat would not suffice for us to understand "what it is like" (subjectively, experientially) to be a bat.

John Searle denies that Dennett has captured the "special causal powers of the brain" that produce consciousness, but is optimistic about future science doing so. His Chinese Room thought experiment (1980), the most written-about thought experiment in the history of philosophy, challenges not just Dennett, but every non-biological form of materialism on a fundamental level, because it concerns the origin of intentionality (or aboutness, upon which accounts of meaning rely). Basically, Searle imagines someone who, like him, knows no Chinese, working in a large room rigged with complex symbolic input-output instructions. When Chinese characters are fed into the room (input), the person uses the instructions (program) to select the appropriate Chinese characters to send out of the room (output). The worker could be an Anglophone monoglot, and the instructions could be all in English. From the outside, though – if everything is set up appropriately – it would seem as though the person inside understood Chinese. Programmatic input-output relations appropriate to the external world therefore do not suffice to ground true meaning. Any mental model that confines itself to describing such functional dynamics leaves something out.

Even this cursory and partial exposition of some of the livelier objections to MDM shows their intuitive appeal. Dennett responds to each of them in numerous contexts (1991, 2005, inter

alia).⁶ The responses are complex and, again, counterintuitive. One argument thematic among Dennett's responses holds that these thought experiments are merely "intuition pumps," designed to exploit existing intuitions rather than providing good grounds for them. Nevertheless, their intuitive appeal gives the anti-MDM camp a distinct rhetorical edge. The reader should bear in mind that while the anti-MDM arguments typically claim allegiance to naturalism, their references to future science and special, as-yet unknown, causal powers of the brain reveal their uneasy fit with a standard scientific worldview. In this, sometimes less visible, sense, Dennett's MDM has its own intuitive appeal. It lies beyond the scope of this chapter to resolve this clash of intuitions; the objective here has been to clarify Dennett's case, the better for the reader to assess it.

This treatment of objections to MDM should not overshadow its alliances. MDM's rejection of a central Cartesian Theater fits well with Bernard Baars' "global workspace" model (1988). Higher-order theories of consciousness such as David Rosenthal's Higher-Order Thought (HOT) theory (2005 and many earlier works) explain consciousness as arising when mental contents themselves become objects of (ipso facto) higher-order mental states. Unsurprisingly, this thoroughly cognitive model receives a sympathetic hearing in Dennett's work. Along similar lines, and from an evolutionary perspective, Antonio Damasio (1999) explains consciousness as the organism registering, as a stimulus, itself in the act of perceptual change (see Dennett 1993: 920 for more commonalities between Damasio and Dennett). Jesse Prinz (2004) conceives of consciousness similarly to Damasio, and, like Dennett, assigns attention a crucial role. If anything, Dennett is more optimistic about the explanatory reach of the latter two projects than their authors.

More recently, Dennett has enthusiastically endorsed Andy Clark's explanation (2013) of how the brain seems to project phenomenal properties out into the world. The "projection" metaphor glosses a functional process that can be elaborated scientifically. The organism is "designed to deal with a set of [Gibsonian] affordances, the 'things' that matter," and this "*Umwelt* is populated by two R&D processes: evolution by natural selection and individual learning" (Dennett 2017: 165–6). Feedback (probabilistic, Bayesian, feedback) from how top-down guesses do against bottom-up incoming data determines what in the environment becomes salient. A lack of feedback would mark the absence of prediction error (data confirm the top-down guesses) and would work as confirmation (167–9; see also Clark 2013). The affordances we experience result from such processes. One might redescribe this, in the language of MDM, as conceiving of the brain as oriented by evolutionary and developmental pressures to probe its own activity, where unlikelier events win competition for attention from such probes.

The end of Section 5 noted that ascertaining the mechanical realization of the relevant functional processes is a matter of ongoing empirical research. Any model that acknowledges that mental content does not require re-presentation of mental content to an audience in a Cartesian Theater to become conscious – that is, eschews conceiving of consciousness as a single temporal stream of consecutive conscious events, and accounts for the construction of an apparent stream of consciousness retroactively – will hold consistent with MDM's central principles. "A wide variety of quite different specific models of brain activity could qualify as multiple drafts models of consciousness if they honored its key propositions" (Dennett and Akins 2008; see Dennett 2005: 133–42 for a selection of neuroscientific models consistent with MDM). By explaining our intuitions about experience, without granting them ultimate authority, the MDM secures the viability of contemporary scientific research into consciousness.

Notes

- 1 That a non-extended substance should have the property of being locatable in extended space is, of course, paradoxical. This observation lies at the heart of the general rejection of Cartesian dualism.

- 2 Even the brain's status as location of the observer is contingent. As Dennett (1981) notes, if one's brain and body were separated, with lines of communication between the two maintained through radio connection, and the brain kept alive in a vat while the body went on a remote mission, one's point of view would be the sensory contacts of the body with its surrounding stimuli, and not the vat.
- 3 This move has smacked of verificationism to many commentators. Dennett's response has mostly been to accept the charge but deny its force. While we do not have room for a full discussion of this response, it is not clear that there is much wrong with his variety of verificationism (what he once referred to as "urbane verificationism") (Dennett 1991: 461–2; see also Dennett 1993: 921–2, 930n; Dahlbom 1993; and Ross and Brook 2002, Introduction).
- 4 Of course, if mechanisms other than attention – perhaps less deliberate or guided, less compelled by stimulation, or more sub-personal than attention – can serve as probes, this should be explicated. Depending upon one's sympathies, this point can be regarded as a complaint about Dennett's account or as a research question motivated by it. The same can be said for any lack of specificity concerning the kinds of memory relevant to consciousness.
- 5 This is a particularly loaded sentence. For an explication of the "vehicles" at issue, see Brook (2000). For a discussion of the ontology of consciousness as including phenomenological *effects*, easily mistaken for inner *causes* of phenomenal experience, see Chapter 14 of Dennett (2017); see also Dennett (2007). Fallon (forthcoming) argues that these claims support a "realist" interpretation of Dennett on consciousness.
- 6 Dennett's numerous comments (1991) on Fodor's "language of thought" (LOT) account (1975) nicely encapsulate his positive arguments concerning intentionality. See also the exchange between Rey (1994) and Dennett (1994). Dennett (1993: 925–8) gives one succinct response to the "inverted qualia" arguments.

References

- Baars, B. (1988) *A Cognitive Theory of Consciousness*, Cambridge: Cambridge University Press.
- Block, N. (1990) "Inverted Earth," *Philosophical Perspectives* 4: 53–79.
- Brook, A. (2000) "Judgments and Drafts Eight Years Later." In D. Ross, A. Brook, and D. Thompson (eds.) *Dennett's Philosophy: A Comprehensive Assessment*, Cambridge, MA: MIT Press.
- Chalmers, D. (1996) *The Conscious Mind: In Search of a Fundamental Theory*, New York: Oxford UP.
- Clark, A. (2013) "Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognitive Science," *Behavioral and Brain Sciences* 36: 181–204.
- Dahlbom, B. (1993) "Editor's Introduction." In B. Dahlbom (ed.) *Dennett and His Critics*, Oxford and Cambridge, MA: Blackwell Publishers.
- Damasio, A. (1999) *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*, New York: Harcourt, A Harvest Book.
- Dennett, D. (1981) "Where Am I?" In D. Hofstadter and D. Dennett (eds.) *The Mind's I: Fantasies and Reflections on Self and Soul*. New York: Basic Books.
- Dennett, D. (1988) "Quining Qualia." In A. Marcel and E. Bisiach (eds.) *Consciousness in Contemporary Science*, Oxford and New York: Oxford University Press.
- Dennett, D. (1991) "The Message Is: There Is No Medium," *Philosophy and Phenomenological Research* 53: 919–931.
- Dennett, D. (1993) *Consciousness Explained*, Boston: Little, Brown.
- Dennett, D. (1994) "Get Real," *Philosophical Topics* 22: 505–568.
- Dennett, D. (1998) *Brainchildren*, Cambridge, MA: MIT Press.
- Dennett, D. (2005) *Sweet Dreams*, Cambridge, MA: MIT Press.
- Dennett, D. (2007) "Heterophenomenology Reconsidered," *Phenomenology and the Cognitive Sciences* 6: 247–270.
- Dennett, D. (2017) *From Bacteria to Bach and Back: The Evolution of Minds*, New York: W.W. Norton & Company.
- Dennett, D., and Akins, K. (2008) "Multiple Drafts Model," *Scholarpedia* 3: 4321. http://www.scholarpedia.org/article/Multiple_drafts_model. Accessed 17 April 2017.
- Dennett, D., and Kinsbourne, M. (1992) "Time and the Observer: The Where and When of Consciousness in the Brain," *Behavioral and Brain Sciences* 15: 183–201.
- Fallon, F. (forthcoming) "Dennett on Consciousness: Realism without the Hysteries," *Tōpoi*.

- Fodor, J. (1975) *The Language of Thought*, Scranton, PA: Crowell.
- Jackson, F. (1982) "Epiphenomenal Qualia," *Philosophical Quarterly* 32: 127–136.
- Nagel, T. (1974) "What Is It Like to Be a Bat?" *The Philosophical Review*, LXXXIII, 435–450.
- Prinz, J. (2004) *Gut Reactions: A Perceptual Theory of Emotion*, New York: Oxford University Press.
- Rey, G. (1994) "Dennett's Unrealistic Psychology," *Philosophical Topics* 22: 259–289.
- Rosenthal, D. (2005) *Consciousness and Mind*, Oxford and New York: Oxford University Press.
- Ross, D., and Brook, A. (2002) "Introduction." In Don Ross and Andrew Brook (eds.) *Daniel Dennett*. Cambridge and New York: Cambridge University Press.
- Searle, J. (1980) "Minds, Brains, and Programs," *Behavioral and Brain Sciences* 3: 417–458.

Related Topics

- Dualism
- Materialism
- The Global Workspace Theory
- Representational Theories of Consciousness
- Consciousness and Attention