

Journal of Cognition and Neuroethics

Evolution Beyond Determinism: On Dennett's Compatibilism and the Too Timeless Free Will Debate

Maria Brincker

University of Massachusetts Boston

Biography

Maria Brincker is currently an Assistant Professor of Philosophy at University of Massachusetts Boston and has previously held an Arts & Neuroscience fellowship at Columbia University. Her interdisciplinary work spans a broad array of topics within philosophy of mind and neuroscience, such as affordance perception, mirror neuron theories, sensorimotor grounding of cognition, typical and atypical social cognitive development, action theory and aesthetics.

Acknowledgements

I want to thank the organizers and participants at the 2014 "Free Will" conference at the Center for Cognition and Neuroethics at the University of Michigan-Flint for their kind feedback on an earlier version of this paper, and for the Kane-Dennett Prize recognition and funding. The ideas of this paper have been stewing for more than fifteen years but this conference supplied the needed impetus to action. Further, I want to thank Elly Vintiadis for her helpful comments and last but not least deception expert Mark Mitton for our many heated discussions and his deep insights into questions of causality and the selective nature of perception. As Bergson, he has tirelessly alerted me to the impossibility of conveying the true dynamics of action with frozen symbols without hiding precisely what needed to be revealed.

Publication Details

Journal of Cognition and Neuroethics (ISSN: 2166-5087). March, 2015. Volume 3, Issue 1.

Citation

Brincker, Maria. 2015. "Evolution Beyond Determinism: On Dennett's Compatibilism and the Too Timeless Free Will Debate." *Journal of Cognition and Neuroethics* 3 (1): 39–74.

Evolution Beyond Determinism: On Dennett's Compatibilism and the Too Timeless Free Will Debate

Maria Brincker

Abstract

Most of the free will debate operates under the assumption that classic determinism and indeterminism are the only metaphysical options available. Through an analysis of Dennett's view of free will as gradually evolving this article attempts to point to emergentist, interactionist and temporal metaphysical options, which have been left largely unexplored by contemporary theorists. Whereas, Dennett himself holds that "the kind of free will worth wanting" is compatible with classic determinism, I propose that his models of determinism fit poorly with his evolutionary theory and naturalist commitments. In particular, his so-called "intuition pumps" seem to rely on the assumption that reality will have a compositional bottom layer where appearance and reality coincide. I argue that instead of positing this and other "unexplained explainers" we should allow for the heretical possibility that there might not be any absolute bottom, smallest substances or universal laws, but relational interactions all the way down. Through the details of Dennett's own account of the importance of horizontal transmission in evolution and the causal efficacy of epistemically limited but complex layered "selves," it is argued that our autonomy is linked to the ability to affect reality by controlling appearances.

Keywords

Free will, determinism, compatibilism, naturalism, evolution, Dennett, causality, emergence, time, appearance-reality, process philosophy, action, choice, autonomy, self, metaphysics

Introduction: A tension in Dennett's compatibilism

Dan Dennett has over the past three decades developed and argued for an evolutionary and naturalist view of free will (Dennett 1984 & 2003). The core idea is that over evolutionary time, and through myriads of small and massively complex biological and cultural selections, the ability of voluntary control—what we call free will—has gradually evolved. Free will under Dennett's interpretation is simply the adaptive ability to anticipate outcomes and to flexibly exert control over factors in the world according to one's preferences, perceptions and deliberations. Beyond his discussion of how the ability of reason-based voluntary control has evolved, Dennett has spent considerable energy arguing that this kind of 'freedom' is 1) utterly compatible with determinism and 2) the kind worth wanting. He, in other words, defends a compatibilist position and works hard

to dispel the incompatibilists, be they hard determinists or libertarians, who all claim that “true” free will must involve some kind of indeterminism, and hence some kind of denial of determinism.

Dennett’s view of free will might indeed be compatible with determinism. However, I shall in this article argue that his historical and perspectival evolutionary story of action choice and normativity would invite us to think beyond the classic determinism and indeterminism dichotomy, toward an emergentist or interactivist account. Further, I suggest that such an alternative would be more empirically plausible and, by way of its temporality, makes room for a free will theory that even some indeterminists might find worth wanting.

Throughout his writings, Dennett employs a series of characteristic toy models or “intuitions pumps” as he calls them: “Intuition pumps are cunningly designed to focus the readers attention on “the important” features, and to deflect the reader from bogging down on hard to follow details” (Dennett 1984, 12). I propose that we might want to use a bit of Dennett’s own methodology on himself and “turn the knobs” on his intuition pumps to see what is doing the work. Following his advice, I will “go slow where others go fast” and argue that his atomistic and compositional metaphors might be doing a lot of the heavy lifting in his compatibilist picture. I point to tensions within Dennett’s own account between his keen eye to the causal role of layered heterogeneous inners, as well as the causal efficacy of pattern recognition in evolutionary and cultural history on the one hand, and his eternalist and compositional atomist approach to questions of metaphysics on the other. The question is whether without his metaphysical assumptions, his evolutionary story would invite a more temporal, emergentist—and arguably more naturalist—understanding of the causal fabric of the world. Thus, his own evolutionary theory might be used to reveal new options in the metaphysical possibility space, and thereby show us beyond the deadlocked dichotomy of classic determinism and indeterminism.

A. Dennett’s naturalism & critique of “the buck stops here” claims

Dennett is a vehement critic of libertarian free will positions, as these typically point to some arbitrary break in the causal chain. He does not exclude the possibility of indeterminate quantum events nor that some level of indeterminacy or randomized processes might be useful—and thus evolutionarily, culturally and individually selected for—but he argues that no *actual* indeterminacy is needed at any relevant level of description to explain action choice. In other words, Dennett argues that indeterminacy

can be seen as either an irrelevant feature of ultimate physics or as a tool within a determinist story. He is particularly suspicious of the notion of “agent causation,” and the idea of any agent-level claim of indeterminacy. He argues that such agent-as-unmoved-mover, “Deus ex Machina,” “the buck stops here” claims neither make any scientific sense nor—as indeterminism is precisely uncontrolled—seem to give us the kind of “free will worth wanting” (Dennett 1984 & 2003). Dennett has similar objections to other “the buck stops here” claims, and e.g. in connection with his analysis of intentionality, he compares the idea of “unmoved mover” to “unmeant meaners” and irreducible “original intentionality” (Dennett 1989, 288). He sees such claims as a sort of unscientific mysticism or at least as arbitrary and empirically intractable. His project is to show that we can make sense of our experience of action control, lived possibility and responsibility without “magic feathers” just like he thinks that consciousness can be understood without reference to a “Cartesian theater” (Dennett 1991).

In spite of his claims to the contrary, many in the conceptual landscape of the free will debate view Dennett’s claim that our free will is compatible with determinism as incoherent and eventually interpretable as an instance of hard determinism and in effect implying that we do not as a matter of fact have true free will.¹ As we shall see, a central unsettled issue is the question of whether, and in which sense we are free to “do other wise.” Dennett says that even if we cannot do otherwise from some sort of metaphysical Gods-eyes perspective, we have real lived options from our own epistemically limited perspective. Critics suggest that this makes free will an illusion, but Dennett insists that free will depends not on metaphysics but on our lived choices. He argues that our ability to perceive and deliberate about competing action possibilities is integral to action, anticipation and control, and these abilities all depend on eons of gradual processes of biological and cultural selection and horizontal transmission. Thus, our processes of decision-making should neither be seen as illusionary nor as causally epiphenomenal according to Dennett, but rather as an integral part of the causal fabric of everything. Dennett even challenges the view that we should care if we could have really done otherwise under the particular cosmic microstructure. He argues that the presence or absence of metaphysical possibilities would be utterly inscrutable from our lived perspective, and therefore should not cause us any worries. “The kind of free will worth wanting” is not metaphysical but practical. The key is our biological, cultural

1. See e.g. his recent review of Sam Harris’ book see here: http://www.naturalism.org/Dennett_reflections_on_Harris%27s_Free_Will.pdf and for Harris’ reply: <http://www.samharris.org/blog/item/the-marionettes-lament>.

and political ability and freedom to act according to our own—equally biological and cultural—preferences. Dennett thus uses, the Gods-eye point of view to argue for the compatibility of determinism and free will, but challenges the psychological relevance—not consistency—of the metaphysical view, which makes determinism unacceptable.

2. The gradual evolution of freedom

Before further discussing the metaphysical aspects of Dennett’s compatibilism, we need to look at Dennett’s empirical and evolutionary account of the historical complexity underlying processes of human action control. I think that the millennia-old free will debate is in dire need of some empirical details about how we actually do seem to choose and control our voluntary actions. Note though—Dennett is not simply interested in revealing the details of the evolutionary histories, but is also concerned with using these as evidence to show that the whole can come to be “freer than its parts”—and by implication that freedom could have emerged in a determinist world.

It seems to stand to reason that nothing composed of such unfree parts could have any more freedom, that *the whole cannot be freer than its parts*, but this hunch, which is the very backbone of resistance to determinism, will turn out, on closer inspection, to be an illusion. (Dennett 2003, 61)

Though we might not agree on the nature of the “parts,” I share Dennett’s belief that freedom has and can evolve—and that in that sense that the whole can be freer than its precursors. I also agree with his critique of essentialist and unchanging categories. The world is filled with borderline cases of everything, and, as Dennett, I see this variance as a crucial feature of our world rather than something that needs to be explained away.

A. Historical precursors and relational categorizations

In *Freedom Evolves* Dennett discusses the chicken-egg paradox of the “first mammal” and the idea that—even though the parent of a mammal must per definition be a mammal—going back in evolutionary time, at some point our ‘parents’ were not mammals. Does that undermine the fact that we are mammals? He writes:

What should we do? We should quell our desire to draw lines. We don’t need to draw lines. We can live with the quite unchocking and unmysterious fact that, you see, there were all these gradual changes

that accumulated over many millions of years and eventually produced undeniable mammals. (Dennett 2003, 127)

Note Dennett's pragmatic terminology: We "can live with" gradual changes and yet at some point the category is "undeniable." The category, thus, emerges gradually and its eventual factuality is based in the relational response—not an essentialism of inherently necessary and sufficient properties. Dennett's advice to us is to "quell our desire to draw lines." I agree. Drawing neat lines is a powerful epistemological activity that often yields problematic metaphysical inferences. However, I shall suggest that 'line-drawing' i.e. discrimination and discernability, might be the engine of causality and change. As we shall see, I am concerned that Dennett's own "intuition pumps"—by insisting on atomistic/digital (all-nothing) and compositional (part-whole) building blocs at the lowest level of reality—might lose the baby with the bath water. That is, he might forget his own insight that category borders typically are products of line-drawing activities, which themselves have to be explained. More on that later, but his critique of essentialist and unchanging all-or-none categories is well taken.

B. The evolution of normative entities and interests

Dennett's project is to show that freedom can evolve—and not freedom in some sense of random undetermined action—but rather freedom as the ability to act according to one's interests and preferences. But what are *interests* and what does it mean to say *someone* has interests? Immediately we have opened two core cans of worms: 1) the issue of normativity/teleology and 2) the issue of what it means to be a self. Here again, we see Dennett's anti-essentialism and his historical approach. The precursors for what Dennett purposively anthropomorphically calls "interests" did not evolve over night but slowly emerge through complex processes. He writes:

The day the universe contained entities that could take some rudimentary steps towards defending their own interests was the day interests were born. The very tendencies of these organisms to preserve this or that (their varieties of *homeostasis*) helped sharpen the definition of their interests. Only certain sorts of homeostasis tended to be self-preserving in the long run; those kinds were replicated hence persisted, and hence gave further definition to the crude primordial "interest" in self-preservation and self-replication. (Dennett 1984, 22)

Like in the question of the primordial mammal, we see how Dennett wants to give a gradual multipronged story of a complex generative dynamic process of evolution. The circularity and self-referential nature of the explanation is also apparent: What does it mean to be a rudimentary *self*? To be an entity that takes steps towards *self*-preservation and *self*-replication. What does it mean to have rudimentary *interests*? To be an entity that acts to protect (preserve and replicate) those *interests*. What breaks the unfruitful circularity of teleological explanation is its fruitfulness: I.e. the hindsight of preservation and replication. Judged from the present, we can look back and reify gradual “interests” and “selves” as precursors just like we can indentify non-mammalian mammalian ancestors. Note however, the different meanings of self-preservation given the inclusion or exclusion of self-replication. What is it that “persists” and that is preserved? In the case of a single organism, it is the survival of the organism and in the case of persistence through replication, it is the persistence of the organization and historical line of this organization through it’s expression in future individual organisms. Is it a bit more of a stretch to talk about replication-based persistence in terms of *interest* and *self*-preservation? Dennett does not seem to think so, as he is rather comfortable with what he calls the “gene-eye-perspective” (and later “meme-eye-perspective”). As a matter of fact, he often takes this non-phenotypical “perspective” at sub-personal levels. But what would it mean to say that genes are selves with interests? Dennett himself points out the difference between us having interest in, and taking steps towards, preserving say a rock formation and the notion that this formation had interests of its own (1984, 22). Thus, doesn’t it even go against Dennett’s own stated definition to think of genes as carriers of interests, as genes are non-autonomous and thus incapable of taking any, even ever so rudimentary, “steps towards defending their own interests”?

Dennett is well aware that a ‘naked gene’ is not an agent—not even in the primitive precursor sense. This is also clear in his discussion of memes. Dennett follows Dawkins and draws an analogy between biological genes and cultural “memes” as replicating recipes for behavior—or as he writes “a meme is an information-packet with attitude.” Language structures are examples of memes, and speech thus a ‘memetically’-based human cultural activity, transferred between people and generations. But like the “naked genes” of viroids, Dennett stresses that they depended on the human organism to be “expressed.”

Memes depend on human brains as their nesting places: human kidneys or lungs wouldn’t do as alternative sites, because memes depend on the thinking powers of their hosts. Being involved in thinking is a memes

way of being put through the paces and tested by natural selection, just as getting one's protein recipe followed and getting the result out in the world is a gene's way of being tested. If memes are tools for thinking (and many of the best of them are just that), they still have to be wielded for their phenotypic effects to show up. You still have to think. (Dennett 2003, 186)

As we see in this passage, Dennett stresses the dependency of meme and genes on organisms for their phenotypical expression and thus evolutionary "selection," and further he highlights their nature as "tools." But he simultaneously takes what he calls the "gene-eye" and "meme-eye perspective" by writing "getting *one's* recipe followed." One might here want to ask what these purposively anthropomorphic "intuition pumps" are doing for Dennett's argument. Firstly, it seems that he wants to use these analogies to problematize the uniqueness of the human agentic perspective. He wants to give a story of gradual evolution of capacities rather than a radical all-or-none appearance of the human intellect. This part of the project I fully support. However, the question is if Dennett is so focused on the notion of "replication" that he downplays the dynamic processes of the autonomous "self" that is to be preserved or perhaps replicated? Dennett is perhaps using the gene-eye point of view as a way to resisting what he sees as mythical and Cartesian notions of a too substantial personal level self. The gene and the meme are "entities" that can be studied, and might be seen as the "things" that are being replicated. Dennett might use these as fix-points in his attempt to show that personal level teleology, preference and selection has biological precursors, and that "the freedom of the whole can be freer than the parts." In other words, he appears to be using the "gene-eye perspective" to imply that the world, minds and actions can be understood via a sort of atomistic compositionality. Ironically, I shall argue Dennett's seeking of anchoring "things" and "parts" brings in a new set of "unmoved movers" and might obscure rather than help our understanding of emergent "selves," autonomy, teleology and subjectivity. But to make this point, we need to look deeper into the evolutionary story.

C. Horizontal transmission & emergence of Bastardized 'inners'

In Dennett's attempt to give a bottom-up story, he does not just look at the last efficient cause of an action but seeks to understand freedom through the messy web of actual causal histories. A part of this deep history are the big leaps of evolution through horizontal integrations of different lineages, of "self" and parasitic "non-self" entities. Dennett here again draws parallels between biological and cultural horizontal

integrations.² I share Dennett's view that there is a continuity between the biological and cultural domain, and further that the details of the various layerings via horizontal integration are crucial to the multiplicity of timescales of regulation and control of multicellular organisms. I shall focus on his analysis of how the evolution of eukaryotic cells through the invasion and survival of one prokaryotic cell inside of another have allowed for new orders of complexity (2003, 144–146). This evolutionary history of layering is worth dwelling on as it might help us understand some basic dynamic characteristics of self-regulating organisms. One such question that Dennett discusses is how prokaryotic cells somehow already came with much more complex gene-regulation processes than they needed. He writes:

Fancier eukaryotic cells, however, to say nothing of us multicellular types composed of these more complex building blocks, need a mind-boggling elaborate system of intermediate steps, checks and balances, so that genes can be turned on and off at appropriate times by the indirect effects of other gene products and so forth. For some time biologists had a classic chicken-and-egg puzzle to contend with: How did this elaborate gene-regulation machinery evolve? Multi-cellular life couldn't even begin to evolve until most of this expensive machinery was in place, but it apparently isn't required for prokaryotic life. (Dennett 2003, 146)

In other words, how could these much needed cellular control mechanisms be in place prior to the first eukaryotic pre-cursor cells? Such control mechanisms, balancing and adaptively applying tools given changing circumstances, allow primitive multicellular organisms some level of autonomy and ability to have responses selected from various repertoires. Thus, their evolution might give us an important clue also to the understanding of our human-level abilities of action control. Dennett explains the emergence of this sort of cellular complexity in prokaryotic cells as follows:

The answer that is now emerging is that it was paid for by a civil war that raged for roughly a billion years of early prokaryotic life. It was an arms race within the genome, with good citizen genes doing battle with those transposons – free loaders who copied themselves repeatedly in the genome without providing any benefit to the whole

2. He writes e.g. "Horizontal transmission of design, of information that can be put to good uses, is a key feature of human culture, and undoubtedly the secret to our success as a species" (Dennett 2003, 145).

organism. This created lots of measures and countermeasures, such as silencing mechanisms and isolation-defeating mechanisms. (Dennett 2003, 146)

This is indeed an interesting history to understand. Note the idea that the large jumps in evolutionary novelty became possible due to active processes of stabilizing the cells in the face of inner turmoil at an earlier stage. Further note, that when brought under control this intruder-created nuisance produced new potential capacities, which would then later be exploited to the organism's advantage in entirely new contexts. A case of what doesn't kill you makes you stronger—here in the specific sense of having expanded the repertoire of the 'self'—and thus, response possibilities. Such inner complexity is expensive as Dennett says, but also in a broader sense opens you up to new kinds of vulnerabilities. I therefore like to call these events "bastardizations," implying at once the multiple origins, otherness, and the always-looming possibility of indigestion, exclusion, isolation, general friction and precarious reception. But it also implies transformation, creative novelty and the birth of a truly new and unique self, the main challenge to whom comes from the quest of social recognition i.e. fitting existing categories. It is thus always a fragile balance of how much can be metabolized without collapsing the autonomy organism.³ Ironically, our subjectivity and great evolutionary assent to freedom has come through a series of debasing and tumultuous internalizations of others.⁴ But this is of course me projecting my interpretation onto Dennett, as his metaphors are generally more about wars, infections and peaceful co-existence than about illegitimate children and the precariousness of subjectivity. He writes:

The eukaryotic revolution draws attention to the fact that even in biological evolution, which Darwin aptly called "descent with

-
3. One might also suggest that the emergence of inner turmoil in moving organisms might have presented a puzzle of how to create coherence and negotiate the now necessary frame-shifts between inner and outer concerns. A puzzle to which conscious feelings and perceptions might have been an evolutionary solution.
 4. We are at present exposed to a radical cultural horizontal transmission through the technological inundation of information and surveillance. These things give us powerful new action repertoires and reach but also transform us—and our worlds—at runaway speeds. The borders of 'selves' and 'societies' are changing in ways that have significant effects on our autonomy, and if we don't evolve new 'membranes' to help us control, regulate and hide our inner multiplicities, then we might very well end up adapting by moving the prime locus of autonomy to a societal level. In other words, we are in a period of externalization of our inners, which as we shall see might represent a loss of freedom in the sense of a loss of the ability to hide.

modification,” there is plenty of room for horizontal transmission of design. The prokaryotic hosts who were first “infected” by their symbiotic visitors got a huge gift of competence designed elsewhere. That is, they didn’t get all their competence by vertical descent from their ancestors via their parent and grandparents and so forth. They didn’t get all their competence from their genes, in other words. They did however pass on this gift to all their offspring and grand-off spring through their genes, since the genes of the invaders came to share the fate of the nuclear genes of their hosts, traveling side by side into the next generation, which was infected at birth, one might say, with its own complement of symbionts. The clear trace of this dual path is still highly salient today, in all multicellular creatures, including us. Mitochondria, the tiny organelles that transform energy in each of our cells, are the descendents of such symbiont invaders, and have their own DNA. Your mitochondrial DNA, which you get only from your mother, exists in each of your cells, alongside your nuclear DNA – your genome. (Dennett 2003, 145)

Given this passage, the crucial question is when something becomes part of “me.” Dennett’s terminology here suggests that ‘I’ am my genetic code but that ‘I’ depend on these ancient invaders, which are replicated in my offspring by an “infection at birth” but stay forever foreigners traveling “side by side” as *their* DNA has never merged with *mine*. But isn’t an alternative interpretation of the same biology, that I exist as a genetic multiplicity, that *my* genomes are a combination of both the mitochondrial and the nuclear DNA? After all, even the nuclear DNA is a bastardization as mentioned above. Interestingly Dennett leaves out the story of what one might consider life’s ‘original bastardization’. One candidate could here be the first trapping of RNA (possibly itself imported via meteorites from Mars⁵) in primitive lipid vesicles, and thus perhaps a primordial self of *self*-replicating life. However, the question is if Dennett would be more prone to think of the earlier RNA duplications as already self-replicating. The difference is that only with the event of the proto-membrane can we talk about some sort of inner environment, and also of clear processes of self-maintenance and autopoiesis. I shall not suggest that we can discern a primordial mammal or a truly original bastardization of life, but merely highlight that Dennett’s evolutionary accounts have a tendency to downplay

5. See Grossman & Webb (2013).

or sideline the dynamics of boundaries and habitats in the evolution of self-replicating individuals. He does in the quote above talk about the necessity of the wider “molecular machinery,” in bringing about the phenotype. But yet he talks as if the gene/meme is the individual or the ‘entity’ that is being (self)-replicated, and thus, as somehow the locus of “interests” and somehow a “self.” In brief, he spends a lot of time on genes and memes as recipes for self-replication, but comparatively little on their contexts. References to membranes, time and autopoiesis are suspiciously missing from the index of his books— as if he tells the story of the chicken but forgets to mention the egg.

These terminological choices are of some consequence when considering the issue of what it means to be an agentive self— and whether the DNA is more important than the boundary that actively establishes and maintains a difference between “self” and “other” not by any inherent pre-given essence or difference but via on-going dynamic and non-instantaneous interchanges. With the gene as the locus of the proto-self self-replication, Dennett avails himself of a relative temporal stability and passive self-containment. If we on the other hand focus to the entire cell or organism we can see that its stability, maintenance and autonomy is a precarious, active and contextual process in constant flux.

D. The spatio-temporal agent is the loop

As we have seen already, Dennett wants to dispel the need for extra supernatural “magic feathers.” He sees our abilities to recognize, anticipate and select actions as tools to maintain some preferred homeostasis or outcomes that we have developed and gradually selectively shaped through both evolutionary time and though our cultural environments. Based on analyses like the above one of the eukaryotic revolution, Dennett defends the idea that our human-level freedom can be understood analogously, as much fancier ways of using meme invader and “trapping reasons,” for purposes of intricate forms of anticipation and avoidance. He argues that all this is possible to understand without any super-natural assumptions or breaks in the causal fabric of reality; “the whole can be freer than its parts,” and we might add *the present more free than it’s past*:

Events in the *distant* past were indeed not “up to me,” but my choice now to Go or Stay is up to me because its “parents” – some events in the *recent* past, such as the choices I have recently made – were up to me (because *their* “parents” were up to me), and so on, not to infinity, but far enough back to give my self enough spread in space and time so that there is a me for my decisions to be up to! The reality of the moral

me is no more put in doubt by the incompatibilist argument than is the reality of mammals. (Dennett 2003, 135–136)

We see here, the analogy between this notion of autonomy of something being “up to me” and Dennett’s primordial mammal argument. The purpose of the analogy might be to pump the intuition that autonomy too is not only a category with degrees—an intuition I share—but also that it is a historical lineage category. I also agree that autonomy is historical, but like we saw in the discussion of the gene eye-perspective, I think we are genetic bastards and that autonomy is continuously re-negotiated in the present. Thus, I am hesitant to endorse the idea that whether my current action is up to me depends that strongly on whether its precursors has the proper kind of history.⁶ However, Dennett also adds that the question of whether we see something as being up to us depends on whether we go “far enough back to give my self enough spread in space and time so that there is a me for my decisions to be up to!” This point about how big or small a perspective we take on our selves is really fascinating and something that Dennett returns to multiple times in discussing moral responsibility.

The general idea of a both spatially and temporally extended self also plays a key role in Dennett’s argument against the incompatibilists, as it is via this notion that we become the authors of our actions and that we play a consequential causal role in the world. Dennett argues that the inference, that our heart-wrenching deliberations, thoughts and actions would be obsolete in a determinist world, comes from the idea of an extensionless and non-material self. In accordance with his critique of the Cartesian theater, Dennett is similarly critical of the idea of some sort of localizable atomistic conscious self, which can be neatly contrasted with the non-self aspects of our physiology. Hence, in his discussion of Libet’s findings and assumptions about the causal role of the conscious self he concludes: “What Libet discovered was not that consciousness lags ominously behind unconscious decision, but that conscious decision-making takes time” (2003, 239).⁷ And a bit further down he writes:

When we remove the Cartesian bottleneck, and with it the commitment to the ideal of the mythic time *t*, the instant when the conscious decision happens, Libet’s discovery of a 100-millisecond veto

6. Note the similarity to Millikan’s teleosemantics and historical notion of “proper function”—and perhaps the challenges her account faces due to its disregard for the influence of synchronic organization. (Millikan 1989).

7. Gallagher (2006) has made similar criticisms of the implausible temporal assumptions of the Libet studies.

window evaporates. Then you can see that our free will, like all our other mental powers, has to be smeared out over time, not measured at instants. (Dennett 2003, 239)

I couldn't agree more, both with the claim that the notion of a "time t" for action is a myth, and the claim that free will has to be smeared out over time. In the further explanation, we can really see the central role the extended self plays in Dennett's compatibility theory both our freedom and responsibility. He writes:

Once you distribute the homunculus (in this case, decision making, clock watching, and decision-simultaneity-judging) in both time and space in the brain, you have to distribute the moral agency around as well. You are not out of the loop; you are the loop. You are large. You are not an expansionless point. What you do and what you are incorporates all these things that happen and is not something separate from them" (Dennett 2003, 241–242)

This is an extremely important idea—if we are not Cartesian dualists or in other ways thinking of free will as external to the material world and the biological body, then, in so far as "we" exist, it is natural to think of our deliberations and decision-making as not only being *in* the causal loop but *being* the loop. It is thus by way of this idea that Dennett claims that his account of voluntary action indeed makes actions, not only up to us, but causally dependent on us. Further, he writes, "you have to spread the moral agency around as well." Thus, the degree to which we see ourselves as large or small corresponds to our respective internalization and externalization of the responsibility. His point seems to be mainly a question of perspective: if we on the one hand see ourselves from the outside we tend to look at the multitudes of "external" causes of our action. But if on the other hand we spread ourselves out, then those same causes seem to "internal" to ourselves and they would thus be our responsibility. The question is whether this leads to an unnecessary moral relativism. Are all perspectives equally morally valid? In other words, does Dennett in his eagerness to avoid the homunculus again forget about the semi-permeable membranes and the phenotypes selection works on?

E. Appearance and reality—the causal effects of hiding and revealing

I want now to highlight another dynamic—and possibly constitutive—aspect of autonomic evolved selves. Namely, the ability to selectively hide and reveal aspects of both our inner and outer worlds and thus purposively control interactions. Dennett

himself talks about intentional deception as an important mental capacity as well as of course our ability to track and gage when we and others are being deceived. He proposes that the difference between appearance and reality is “fatal” to all organisms, but only we humans have the ability to reflect on and deliberately “bridge the gap.”⁸ Dennett is probably right that our ability to voluntarily take steps to make reality appear and not appear sets us apart. However, he is also well aware that many unconscious steps towards controlling the gap between appearance and reality are at the core of the evolution of life. He writes e.g.: “Mother nature abides by the “Need to Know” principle” (1984, 24). Thus the issue is not simply the human ability to deliberately “bridge” —i.e. eradicate— the gap but rather to purposively control and exploit the gap.⁹

Further, I would argue that one could see this control as a core ingredient in the emergence of teleological evaluative selves and thus as a key to any evolutionary story of the freedom of action. The question is if it isn’t the “hiding” of the genetic material behind a membrane that allows for steps towards self-preservation? If the world simply triggers the replication is it self-replication? Independently of how one decides to categorize the primordial mammal or the beginning of self-replicating life, I think one gets the picture wrong if evolving processes of control are not seen as dependent on and contributive to abilities of hiding and revealing.

Dennett’s story of horizontal transmission in both biology and culture is fascinating to see against this background—as we then see that the self-other dynamic is one of control. What is “me” is not my just my nuclear genome and it is not the conglomerate of my human stem-line cells with both mitochondrial and nuclear DNA, but rather a much more genetically and culturally messy—and constantly changing—set of processes. Are the symbiotic microbes of my digestive system part of me? Is my foot? Is my language?

8. “The difference between how things really are is just as fatal as a gap for them as it can be for us, but they are largely oblivious to it. The recognition of the difference between appearance and reality is a human discovery. A few other species—some primates, some cetaceans, maybe even some birds—show signs of appreciating the phenomenon of “false belief”—*getting it wrong*. They exhibit sensitivity to the errors of others, and perhaps even some sensitivity to their own errors as errors, but they lack the capacity for the reflection required to dwell on this possibility, and so they cannot use this sensitivity in the deliberate design of repairs or improvements of their seeking gear and hiding gear. That sort of bridging of the gap between appearance and reality is a wrinkle that we human beings alone have mastered” (Dennett 2003, 165). Note that Dennett’s thoughts on human and primate abilities to appreciate false belief have lead to a whole industry of test paradigms. See also Brincker (2014) for a recent theory of the development of “false belief” implicit and explicit performance in humans.

9. See also Dennett points about self-deception (1984, 48).

Are my contacts? The point is not to answer these questions in the abstract. The point is that our bodily reactions, engagements and actions shape the answers as they answer. I am always a self—or rather selves—in the making. My complexities and messy causal histories determine me and yet scaffold my freedom.

F. Options & evitability

Dennett thus, builds forward to his notion of freedom as the ability to recognize, anticipate outcomes and actually choose what you prefer—given your also evolved preferences.

This process, he argues, can be seen as entirely deterministic: You recognize what he calls “a “special interest opportunity” and given your anticipation of outcome compared to your preferences you select either to act towards or to avoid. He writes

So a real opportunity is an occasion where a self-controller “faces” —is informed about—a situation in which the outcome of its subsequent “deliberation” will be a decisive (as we say) factor. In such a situation more than one alternative is “possible” so far as the agent or self-controller is concerned; that is, the critical nexus passes through its deliberation. (Dennett 1984, 118)

Thus, Dennett takes possibilities for action to be recognized opportunities that one could either decide to pursue or not. Accordingly, he argues that determinism does NOT, as most claim, make actions “inevitable.” Rather, we have evolved—possibly in a determinist fashion—to be very fancy evators or avoiders.

Both hard determinists and Libertarians here typically object and say that this is not true “evitability” as you “could not have done otherwise”! Dennett says no not in *these exact circumstances*, given these precise preferences, perceptions and anticipations this precise action was selected by your “will” if you want. Had it been a different—even ever a so slightly different—universe you could have done otherwise. And Dennett says that we always sneak in such different circumstances when we say we could have acted differently. Thus e.g. *yes had I wanted to* I could have picked a different action...but his point is that this different volition would exactly be a product of a different causal world—not this world.

I will in the following sections analyze Dennett’s metaphysical arguments and the intuition pumps he uses to defend this idea that other outcomes could only come about in other possible worlds. A closer look seems to reveal a tension within Dennett’s compatibilism. As we have seen in his evolutionary story he points to 1) the historicity

and precursors of categorizations, and the causal efficacy of 2) our “loop” of perspectival epistemic limitations, spread-out layered selves and 3) of the gap between appearance and reality. Yet, he assumes a metaphysical base level of non-perspectival causality where appearance and reality always neatly coincides. I shall suggest that he overlooks a set of metaphysically possible worlds in which all causal interactions are limited and there thus—even “under the aspect of eternity”—is no base-level, but always precursors and “back stages” to use Dennett’s own term. But I get ahead of myself.

3. Dennett’s metaphysical view from eternity

Dennett argues that from our lived perspective we have multiple possibilities for action, but these are not possibilities to actually change the course of the world. Quoting from *Elbow Room*:

But if we want to change the course of history we are in for a big disappointment. For no one can change the course of history – for reasons that have nothing to do with determinism. At the beginning of the chapter we imagined all of space and time, past, present and future laid out before us (“*sub specie aeternitatis*” in philosophical parlance: under the aspect of eternity). If the scene we thereby imagine is supposed to be the *actual* course of history through eternity, then – look, and see – the image has no branchings. Only one actual thing happened whether or not what happens is determined to happen, so the part of our image we label “Future” consists of the events that actually happen – happen to happen – in the fullness of time. (Dennett 1984, 124)

Dennett is right that branches are hard to make sense of—as Bergson (1889) pointed out their imagery assumes the hindsight of all the alternative actions as actually having been carried out. However, many would say that the kind of free will worth wanting is precisely one whereby my actions can change or add to the course of world history. Dennett would likely respond that my actions do participate in the causal history of the world and thus influence it, but “under the aspect of eternity” I have changed nothing. I argue that this view from eternity is a fancy vending machine.

A. The metaphysical stance

I have admittedly always been struck by the oddness of the age-old “metaphysical stance,” above conceived of as “imagining all of space and time” “laid out before us.” It is supposed to be the ultimate exercise of objectivity of abstracting away our situated perspective—and yet it seems impossible to formulate this metaphysical notion in abstraction from some sort of perceiver or subjective point of view. This feature of an entirely external non-specified perceiver is clearly seen in the expression, “the view from nowhere” and the “God’s-eye perspective”—or here as Dennett writes “under the aspect of eternity.” The purpose of the non-specificity of the perceiver is presumably that any specification would impose a limitation to the view and thus prevent the access to the totality of being, the “Ding an sich.” It is thus a paradoxical mind-bending activity of imaging the mind-independent world: the appearance of reality independently of appearance, i.e. the ultimate bridging of the gap between appearance and reality.¹⁰ But what could it mean to “see” (not to mention smell), the entire “actual course of history”? As Akins (1996) reminds us, our senses are narcissistic, and it seems that the metaphysical stance invites us to abstract from *any kind of categorization, recognition or reification process* that we would rely on in actuality. This “metaphysical stance” exercise is at least as old as the ancient Greeks and likely as old as metaphysical thinking, and is constantly appealed to in the free will debate as a way of exposing the inconsistency of a determinism and the notion of free will as the ability to do otherwise or “agent causality.”

It should be stressed that it is a pivotal part of Dennett’s own compatibilist argument that we can *never actually* reach a God’s eye point of view—and that this epistemological limitation is precisely what makes our action-opportunities real from our perspective. Laplace’s omniscient intellect is impossible in reality and this makes a difference to Dennett, as it guarantees the practical unpredictability of the world. Even if the world is entirely determinist, as long as it is inscrutably so and forever hidden, it is irrelevant to our freedom.

Yet, Dennett appeals to this admittedly impossible vantage point to deny the very possibility of metaphysical possibility. So, we need to understand how this image is doing

10. See also Putnam’s critique of what he calls the “externalist view.” He argues that there cannot be “exactly one true and complete description of the ‘the way the world is’...there can be no God’s eye view of reality” and writes further: “What we have here is the demise of a theory that lasted for over two thousand years. That it persisted so long and in so many forms in spite of the internal contradictions and obscurities which were present from the beginning testifies to the naturalness and the strength of the desire for a God’s eye view” (Putnam 1981, 74). See also Hornsby on the nomological character of causality (Hornsby 1997, 78–80).

work in his argument. I will therefore look closely at his metaphysical images and models, which are used 1) to evoke the idea that the future is closed, that we can change nothing and 2) to explain the determinism, which he claims to compatible be with lived freedom.

B. Laplace's infinite intellect

Before we turn to Dennett's use of the metaphysical stance, I want to say a bit about Laplace's famous "metaphysical stance" image of determinism. He writes: "We may regard the present state of the universe as the effect of its past and the cause of its future." And to explain this determinist causation from "the present state" to the "it's future," he continues:

An intellect which at a certain moment would know all forces that set nature in motion, and all positions of all items of which nature is composed, if this intellect were also vast enough to submit these data to analysis, it would embrace in a single formula the movements of the greatest bodies of the universe and those of the tiniest atom; for such an intellect nothing would be uncertain and the future just like the past would be present before its eyes. (Laplace 1814)

Thus, the way Laplace conceives or "sees" the past as fixing the present and future is through an exhaustive knowledge of a time-slice of a frozen universe set in predictable motion. Note that the role of the "intellect" is as the measuring stick, the background (or foreground) in relation to which—before which—the world is arrested and everything is reified. In other words, the picture does not work without its implication of an intellect or some kind of "reifyer" which—as any measuring stick—itself is external to the universe, not included, not itself "given." It is the full transparency (appearance=reality) of everything *for* the intellect that allows for the exhaustiveness of the knowledge and thus predictive completeness. With limited and perspectival transparency uncertainty would not be eradicated and the world would appear only partially and thus not be transparently "present" as a totality. Interestingly, Laplace's image imposes a temporal limitation to the otherwise vast intellect. The intellect need not know everything "under the aspect of eternity" but only know everything at "a certain moment," and then given the deterministic analysis of forces and motions the temporal transparency appears: everything becomes "present before its eyes."

The notion of time slices, moments, instants, as we shall see in Dennett's analysis, seem to be a fixture of determinism formulations. Laplace takes the instantaneous transparency of everything for granted. But it is an open question if all causalities

and regularities of “the greatest bodies of the universe and those of the tiniest atom” could possibly be arrested and “revealed”—even to an infinite intelligence—on one homogeneous and near instantaneous temporal scale. This is important, as the purpose of image is to argue that *the actual world* we live in—behind the appearances—is fully determined and timeless in the sense that it could have been predicted and instantaneously “present before” us if we knew the initial conditions of the universe.

C. The view from eternity & the denial of possibilities

As we have seen, Dennett proposes a historical, empirically-based theory of how our ability of voluntary action has evolved and how it can be understood as arising within a causal understanding of nature. Free will, he argues, does not break the causal web as most have assumed, but rather acts within it. However, note that his metaphysical view of the world does not depend on determinism. He writes: “The image we should have when we imagine this universe to be nondeterministic ought to be indistinguishable from the image we imagine when we conjure up a deterministic universe through time” (Dennett 1984, 124). How so? The key to Dennett’s suggestion, that the future’s closedness does not depend on determinism, seems to lie in his eternalist “bloc universe” view. Though a defender of determinism as compatible with free will, he is agnostic about whether the world from an inaccessible point of view unfolds in a determined or indetermined fashion. Thus, the reason why he sees the metaphysical stance as excluding changes to the course of history is *not* that we are caught in a determinist causal web, in which all actions in theory could have been predicted from initial conditions as Laplace’s famous thought experiment would have it. He continues:

If the past is unchangeable, the future is unavoidable – on anyone’s account. The future consists, *timelessly*, of the sequence of events that will happen, whether determined to happen or not, and it makes no more sense to speak of avoiding those events as it does to speak of avoiding the events that have already happened. (Dennett 1984, 124)

Under the eternalist view, there is no passage of time, just a totality of moments or temporal extension. Time is a dimension like space, and therefore the past and future are equally fixed and “timelessly” existing. Dennett asks us to “look” at the totality of time: “past, present and future laid out before us.” Note that Dennett here is assuming a different extent of transparency (appearance=reality) than Laplace did. He immediately jumps to the view from eternity, which Laplace only gained through the application of deterministic analysis. From that vantage point, causes and laws are irrelevant and

the indeterminist and the determinist universe are equally transparent. So, what has Dennett shown with his eternalist view? That the passage of time changes nothing? That the “present” holds no possibilities, nothing more than what “will happen”? It seems to me though, that these conclusions are assumed in the formulation of the timeless eternalist picture. In other words, isn’t the claim that there could be local possibilities, which can cease to exist, precisely denying that one can see everything from nowhere, from eternity, from God’s house, you name it? If one assumes that *everything* is “seen” from this metaphysical stance then one limits the possible metaphysical options accordingly. Only worlds of full eternalist transparency are possible. The transparent moments could be “sequenced” in an either determinist or indeterminist fashion, but there shall be nothing hidden, no time, no change of what has already been assumed. Any view that takes time—and emergence for that matter—to be real is excluded as impossible. Thus, Dennett like Laplace seems to use his metaphysical stance to assume what he wants to conclude. He assumes that reality across eternity could be given fully transparent and coherent. Hence, he claims, there were never any other possibilities.¹¹ In “the fullness of time” he writes above, only the actual history will exist. That might be so, but isn’t it at least possible that even God only sees a partial truth? Could one imagine that some possibilities—or precursors—“existed” (in a non-reified sense), but never had any consequence¹² and thus are inscrutable, hidden even from the “fullness of time” perspective? This is a hard idea to entertain because we can never describe any such possibility without reifying and recognizing it, and thus installing it forever in the causal web. We might not therefore be able to conceive of this metaphysical idea from the metaphysical stance, but the classic formulation of the paradoxical stance has problems of its own and perhaps shouldn’t be our measure of the possible.

D. Discrete and compositional determinism

Dennett does not just want to argue that only the actual course of history exists, he also wants to show that determinism is compatible with his notion of free will as evolving. Over the course of his books he presents different formulations of what

11. Dennett’s position sounds and smells like a version of actualism, i.e. the position that only what is actual is possible. However, he vehemently argues that it is not—because he allows for lived and situated possibilities also interpreted in terms of capacities of objects. See (Dennett 1984, 144).

12. I shall not here be able to develop an emergentist or interactivist view of causality, but note that already Hume plays around with the idea that causing means producing, and that a cause can only be reified or understood through its effects, it is the “by which” (Hume 1750).

determinism is and I shall look at a few of them. In *Freedom Evolves* he starts out his discussion with Inwagen's definition that the claim of determinism is that "there is at any instant exactly one physically possible future" (Inwagen 1983, 3). He writes that this "not a particularly difficult idea." I beg to differ. As discussed in relation to Laplace, it is not clear how we are to actually think of "instants" that somehow contains something that fixes "physical futures"? Again a crux of my worry has to do with the timelessness of the material universe assumed under classical determinism, and the notion of a universe "time slices," "instants," "time t" etc.¹³ We saw that Dennett himself talks about the notion of "time t" as a "mythical ideal" when it comes to the empirical study of the mind, however he seems perfectly content with it when defining determinism. The idea of an all-encompassing and extensionless time-slice of the arrested universe is in many descriptions of determinism operationalized as a "state description." Dennett writes: "A universe is *deterministic* if there are transition rules (the laws of physics) that *determine exactly* which state description follows any particular state description. If there is any slack or uncertainty, the universe is *indeterministic*" (Dennett 2003, 28–9). But interestingly Dennett quickly gives up on this definition as a useful tool for his argument. He writes: "There are too many fudge factors in this simple vision as it stands: How exact must a state description be? Must we plot every sub-atomic particle, and just which properties of the particle need to be included in the description?" (Dennett 2003, 29). I agree! The classical definition might not be feasible due to the challenge of pinning down the ultimate reality—or its "rules." Dennett of course sees this as merely a question of practical feasibility, and he then solves the issue by going a way from a material world model to a digital model without any "fudge." He writes:

We can anchor these slippery factors arbitrarily by adopting another simplifying idea, W.v.O. Quine's proposal that we restrict our attention to simple imaginary universes, which he calls "Democritean" universes, in honor of Democritus, the most inventive of the ancient Greek atomist. A Democritean universe consists of some "atoms" moving about in "space." That is all. The atoms in a Democritean universe are not modern atoms full of quantum complexities but truly *a-tomic* (unsplittable, unsliceable) atoms, tiny uniform points of matter with

13. It is beyond the scope of this article to enter the deep debates within the philosophy of time, but I do want to motivate free will theorists to interrogate their typical treatment of the world as timeless. It is particularly ironic when the question at hand is how we as agent might effect *change* in the universe. For great discussion of the problems of the denial of temporality see Crome (2007) & Smolin (2013).

no parts at all, rather like those postulated by Democritus. The space they inhabit must be made ultra simple, too, by *digitizing* it. (Dennett 2003, 29)

What has been described here? Dennett suggests that these “unfudged” imaginary all-or-nothing, uniform, ahistorical unsliceable atoms are a kind of “matter.” But in which sense? He also writes that they “move” in digitized space, but how so? Dennett has reduced the “universes” to sets of successive “states.” In the quest to “unfudge” he seems to have defined away materiality, temporality and change. But do we not precisely need the fudge to portray a material and temporally changing world—i.e. is he losing the material world he seeks to model in these first simplifying moves?¹⁴ Dennett does not seem to think so. He uses this Quinean notion of Democritean universes to construct a “vast” but finite possibility space of “all possible worlds,” namely all possible combinations of state descriptions in the digitized space given a predefined finite number of successive “instants.” Pacing Borges, he calls this the Democritean Library. Each world represents one view from eternity, just simpler—and finite—as time and space are here explicitly discrete, and matter digitized. Given this finite Library of finite logically or “physically” possible and fully transparent universes, Dennett can now clarify Inwagen’s definition. Determinism holds in the subset of logically possible worlds in which:

“There is at any instant exactly one physically possible future. To say that determinism is true is to say that our actual world is in a subset of worlds that have the following interesting property: There are no two worlds that start out exactly the same (if they start the same, they stay the same forever – they are not different worlds at all), and if any two worlds share any state description exactly, they share all subsequent state descriptions.” (Dennett 2003, 67)

Again, as we have already seen, the eternalist view does not pertain uniquely to determinist universes but relies simply on the fact that everything is fully transparent and scrutable in each possible world. The news is that a “set” based definition of determinism is now not based on fixed transition rules, but rather on the notion that we browse the Democritean library, and compare and contrast “worlds.” The very idea of a “world”

14. This is by no means a new insight but a paradox, which has been known at least since Parmenides and Zeno, and Democritus atomism was a compositional attempt to reconcile timeless being with our experience of time and change. See Crome (2007) for a fascinating analysis of the classic paradoxes and the conclusion that making sense of time and change might involve going beyond discrete math.

depends on the discerning intellect in the metaphysical stance, but now we are also asked to make “sets” of “possible worlds.” This requires a meta-metaphysical stance, as I now need to step back from the individual world and imagine the whole vast Library of possible worlds to find out, which worlds have respectively overlapping and differing states. If two possible worlds after one overlap stay the same through eternity, then they are determinist worlds. With this definition the determinism—or indeterminism—of our actual world can be settled on the basis of sets of entirely discrete and unchanging digitized sequences, without any of “fudge” of describing or applying laws to create movement and change. In other words, the meta-metaphysical stance allows us to categorize across worlds and thus define determinism without ever introducing matter in lawful motion. With this toolbox Dennett’s challenge is just to show that life can evolve in such “worlds”, and this is where Conway’s game of life comes into the picture.

E. Conway’s game of life

One of Dennett’s favorite models to show how complexity and avoidance can come from determinist underpinnings is Princeton Mathematician John Conway’s digital “game of life.”¹⁵ The game consists of 1) a very simple finite two-dimensional world—a grid, with cells capable of being ON or OFF and discrete temporal instances. 2) A set of initial conditions where some and not all cells are ON - to get the “physics” going and 3) a set of very simple timeless and universal rules, which Dennett calls the “life Physics”:

Life physics: For each cell in the grid, count how many of its eight neighbors are on at the present instant. If the answer is exactly two, the cell stays in its present stage (ON or OFF) in the next instant. If the answer is exactly three, the cell is ON in the next instant whatever its current state. Under all other conditions the cell is OFF. (Dennett 2003, 36)

How does the transformation take place? How are the rules followed? From the perspective of each of the finite number of cells—we get an “input” of 1) number of ON cell neighbors and 2) ON or OFF status - which is categorically adjusted according to the rules of the “life physics.” Note again that appearance and reality always neatly coincides. Dennett actually acknowledges this and writes “One of the delights of the Life world is that nothing is hidden in it; there is no backstage.” It is a neat world of zero ambiguity,

15. See Gardner’s (1970) for the original description in *Scientific American*.

where nothing exists, which is not recognized and responsive to the “life physics” at each instant.

Dennett, as mentioned, uses this simplified universe as an “intuition pump” for us to understand that life-like complexity can emerge from utterly deterministic rules. He points out that from an “intentional stance,” we can recognize rather interesting patterns when we look at the unfolding changes of the “life world” from various distances and tempos. And he sees these higher-level patterns as in a sense real. Here is a quote from his discussion of pattern recognition in *the Intentional Stance*:

Is the pattern that enables you to make the prediction “real”? So long as it lasts it is... The pattern may owe its existence to the intentions (clear-sighted or confused) of the machines designer, but its reality in any interesting sense – its longevity and robustness – is strictly independent of the historical facts about its origin. (Dennett 1987, 39)

The pattern is real and its reality is in its longevity and robustness. But what would that mean in the “Life world” as these patterns are indiscernable to the “life physics”? Is the reality in their availability to be recognized by *us* as external viewers? He writes:

Whether one can see the pattern is another matter... I claim that the intentional stance provides a vantage point for discerning similarly useful patterns. These patterns are objective – they are *there* to be detected – but from our point of view they are not *out there* entirely independently of us, since they are patterns composed partially of our own “subjective” reactions to what is out there; they are the patterns made to order for our narcissistic concerns (Akins 1986). (Dennett 1987, 39)

Thus, pattern recognition is relational as it is “composed partially of our own “subjective” reactions.”¹⁶ The “interesting” patterns of the game of Life also clearly depend on us as viewers, the distance and speed by which we watch, as well as all the other recognition repertoires we bring with us (think Rorschach test). But is reification also relational at the “physical level”? It seems that the ON and OFF status of the cells must depend on their “recognition” as such, not by us but by the “life physics.” In this

16. His description of the patterns and their objectivity and yet relational dependence can be seen as somewhat analogous to Gibson’s view of affordances as objective but yet relative to the perceiver in their reification. See e.g. Gibson (1977).

connection it is curious that another of Dennett's favorite toy models is the "two-bitser"; a soda vending machine that recognizes and responds to certain coins and not others (Dennett 1987). Here he seems to propose, along the lines of Quinean inscrutability,¹⁷ that if a two kinds of coins ("fake" and "real" quarters) cannot be discriminated, then there appears to be no further fact of the matter of which kind they are. The ultimate test is the actual interaction between coin and machine: if the soda comes out then it is a quarter.

In the two-bitser-case the interactions depend on a wealth of material details and complexities of the coins and the machine—which go beyond the recognition itself. I.e. no coins are actually identical material tokens and yet treated as such from the limited (and limiting) "perspective" of the machine. The material opacity has been eradicated in the game of life, where there is "no back stage" for either grid, cells or physics that goes beyond the recognition itself. The "Life worlds" are fully transparent two-dimensional worlds. Appearance equals reality and *every existent difference* is recognized and responded to with transition rules by the "physics," the omniscient Gods-eye-grid.

Dennett introduces the 'game of life' to show that complexity can emerge from strict physical determinism. But as even the "physical level" is digital it is not much like a material world in any traditional sense. There are not any persisting atoms nor any other "matter". Rather the cells go through instantaneous birth and annihilation without any component precursors or remains. Ex nihilo—in nihilo—like Leibnizian monads.¹⁸ No continuous becoming and only something like instantaneous symbolic being. This digitalized all-or-nothing idea is necessary in Dennett's (and Leibniz') model to "unfudge" and stop the regress of further compositionality.

The question that I would like to raise is whether it is plausible that the ultimate level of universe would have any of these "life world" features? Note all the "unmoved mover" and "unregulated regulator" aspects of the Game of life. Where do the eternal

17. See e.g. Quine (1960).

18. Interestingly Leibniz monads are in many ways like digital pixels. See his monadology 1–6, in particular: "3. Now, where there are no constituent parts there is possible neither extension, nor form, nor divisibility. These monads are the true atoms of nature, and, in a word, the elements of things. 4. Their dissolution, therefore, is not to be feared and there is no way conceivable by which a simple substance can perish through natural means. 5. For the same reason there is no way conceivable by which a simple substance might, through natural means, come into existence, since it can not be formed by composition. 6. We may say then, *that the existence of monads can begin or end only all at once, that is to say, the monad can begin only through creation and end only through annihilation.* Compounds, however, begin or end by parts" (Leibniz, [italics mine]).

laws of physics come from? Where does the space grid come from, where do the pixels' coloring come from or disappear too? How are "instants" aligned, i.e. how are we to understand simultaneity or the idea of the "state of the universe at time t "? Dennett is a devout naturalist but yet these assumptions instill a finite border beyond which the "why" question does not apply. Dennett could here remind us that this model is purposively "unfudged." But as we have seen in our earlier analyses, it is not just Conway's model that introduces "unexplained explainers"—all assumptions of this "no back stage" collapse of appearance-reality collapse do. In other words, all Dennett's proposed metaphysical images do.

Beyond the question of the plausibility of "the buck stops here" assumptions of the framework, there is Dennett's question of whether the game of life could succeed in simulating evolution and freedom. In other words, returning to the issue of whether an unfudged model can show the complexity of life and pass the "Turing test" of life so to speak. Remember that in the two-dimensional discrete and determinist "Life worlds" all transitions are based on the "physical level" where appearance and reality never come apart. We can see the higher-level patterns, but the "life physics" cannot, and being external to the "Life world" our recognition is without consequence. In our lived world of epistemic agents and vending machines Dennett accepts the causal power of discrimination based on epistemic limitations. We are the loop. Dennett also insightfully proposes that whereas some interactions amplify variation, digitizing can be seen as a way of absorbing micro-variation. He writes:

Surely, the result of a coin flip is the *deterministic* outcome of the total sum of forces acting on the coin...but this total sum has no predictable patterns in it. That is the point of a randomizing device like a coin flip, to make the result uncontrollable by making it sensitive to so many variables that no feasible, finite list of conditions can be singled out as the cause...*It accomplishes just the opposite of digitizing in computers: Instead of absorbing all the micro-variation in the universe, it amplifies it...* (Dennett 2003, 85 [italics mine])

This is an important point in regards to our metaphysical appearance and reality questions. Along these lines Conway's Game of life, and other digitized representations, like the two-bitser can be seen as owing their discrete all-or-none outcomes to condensing variation absorbing processes. In the case of Dennett's use of the Game of life as a metaphysical illustration, we are asked to disregard the physical computer, pretend there is no backstage, and limit our view to the "front stage" of the life world. We do

this by taking the deterministic laws, digitized “matter” and finite and discrete grids as “uncaused causes.” But given Dennett’s own analyses of relational categorization and pattern recognition would one not expect these digitized “Ex nihilo—in nihilo” creations and annihilations to precisely depend on a perspectival view? In other words, from the perspective of the two-bitser the quarter is either present or not. From the “grid-eye-perspective” of the “Life world” a pixel is either on or off. But like Dennett writes here such “unfudging” is generally the product of interactions absorbing variations. From a different perspective one might be able to discriminate two coins, which the two-bitser’s recognition systems cannot, and thereby amplify consequences—e.g. by arresting people and so on. The point I am making is that each of Dennett’s models gain their discreteness via the introduction of an ultimate “uncategorized categorizer,” be it the God’s eye point of view or the “grids” of ideal physics, or a giant immaterial two-bitser, in brief, the supreme perspective, beyond which no other discriminations or reifications can be made. Further, in determinist models like the game of life, causality is then restricted to this level of ultimate reification. Remember we can perceive, enjoy and name the patterns, but only influence them as “Deus ex Machina” by starting a new world. Thus, by way of an algorithmic Life physics-based occasionalism the world is recreated at each instant given the prior reification.

I find it surprising that Dennett, on the one hand, notes that unfudging categorization is the *product*, the *output* of a variation absorbing process, and yet on the other assume an ultimate miraculously unfudged reality—be it determinist or indeterminist—in which “back stages” are per definition excluded as impossible. In other words, the question is why he does not pursue the possibility that a supreme causally-*efficacious* “uncategorized categorizer” might be impossible, not simply in practice due to our limited perspective, but due to the fact that reification needs a limiting perspective. It might be that Dennett is so focused on showing that compatibilism is *possible*, that life and freedom could conceivably have evolved from a determinist universe without “back stage,” that he fails to consider whether this is a *plausible* story? At least this is what we see in his discussion of the “Game of life.” He argues, based on Conway’s proof, that one could embed a Universal Turing Machine in the life world. And therefore that a sort of self-replicating life *could* have evolved in a gigantic pixel space. He does worry about how to simulate variability, noise and mutation—i.e. re-create the fudge—in such a space: “Can a two dimensional world be noisy enough to support open-ended evolution, while still quiet enough to permit the designer parts to do their work unassailed? Nobody knows” (Dennett 2003, 50). He is right, we don’t know. But we also don’t know if God created us—and the fossil record. Some things might be possible but too unlikely to

spend time on, especially in the face of an alternative, that is. As it stands, I take it as a live option that there are no unmoved movers. We shall now look a bit at what it might mean if that were the case and we perhaps had interaction, hiding and revealing, and absorption and amplification all the way down.

4. Alternative views from within—emergence & interaction.

As mentioned Dennett suggests that his eternalist argument against possibilities works—ironically, for all possible worlds—and thus, whether the world is deterministic or indeterminist (Dennett 2003). In this way, he follows the free will debate consensus and takes classic determinism and indeterminism as the only metaphysical frameworks in town. Like Hume famously wrote: “As objects must either be conjoined or not, and as the mind must either be determined or not to pass from one object to another, it is impossible to admit of any medium betwixt chance and an absolute necessity” (Hume 1750). But is this true? Are necessity and chance, determinism and indeterminism our only metaphysical options? If we let go of the metaphysical stance assumption that the actual world has an ultimately reified causal level, then it seems that the possibility space would be open for other metaphysical options. One being that causality somehow takes place from within and exploits the gaps between appearance and reality. I shall in this article not attempt to develop such an alternative framework, but simply advocate for its possibility and briefly alert to the fact that I am not the first one objecting to the determinism-indeterminism ultimatum. Within the traditions of process philosophy, pragmatism, emergence and interactivist approaches many have pointed to metaphysical conceptions that does not fit the traditional eternalist metaphysical picture.¹⁹ In a recent article Mark Bickhard defends the empirical plausibility of a process based interactivist and emergentist view. He writes:

Process, in fact, is now the dominant language of science. Every science has progressed beyond an initial conception of its phenomena in substance terms to understanding that they are in fact process phenomena. Fire is no longer modeled in terms of the substance phlogiston, but instead in terms of the process of combustion; heat no longer in terms of caloric, but in terms of random kinetic processes;

19. See Vintiadis (2013) and Seibt (2012) for helpful recent overviews of respectively the emergentist tradition and process philosophy, but see also classic philosophy texts like e.g. Bergson (1896) and Whitehead (1929/1978), and of ecologists and theoretical biologists like e.g. Bateson (1979) and Rosen (1991).

life no longer in terms of vital fluids, but in terms of special kinds of far from thermodynamic equilibrium processes. And so on. Every science, that is, with the exception of the sciences and philosophies of mind and persons. Here substance and structural views are still dominant. (Bickhard 2009, 553)

Thus, Bickhard points out how within nearly all sciences process-based and relational phenomena are taking center stage.²⁰ Interestingly, for our purposes here he is well aware that many metaphysicians have a hard time letting go of atomistic, determinist and eternalist assumptions and getting their mind around the notion of a process metaphysics.

The shift to a process metaphysics, however, induces major changes in our overall framework of assumptions: First, change becomes the explanatory default, and it is stability that requires explanation. Similarly, processes, unlike atoms or the “stuff” of substances, do not have inherent boundaries, and boundaries too, therefore, must be explained, not assumed. Second, processes have their causal powers in virtue of their organization. Organization cannot be delegitimated as a possible locus of causal power without eliminating all causality from the universe. But, if organization is a potential locus of causal power, then so is higher level organization. In particular, there is no metaphysical block to the possibility of emergent causal power in new organization. And third, if emergence is a metaphysical possibility, then the door is open to the possibility that normativity and mind are emergent. (Bickhard 2009, 553–4)

In other words, to use another of Dennett’s favorite expressions—we need to perform a “strange inversion of reason” to understand process metaphysics (Dennett 2003, 47). Of course Dennett takes the Darwinian inversion of reason to be a bottom-up compositional story, as opposed to a top-down story of intelligent design. What I, via Bickhard, propose is instead an inversion of reason away from reductionist foundationalism and “unexplained explainers” (be they large or small, physical or immaterial) to a view from within—all the way out so to speak.

20. See e.g. Witzany (2014) for recent insights in the “pragmatic turn in biology.”

Such accounts can and have been developed in different ways, typically given differing critical starting points. Physicist Lee Smolin and philosopher Roberto Unger have in recent works (Smolin 2013, Unger & Smolin 2014) challenged assumptions about a-historical pre-given deterministic laws existing externally to the fabric of actual processes. If rather laws are inherent to and evolving with the concrete physical processes and conditions, then we get a radically different “view from eternity” than the one Laplace and Dennett avail themselves off. Smolin (2013) also points to several much earlier proponents of the ‘laws evolve’ view. Dirac writes in 1939: “At the beginning of time the laws of Nature were probably very different from what they are now. Thus, we should consider the laws of Nature as continually changing with the epoch, instead of as holding uniformly throughout space-time.” Given Dennett’s view of freedom as evolving, why would he not also entertain the idea that laws and “transition rules” might evolve as well? Why the externalist view of a world that comes pre-categorized? This is precisely the aspect of universal laws that Charles Sanders Peirce found the most problematic all the way back in 1891:

To suppose universal laws of nature capable of being apprehended by the mind and yet having no reason for their special forms, but standing inexplicable and irrational, is hardly a justifiable position. Uniformities are precisely the sort of facts that need to be accounted for...Now the only way to account for the laws of nature and for uniformity in general is to suppose them results of evolution. (Peirce 1891)

This point could not be more pertinent to the issue at hand, and the metaphysical assumption of “unexplained explainers” that each of the formulations of determinism that we have met contains. Like Bickhard wrote above, given a process view it is stability and borders, not change, that need to be explained, Peirce similarly suggests that it is uniformities and laws that call for an explanation. Further, Nancy Cartwright has from an internal stand point of how actual science unfolds, advocated for an alternative metaphysical view, which she calls “the dappled world” (Cartwright 1999). The idea is here not only that our current laws of nature “lie,” but that the world might simply be such that not even from the metaphysical stance are there universal and non-local regularities to be found. She has also worked on an Aristotelian inspired notion of “capacities” that could be congenial to an emergentist or “internalist” account of causation and thus an evolutionary view of action (Cartwright 1994).²¹ Thus, even though as Bickhard writes

21. Cartwright is one the originators of the so-called Stanford school within the philosophy of science, which in

above the field of philosophy of mind seems to have stubbornly insisted on outdated and implausible metaphysical assumptions, alternatives do seem to exist—even if they are not neatly reified.

5. Conclusion: Evolving Dennett's story beyond determinism

Given Dennett's own analyses of the causal efficacy of epistemic limitations, hidden complexities and two-bitser inscrutability, his philosophy seems to show us beyond the external metaphysical stance. If there is no layer of reality without a back stage, where appearance always equal reality, then we might suggest that all causal interactions "amplify" certain variations and "absorb" others. In other words, Dennett's account is in many ways congenial to an internalist—or perspectival, interactivist and emergentist—inversion, where *all* causal effects might rely on some recognition and response. If the world—like the front stage of Conway's game of life—were such that all reality appears at each instant, i.e. reality and appearance never come apart, then this "internalist" idea of causality would make no difference. But imagine a world—perhaps like ours—with mind-boggling multiplicities of spatio-temporal timescales and ways of hiding and revealing. What if no fundamental level is to be found—but rather each level, each interaction has its own "dark matter" and even the "Higgs field" is not an "uncaused cause." Maybe in such a world there are no neat "states" and "instants," where everything can reveal all its aspects or properties. In such a world there would always be some gaps between appearance and reality, between what is of consequence in current interactions and what "it there." Such a temporal world of situated causalities I think would be more conducive to evolution, emergence, change and the odd combination of noise and quiet Dennett is looking to program into Conway's game of life.

If Dennett's compatibilism and his evolutionary account of free will are revisited with such considerations in mind then it looks as if parts of the possibility space have been ignored. Maybe a biological and historical view like Dennett's allows us to invent a new free will position that does not rely on traditional libertarian routes of claiming extra-physical determining forces (dualist) or non-determined (random) effects. The question is whether a dynamic pluralistic world always getting its causality from actual interactions, with limited relational recognition, can provide the grounds we need to re-interpret how real-time action choice and determination might be possible at the human level. More specifically, whether the causal consequences of teleological perception,

many ways is united not simply by geography but by its challenge to eternalist, essentialist and reductionist views. See e.g. Dupré (1995).

memory, anticipation and interaction feedback, can give rise to forms of emergence that we might think of as having downward or—as I prefer—outward efficacy.²²

Some might want to argue that this proposal is not actually that different from Dennett's, as he might be read not necessarily as a classic determinist but as a metaphysical agnostic. This might be so, but then this article is making a significant contribution to the clarification of Dennett's position and its commitments and explanatory merits. His texts are, as we have seen, ripe with reductionist and eternalist models, which indicates that he at least in practice is ignoring other metaphysical possibilities such as emergentist, interactivist and process positions.

I propose that the possibility of such positions matter—also to our conception of our own actions. Instead of saying, as Dennett does, that we have perceived opportunities because we are epistemologically limited, we can now say that our epistemological limitations are *as casually efficacious as anything else can claim to be*. The point is that in a world that cannot be reduced to a level of ultimate and inherently transparent compositional “parts,” every interaction might be only partially revealing. The world of epistemology and metaphysics, of appearance and reality, phenomena and noumena, now interacts as no processes of reification are externalized from the causal web. This might allow for a changed perspective on how our action choices matter. The question is if our evolved abilities to predict, and flexibly integrate horizontal cultural influences, and inhibit and hide impulses, are the sources of emergent effects in the present, i.e. shape the causal fabric of the world in ways that are as determining as they are determined. The proposal is thus not that our actions are “indetermined” in any classic sense but rather that they both—to use Dennett's terminology—*amplify and absorb variability*. We can actively and purposively *hide* and *reveal* causally efficacious aspects of “ourselves” and the world.²³ Dennett is right to highlight our complex and bastardized genealogies, as this heterogeneity and multiplicity of the self holds “it's” generative power. It is not just that we “are the loop” as Dennett says, but further that we constantly dynamically change and choose what pertains to *our* causally efficacious loops. Dennett aims to show that freedom can evolve from entirely unfree atomistic parts, and accordingly likes to describe

22. An important aspect to explore of such a hypothesis would be whether in a world without a bottom layer, one might expect both order and disorder—regularity and variability—to be constantly be produced as irreducible actual concrete interactions would determine the outcomes in real time. This does not seem to be the case in Conway's game of life—here the finite determinist worlds move towards more order.

23. Note here the similarity to Merleau-Ponty's (1964) the insistence on the role of invisibility in visibility as well as his critique (1965) of what he sees as Bergson's (1896) purely positive process philosophy.

us in mechanistic terms as composed of “micro-robots” and “micro-factories.” However, we might—perhaps like the bacterium—in part owe our freedom to not being the sum of a fixed set of finite parts, but rather constantly self-creating open systems. Thus, autonomy and emergent novelty might be possible.

I want to reiterate that I do not claim to have a theory of free will. My aim in this present article is merely to reveal theoretical possibilities that have been hidden by the insistence on fully transparent discrete and eternalist metaphysical perspectives. If there is no floor to the universe, then there might always be a gap between appearance and reality. This would undermine all classic formulations of determinism it seems. However, I tried to show that the potential absence of a metaphysical “back stage,” suggest that our action choices might indeed happen in real time and determine what is to come, possibly by way of new emergent effects. Thus, the elbow room of human agency comes not only from the fact that our procedural selves are “smeared over time and space” but also from the constant games of peek-a-boo where our worlds and we—in all our bastardized multiplicities—are never fully hidden nor fully transparent or expressed.²⁴ These ideas might allow for the development of an ignored kind of free will worth wanting, and I shall be curious to see if some indeterminists would agree. After all, as Dennett is, I am puzzled by the desire to have one’s actions to be without any precursors. And, I wonder whether many do not simply desire—as James (1897) famously expressed it—“that the issue is decided nowhere else than here and now. That is what gives the palpating reality to our moral life and makes it tingle as Mr Mallock says, with so strange and elaborate excitement.”

24. Bergson - and other who have made similar claims - might thus exactly be wrong to propose that the free action was an expression of the entire ‘self’. As he his self later pointed out; all actions simplify. (Bergson 1889)

References

- Akins, Kathleen. 1996. "Of sensory systems and the "aboutness" of mental states." *The Journal of Philosophy* 93 (7): 337–372.
- Bateson, Gregory. 1979. *Mind and Nature—a necessary unit*. New York: Bantam New Age Books.
- Bergson, Henri. 1889. *Essai sur les données immédiates de la conscience*. Paris: F. Alcan.
- Bergson, Henri. 1896. *Matière et mémoire*. Paris: F. Alcan
- Bickhard, Mark. 2009. "The Interactivist model." *Synthese* 166 (3): 547–591.
- Brincker, Maria. 2014. "Navigating beyond "here & now" affordances—on sensorimotor maturation and "false belief" performance." *Frontiers in Psychology* 5 (1433). doi: 10.3389/fpsyg.2014.01433.
- Cartwright, Nancy. 1994. *Nature's Capacities and their Measurement*. Oxford: Oxford University Press.
- Cartwright, Nancy. 1999. *The dappled world: A study of the boundaries of science*. Cambridge: Cambridge University Press.
- Crome, Victor J. 2007. *Zeno's Paradoxes and the passage of time*. Doctoral dissertation. New York: City University of New York.
- Dennett, Daniel C. 1984. *Elbow room: The varieties of free will worth wanting*. MIT Press.
- Dennett, Daniel C. 1987. *The intentional stance*. Cambridge: MIT press.
- Dennett, Daniel C. 2003. *Freedom evolves*. New York: Viking Penguin.
- Dirac, Paul. 1939. "The relation between mathematics and physics." In *Proc. Roy. Soc. Edinburgh* 59 (II): 122.
- Dupré, John. 1995. *The disorder of things: Metaphysical foundations of the disunity of science*. Cambridge: Harvard University Press.
- Gallagher, Shaun. 2006. "Where's the action? Epiphenomenalism and the problem of free will." In *Does Consciousness Cause Behavior? An Investigation of the Nature of Volition*, edited by W. Banks, S. Pockett, and S. Gallagher, 109–124. Cambridge: MIT Press.
- Gardner, Martin. 1970. "Mathematical Games—The fantastic combinations of John Conway's new solitaire game "life."" *Scientific American* 223 (October 1970): 120–123.

- Grossman, Lisa, and Richard Webb. 2013. "Martian chemistry was friendlier to life." *New Scientist* 219 (2933): 14.
- Hornsby, Jennifer. 1997. *Simple Mindedness—In defense of naïve naturalism in the philosophy of mind*. Cambridge: Harvard University Press.
- Hume, David. 1750. *An enquiry concerning human understanding*.
- Inwagen, Peter van. 1983. *An essay on free will*. New York: Oxford University Press.
- Laplace, Pierre-Simon. 1951. *A Philosophical Essay on Probabilities*. Translated by F.W. Truscott and F.L. Emory. New York: Dover Publications.
- Leibniz, G. W. (1893) 1989. "The monadology." In *Philosophical Papers and Letters*, edited by Leroy Loemker, 643–653. New York: Springer.
- Merleau-Ponty, Maurice. 1964. *Le Visible et l'Invisible*. Paris: Gallimard.
- Merleau-Ponty, Maurice. 1965. *Éloge de la philosophie*. Paris: Gallimard.
- Millikan, Ruth Garrett. 1989. "In defense of proper functions." *Philosophy of Science* 56 (2): 288–302.
- Peirce, Charles Sanders. 1891. "The architecture of theories." *The Monist* 1 (2): 161–176.
- Putnam, Hilary. 1981. *Reason, truth and history*. Vol. 3. Cambridge: Cambridge University Press.
- Quine, W.V.O. 1960. *Word and Object*. Cambridge: MIT Press.
- Rosen, Robert. 1991. *Life Itself—a comprehensive inquiry into the nature, origin and fabrication of life*. New York: Columbia University Press.
- Seibt, Johanna. 2012. "Process Philosophy." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta.
- Smolin, Lee. 2013. *Time reborn: From the crisis in physics to the future of the universe*. Boston: Houghton Mifflin Harcourt.
- Unger, Roberto Mangabeira, and Lee Smolin. 2014. *The singularity of the universe and the reality of time*. Cambridge: Cambridge University Press.
- Vintiadis, Elly. 2013. Emergence. *Internet Encyclopedia of Philosophy*.
- Whitehead, Alfred North. (1929) 1978. *Process and reality*. Corrected eds. David Ray Griffin and Donald Sherburne. New York: Free Press.
- Witzany, Guenther. 2014. "Pragmatic turn in biology: From biological molecules to genetic content operators." *World Journal of Biological Chemistry* 5 (3): 279–285.