

## The Emergence of Free, Intentional Control: Reply to Haselager

Timothy O'Connor

### ABSTRACT

Pim Haselager argues that contemporary science and philosophical reflection together give us reason to revise our commonsense experience and conception of human action as an exercise of autonomous, causally undetermined control. The present chapter rebuts that argument. It contends that our commonsense conception of action is internally coherent and can be comfortably integrated with the sciences of human action to date. The currently fragmentary findings of psychology and neuroscience reveal greater *complexity* to the processes constituting human action than the broad template that introspection indicates, but they can be smoothly mapped onto it. Given that this is so, the default status of commonsense philosophy implies that we remain justified in our pre-theoretical understanding of ourselves as free and morally responsible agents.

### KEYWORDS

free will, intention, consciousness, emergence, reduction, indeterminism, probabilistic causation, chance

### BIO

Timothy O'Connor is Professor of Philosophy at Indiana University, Bloomington and a member of its Cognitive Sciences Program. He is the author of *Persons and Causes: The Metaphysics of Free Will* and has edited six books and dozens of scholarly articles on the topics of free will,

consciousness, and the emergence-reductionism debate. He is currently part of a four-year, sixteen-person project of neuroscientists and philosophers who are collaboratively conducting a series of studies pertaining to the neuroscience of free will.

Pim Haselager explores the question of how we should think about free will and more generally human intention-in-action in light of tentative and slowly developing frameworks in neuroscience and embodied cognition. He starts with a note of sympathy for Reidian 'common sense' epistemology, which contends that we are rationally entitled without independent evidence or argument to certain general beliefs about the world and the cognitive tools that we bring to our encounter with it. We have to start somewhere in coming to learn things about the world and ourselves. The reasonable starting points are, roughly, those that we find ourselves accepting instinctively and that appear to be essential to our ability to engage in critical inquiry and action. Among these instinctive beliefs is our self-conception as purposive agents: confronted by more than one option that we have reason to pursue, and being uncertain about what to do, we deliberate and choose. Choice is a kind of uncertainty-resolving intention to act that – so it seems – initiates and (in complex cases) guides the completion of the action. It seems to us, then, that we are *effective, purpose-driven choosers*. But further, it seems that our choosing is *free*: having chosen one option, it was open to us to have chosen differently. Nothing necessitated that particular option, and partly for that reason, we are responsible for having chosen as we did.

Haselager allows that this description captures our instinctive conception and experience of deliberate agency, and that it is *prima facie* reasonable to accept it. However, he adds that

epistemic starting points may be overturned by empirical evidence, and he quickly proceeds to suggest that our commonsense conception of agency is in need of significant revision. Why, exactly?

It is not that our well-confirmed sciences pertaining to the human mind demonstrate its falsity. Haselager (correctly) acknowledges that they are too immature to have done that. Instead, he thinks, there is a broad metaphysics of nature that is more promising than any alternative. This framework is 'scientific materialism,' which sees all of nature (human action included) as constituted by 'physical-law abiding interactions of material elements' (3). High-level organized phenomena are one and all determined from the bottom up, even though scientific practice necessarily starts from our macroscopic perceptual level and tries to burrow down. Furthermore, in thinking about human choice and action, Haselager has 'not seen' any explanation of how macroscopically significant indeterminism might be anything other than mere 'chance' (3), and so he struggles to understand "what 'could have done otherwise' might mean, or why it is considered to be so valuable" (9). So in theorizing about human freedom and moral responsibility, we should discard this naive but possibly incoherent notion of an 'open future' that we partly determine and instead focus on the undoubted *flexibility* of our responses to stimuli over time – on the fact that we learn from experience and often adapt our responses to similar future circumstances.

It should be mildly puzzling that we are naturally disposed to think of our own agency in *profoundly* confused ways, given its intimacy to us and its importance to both our moral outlook and our effective engagement with reality. Indeed, I do not think that we are so confused. Causal

indeterminism is a merely negative condition on our freedom. It is the precondition on something positive: the exercise of a two- (or multi-) way power of choice that is guided but not determined by reasons we have (from our subjective perspective, however misguided) to act in particular ways. We humans are psychologically complex creatures, and we often enough have competing motivations that point towards different courses of action. These motivations are causal as well as rational influences, and so they make us more or less likely to choose one way rather than the other. Even so, we must choose. In choosing one way (and in experiencing the consequences), we influence incrementally the probabilities pertaining to future choices; over time, we typically come to have fairly settled dispositions concerning certain types of goals and outcomes.

I take the previous paragraph to be consistent with a reflective common sense perspective. Contra Haselager, I think it is readily understandable from a broadly scientific perspective, although it requires abandoning the ontologically reductionist straitjacket that he builds into 'scientific materialism' and seeing greater complexity in the relationship of the differently-leveled structures studied by our successful sciences. First, it should be noted that probabilistic causal influence is now a firmly-established part of science. Sometimes, of course, this merely reflects relative frequencies of common variables over ensembles of processes that are similar in coarse-grained ways but whose fine-grained conditions on the underlying driving mechanisms differ. But probabilities may also reflect basic stochastic propensities – individual-case, weighted tendencies of causal agents towards possible outcomes. Certain quantum phenomena such as radioactive decay are naturally understood in this way. Whatever the facts turn out to be, we can certainly *conceive* choice in a similar way. Haselager says that, if actual, this would be mere chance, and so of no help in understanding how we might control our actions in a way that

grounds moral responsibility. He is not alone in this assessment: it represents an amply-defended perspective on 'the philosophical problem of free will.' However, we have to ask what 'chance' means, something that he (and his philosophical supporters) nowhere define. I suggest that our ordinary notion of a chance occurrence is of one that is causally undetermined *and not intentionally controlled*. By that definition, an agent's motivated forming of an intention to act, even if not determined by antecedent causes, is not something that comes about by chance. Perhaps Haselager is supposing that chance occurrence may be defined more simply as anything that is causally undetermined, for the reason that control over an outcome is necessarily exercised *antecedently*. So any causal conditions that fail to determine an outcome thereby fail to (fully) control it. But choice is not a mere outcome, it is an *exercise* of control. It is something we do, not something that happens to us. Or, again, so it naturally seems. So when Haselager asks, prior to the consideration of scientific evidence, what could have done otherwise 'means,' and why it might be valuable, the common sense answer is quite simple: it means that we exercised a power to settle which course of action we would follow among a plurality of alternatives, and it is valuable because it constitutes a form of limited autonomy. We are not merely complex and changing conduits of causal influences from the past, flexible decision-makers *over* time but always pre-determined in how we decide on any occasion. We are fundamental difference-makers – in the necessarily constrained manner of conditioned creatures rather than in godlike, *ex nihilo* fashion – yet still freely and creatively responding *at* a time. Such undetermined activity would not be inexplicable. There need only be antecedent influences (in the form of goals or desires) that motivated the course of action taken rather than the contemplated alternatives without their having precluded those otherwise-motivated alternatives.

The 'logic' of explanation of undetermined outcomes generally has been fruitfully developed by philosophers of science in the advent of indeterministic theories in physics and elsewhere.

Let us turn from Haselager's conceptual misgivings to the more empirical question of whether the totality of present evidence from a range of sciences supports a reductionist conception of physical reality as wholly fixed from below by basic and impersonal physical forces. On this vision, all the structures and activity of chemistry, biology, neuroscience and psychology, sociology, and ecology are but the outworkings of a 21st century version of Democritus's atoms in the void. The descriptive concepts and principles of these restricted-domain sciences are not derivable from physical theory, Haselager allows, but the structural and dynamical *realities* they describe are wholly derivative: they are as they are in all respects because the basic physical realities are as *they* are. Physics is a complete and autonomous domain, determining but not determined by the structured realities. There is much to say about this reductionist vision, but here I can only comment briefly. This vision dominated scientific thinking in the first part of the 20th century, between the breakthroughs of quantum mechanical explanations for chemical structure and the uncovering of the biochemical basis of life. However, subsequent developments strongly challenged it, as the power of complementary, top-down powers of organized structures to constrain and channel the behavior of their elements became clear in a variety of domains. Robert Laughlin (solid-state physics), Denis Noble (systems biology), and George Ellis (cosmology) are among the prominent voices articulating an alternative vision, on which complex systems of many kinds exhibit reciprocal and inter-dependent bottom-up and top-down

behavior in their evolving trajectories over time.<sup>1</sup> While neither vision is conclusively demonstrable, and (given measurement and computational constraints) likely never will be, the 'emergentist' vision is arguably a more natural fit with our evidential reality: profound successes in fundamental physics in describing artificially isolated interactions of small numbers of particles; equal success in delineating and confirming top-down explanations in many complex domains; and little prospect of ever being able to directly apply quantum theory to the entirety of large-scale composed systems involving billions or trillions of interacting particles. Certainly the emergentist vision is a more natural fit with available evidence if one gives any significant credence to our commonsense perspective. Our encounter with the world begins with our conscious experiences, thoughts, emotions, and other intentional states. All of these present themselves to us as having distinctive intrinsic character that is quite unlike the character of our non-conscious environment, and as continuously impacting our behavior and other, subsequent conscious states. To take them at face value is to implicitly endorse an ontologically emergentist understanding of their place alongside the subconscious neural machinery that sustains and modifies them in turn. To endorse, alternatively, a reductionist thesis regarding conscious states is to implicitly endorse a kind of illusionism, according to which our pervasive impression of directly apprehending the intrinsic character of conscious states is simply an illusion for future cognitive science to explain. (On such a view, conscious states are neural, which fact is not given to us in conscious experience. Our awareness of such states and ability reliably to compare them

---

<sup>1</sup> See Laughlin (2005), Noble (2006), Ellis (2016), as well as the multi-authored volume edited by Murphy, Ellis, & O'Connor (2009) and the special journal issue edited by Ellis, Noble, and O'Connor (2012). For systematic analysis of the concept of emergence and its possible scientific applications, see O'Connor and Wilson (2020).

with previous such states is wholly constituted by subconscious mechanisms, as is the pervasive illusion of our grasping a distinctive qualitative and subjective character.)

A serious problem for the scientist who would endorse the reductionist's rejection of our commonsense perspective on conscious experience and agency is that this perspective is quite fundamental to the epistemology of science itself. Scientific theories, models, and results are themselves the products of scientific *activity*: of human persons acting in certain coordinated, purposive ways and communicating their activities and results to one another. The nature of this activity is not part of the content of such theories, as the theories are not (in most cases) about such activity. However, we who embrace the theories and conclusions drawn in any scientific domain must tacitly accept that the theorists and experimenters were guided by purposes known to themselves and successfully carried out those purposes by knowingly intending a variety of specific behaviors thought to advance those purposes. Absent reasonable belief in their reliably self-aware agency, our confidence in what took place and what it resulted in would be undermined: we would have words on a page or a screen, but we would not be in a position to assess what they were indicative of. Or, this would be our predicament absent some alternative account that insured the reliability of the products, despite the unwitting behavior of the scientists yielding them. No such alternative account readily suggests itself, and any candidate that might be presented for our consideration would require us to trust the one providing it, raising the same problem all over again. Haselager seems dimly aware of the self-undermining problem here when he likens the revisionist perspective that he envisions cognitive neuroscience as possibly producing "a snake biting its own tail" (2). But it would more accurately be characterized as the paradox of a snake *swallowing* its own tail. Or, to switch the metaphor, a



scientist who comes to reject our fundamental self-understanding as agents who are regularly consciously aware of efficaciously intending actions is sawing off the branch on which he sits.<sup>2</sup>

Now, affirming that we form conscious intentions that are efficacious in initiating and guiding action is not to affirm that we ever do so *freely*. But the sciences of brain and behavior do not (yet) pose serious challenges to our freedom. At any rate, no such challenge endures from the one set of studies that Haselager discusses, viz, those pioneered by Benjamin Libet and refined by many others on the timing of conscious intention or will. Haselager appears to agree, up to a point. However, it is unclear how widely he would apply the conjecture of Schurger et al. (2012) that the timing of action-initiation may result from the passing of a threshold in *random* neural activity. This conjecture has some intuitive as well as experimental plausibility in Libet-style experimental scenarios where agents are asked to 'spontaneously' make arbitrary decisions, but no good basis has been given to generalize this to the seemingly very different, non-arbitrary and deliberate choices that are better exemplars of our commonsense understanding of free will, as when I must choose between enjoying a leisurely afternoon in the park and getting some pressing work done, or between defending an unfairly maligned individual and maintaining good relations with persons in a position to advance my career.

Haselager closes his discussion by reciting in compact form the familiar scientific 'creation myth' that is the credo of many would-be deflators of the human pretension to which pre-scientific people are supposedly naturally inclined:

---

<sup>2</sup> For a more elaborate development of an argument for a closely-related conclusion, see Richard Swinburne, *Mind, Brain, and Free Will*. Oxford: Oxford University Press, 2013, 117-120.

The scientific picture, incomplete as it may be, indicates that we are bipedal great apes, possessing linguistic skills often used for confabulating societally acceptable reasons for actions that are generally driven by subconscious processes, on a small planet circling around an insignificant sun in a remote area of the universe. The revisions of intentions and free will examined in this chapter may undermine other aspects of a cherished human self-perspective.

(9)

What is true within this myth is that an important aspect of the advance of scientific understanding, especially in physics, has come from learning to prescind our view of the cosmos from our *perspectival* limitations with respect to time, space, and history. However, the resulting advance in knowledge concerning the cosmos and its (and our) history does not overturn our basic self-understanding or speak one way or another concerning our fundamental significance. An alternative framing of the story of the appearance of human beings on this planet that is at least as consonant with the established facts and well-confirmed theories in 2020 as Haselager's is this:

Sundry basic features at the origin of our cosmos were exquisitely fine-tuned for the building blocks necessary for the subsequent appearance of life on our planet and its slow evolution towards greater complexity. Along one branch of that fecund biological bush there emerged a variety of complex creatures unified by the striking new feature of conscious awareness, and culminating in ourselves: purposive, flexibly intelligent, self- and other-aware creatures. Our nature contains strata of our many precursors, such that we are curious hybrids, governed by a patchwork of mammalian machinery and drives interacting with our distinctively personal capacities and attitudes. Curious but also wondrous: planet-bound, bipedal great apes by

heritage yet, astonishingly, also capable of contemplating, probing, and coming to understand the entirety of the vast cosmos whose elemental traces our bodies bear – and capable, too, of contemplating and debating the possible purposes that may govern our remarkable species.<sup>3</sup>

## References

- Ellis, G.F.R. (2016). *How Can Physics Underlie the Mind?: Top-Down Causation in the Human Context*. Berlin: Springer-Verlag.
- Ellis, G.F.R, Noble, D., & O'Connor, T., eds. (2012). *Top-down Causation: An Integrating Theme Within and Across the Sciences?* Themed issue of Interface Focus: A Journal of the Royal Society 2 (Feb), 2012, 1-140.
- Laughlin, R. (2005). *A Different Universe: Reinventing Physics from the Bottom Down*. New York: Basic Books.
- Murphy, N., Ellis, G., & O'Connor, T., eds. (2009). *Downward Causation and the Neurobiology of Free Will*. New York: Springer Publications, 2009.
- Noble, D. (2006). *The Music of Life*. Oxford: Oxford University Press.
- O'Connor, T. "Emergent Properties," *Stanford Encyclopedia of Philosophy*, forthcoming.
- Schurger, A., Sitt, J. D., & Dehaene, S. (2012). "An Accumulator Model for Spontaneous Neural Activity Prior to Self-initiated Movement," *Proceedings of the National Academy of Sciences* 109(42), E2904–13. doi:10.1073/pnas.1210467109.

---

<sup>3</sup> This publication was made possible through the support of a joint grant from the John Templeton Foundation and the Fetzer Institute. The opinions expressed in this publication are those of the author and do not necessarily reflect the views of the John Templeton Foundation or the Fetzer Institute.