

JAIL BREAK: TALLIS AND THE PRISON OF NATURE

THOMAS W. CLARK

Abstract: In *Freedom: An Impossible Reality*, Ray Tallis argues that we escape imprisonment by causal determinism, and thus gain free will, by the virtual distance from natural laws afforded us by intentionality, a human capacity that he claims cannot be naturalized. I respond that we can't know in advance that intentionality will never be subsumed by science, and that our capacities to entertain possibilities and decide among them are natural cognitive endowments that supervene on generally reliable neural processes. Moreover, any disconnection from the multi-level determinants that account for human behavior cannot augment, but would likely undermine, effective human agency. Our full inclusion in nature, understood in terms of a pragmatic, explanatory determinism, is therefore not a prison from which we need to escape.

Key words: free will; determinism; intentionality; materialism; naturalism; rationality.

Introduction

In *Freedom: An Impossible Reality*, Ray Tallis seeks to place humankind outside nature to free us from subjection to natural laws. Why should we want such freedom? Perhaps to be truly originative agents that can be credited and blamed in a way not possible were we and our actions fully traceable to non-self conditions. Although Tallis counts himself a compatibilist and acknowledges we are by no means exempt from causation, determinism in the natural world is nevertheless the arch enemy of free will in this book. To avoid what he regards as imprisonment by determinism, we have to stand apart from nature in some respect, so the compatibility involves separate but co-existing domains of the natural (law-like materiality) and unnatural (free human mentality).¹

According to Tallis, this divide is secured by the virtual space of our intentionality. By inhabiting the conscious mental realm of considered possibilities, we escape the mere actuality of our bodies and so can originate action in a way not available to lesser creatures, slaves as they are to physical determinism; *we* are not “mere conduits” of causation (Tallis, 2021, p. 107). In short, free will on Tallis's account depends on the claim that the mental realm of conscious intentionality cannot be naturalized – cannot be brought

¹ See Tartaglia, 2020, the chapter “Freedom” (119–135) for an idealist conception of human freedom equally antithetical to materialism and determinism.

within the explanatory ambit of science. The very formulation and application of science is, he suggests, unnatural since we must stand apart from nature to do it, and science as a description of natural laws cannot account for intentionality or consciousness.² There cannot be a science of conscious intentionality, which therefore remains a mystery (section 4 below).

This is a very strong claim indeed about what no one doubts: that we are preeminently reasoning creatures with abilities to imagine possibilities, choose among them, and act to realize them. The conscious entertaining of concepts, propositions, and counterfactuals, all presented in terms of the qualities of phenomenal experience (qualia), obviously depends on the physical brain in ways yet to be understood, but Tallis claims this dependence somehow manages to evade the possibly deterministic constraints of its material basis. Moreover, natural material causation itself, Tallis argues, is not fully necessitating in the way determinists sometimes suppose – there is causal slack in nature (section 2). The upshot: our behavior is not fully traceable to our physical status, but has explanations that place us outside a naturalistic determinism, the practical import of which is that we could have done otherwise in an actual situation just as it transpired up to the moment of choice (section 5).³ This endows human agency with a kind of contra-causal origination not available to what’s fully traceable within the material realm, and gives Tallis’s compatibilism a strong libertarian flavor, even as he disavows the ultimate free will championed by libertarian philosophers such as Robert Kane (Tallis, 2021, p. 11).

I hold the opposite view: our evident capacity for reasoning, practical and abstract, and the mental phenomena involved in its exercise, do not causally transcend what the brain and body do, situated as they are in a material and social environment. That we routinely and confidently traffic in possibilities and abstractions in concert with our peers to achieve all sorts of far-flung projects is indeed a marvel of mental representation and counterfactual consideration. In simulating possible outcomes it’s true that, as Tallis argues, we operate at *virtual* distance from the world, but we’re in no sense at an *actual* distance from our neural materiality, upon which the conscious simulation obviously depends (Metzinger, 2003). All our imaginings and deliberations take place in actual, neurally embodied situations, one neural event flowing seamlessly into the next (Tallis acknowledges the unobstructed flow of material nature⁴), the brain functioning as the staggeringly complex engine of mentality, situated within an equally complex social milieu. There’s no reason to suppose that human intentionality, brain based as it must be, transcends whatever causal relations hold among its material constituents, even if we don’t yet have, or may never have, the full story relating the material to the mental completely figured out. Moreover, I will argue that we needn’t transcend determinism, understood as involving multiple levels of explanatory determinants (section 3), to secure true human agency.

² Tallis, 2021, pp. 29-50, section titled “The Unnatural Nature of the Science of Nature.”

³ “...what we have done is one of several possibilities genuinely open to us such that we could have done or chosen otherwise.” (Tallis, 2021, p. 12).

⁴ “...the natural world unfolds according to unbroken – and hence seemingly unbreakable – habits that science unpacks as laws.” (Tallis, 2021, pp. xiii-xiv).

Questioning natural causal necessity

Tallis's defense of free will is quintessentially philosophical, analyzing concepts of causation, natural laws, and intentionality to purport that they can't be reconciled. Following Hume and other philosophers, he demotes the causal relations observed in physical nature to mere tendencies ("habits"), not necessities with coercive power over us (chapter 2, "Bringing the laws on side"). It's science, a human description of nature, which traffics in laws, not nature, and Tallis argues it's the existence of science itself – a manifestly representational project – that proves we are not fully subject to causal necessity.

If nature is indeed not strictly law-like, this might give us a kind of causal openness at the physical level of the brain. But it isn't clear what that would buy us when reasoning, or simulating possibilities, or exercising any other specifically human mental capacity. Tallis does not say how the absence of causal necessity in the brain, or anywhere else in nature, would enable us to do better mathematics, or construct better space telescopes, or think better thought experiments. And indeed, causally unnecessitated neural or otherwise physical events, should they exist, can't constitute the free will Tallis is concerned to defend, which is at the *person* level of conscious experience and reasoning. It's intentionality – the representational *aboutness* of our conscious deliberations concerning possible futures – that ostensibly breaks us free from natural physical determinism. I therefore don't think Tallis need worry about the possibility that law-like physical necessity might rule at the neural level or below, with whatever quantum indeterminism might intrude, which indeterminism he agrees wouldn't help with free will.⁵

The multiplicity of explanatory determinants

What has more bite is his claim about the mental control of behavior: that our intentionality can't be reduced to physical causation and therefore escapes material determinism, thus making us true originators of action. Tallis is surely right that causal explanations of the "push-pull" physical variety can't perspicuously account for human behavior at the intentional level. But does our mental prowess put us outside physical nature, as he claims? And by what route? Here we should distinguish between being causally constrained by the neurally-instantiated processes that subserve our reasoning capacities, which we obviously are, versus seeking reductive causal *explanations*, at the neural level, of goal-directed behavior and intentionality. A reasons-based explanation of what I do in terms of beliefs and desires may not be reducible to a physicalist causal explanation in terms of neural-muscular processes, but this doesn't mean that such an explanation necessarily lies outside or transcends the possibly deterministic operation of my brain and body. Rather, it trades in a higher-level explanatory space, what Daniel Dennett (1987) has called the intentional stance, that we apply to ourselves given our rational and simulational powers and our folk theory of mind: attributing to others beliefs, desires, intentions, etc. Dennett claims, and I concur, that intentionalist mental explanations of behavior are perfectly consistent with

⁵ "Quantum indeterminacy would not seem to open a space for us freely to decide to carry out a criminal act or refrain from doing so." (Tallis, 2021, p. 23)

the physical explanatory level, albeit perhaps not in principle *reducible* to it (opinions differ on in-principle reducibility; no one supposes it's possible in practice). We can therefore safely traffic in (at least) two explanatory spaces, one physical, one mental-intentional, that run more or less in parallel, using whichever one best suits our purposes.⁶ We need not suppose intentionality outruns physical determinism to justly and usefully take the intentional stance when explaining ourselves and others. Rather, we can in good conscience talk about reasons, if not as physical causes, as being among the higher-level, intentionality-based *determinants* of our behavior.

The mental content implicated in intentionality, such as reasoning involving concepts and counterfactuals, supervenes on the physical brain, so in order for me to have simulated possibilities otherwise, my brain must have been physically otherwise. An interesting and I think open question is whether there might be regularities at the level of neurally carried representational content that determine, in a quasi top-down manner, how one brain state follows from another, whether or not we talk about this in terms of causation (Tallis objects to construing reasons and propositional attitudes as causes, although others such as Donald Davidson have not.)⁷ If there are, then it's possible that our conscious simulational processing, albeit neurally realized, has what we might call *determinative priority* over lower-level mechanisms: it determines how those mechanisms are recruited from moment to moment to support our cogitations. If so, we wouldn't be able to predict intentionality-characterized (belief-desire) behavior without taking the content level into account. This comes close to the independence from strictly material determinism that Tallis wants, and, importantly for Tallis, it makes purposive action clearly distinct from "bottom-up" physical events or happenings (e.g., the effects of psychoactive drugs on the brain), even if intentionality is grounded in neural processing.

Note, however, that such grounding tethers the intentional level to its physical underpinnings even if the transitions between brain states might be best understood as determined by representational content. Although intentionality is likely not *explicable* at the neural level alone since it involves complex past and present interactions between the person and world which, in concert with genetically installed representational capacities, help determine the "aboutness" of mental states, it nevertheless remains causally bound and limited by its neural supervenience base. The sequential regularities involved in content-level representations as they occur to us in a mental simulation, for example when I consider the feasibility of adopting a universal basic income (UBI), cannot contravene the scientifically described laws – physical, chemical, or biological – that govern the neural vehicles carrying such content. And of course, I rationally want the neural processes themselves to be reliable and law-governed at the physical levels, not subject to any sub-personal indeterministic whim. Such whim would compromise, not strengthen, the machine level substrata of my reasoning.

Further, note that content-level transitions too normally involve a regulative determinism, but at the level of rationality (e.g., logical implication) if not physical causation.

⁶ On this point, but in the context of explaining consciousness, see Clark, 2019, section 7, 'Phenomenal-physical parallelism.'

⁷ (Tallis, 2021, pp. 106–132) section 4.3, 'Actions, causes, and propositional attitudes.'

Daydreaming aside, trains of thought are often (and fortunately) governed by formal and informal rules of inference, abduction, and generalization, learned over a lifetime, which can in principle explain why – since they *determined* why – I ended up choosing a particular option from a menu of considered possibilities, e.g., my favoring a UBI. We rationally *don't* want any indeterministic slack or noise inserting itself into what we hope is a reliable process of deliberation in which goals and reasons determine and explain behavior. And, crucially, we don't want to be free to have chosen otherwise given what those were at the time, even though we might regret the choice made (see section 5). *That* sort of freedom would only put us outside our own mental processing, leaving us at a serious practical disadvantage.

Finally, we can also see that mental simulation itself usually involves the anticipation of robust regularities holding across multiple explanatory domains, whether it's how we expect physical objects to interact, people to behave, institutions to function, or social and economic trends to more or less follow previous patterns. In deliberating about what to do, either alone or in concert with others, we don't usually imagine things are going to go off the causal, or more broadly, explanatory, rails; rather, very much the opposite. We are, in effect, what I would call *pragmatic determinists*, who normally apply our best anticipatory guess about possible outcomes in light of what's determined them in relevantly similar past situations, where the determinants fall into various domains, such as physical, psychological, rational, and social.

Mystifying the mental

The argument might be put that since we don't yet have a complete physical account of what underlies our mental capacities, whether we're talking about intentionality or phenomenal consciousness, then we can't claim that such capacities are necessarily constrained by the physical determinism that might obtain in the brain (remembering that Tallis rightly discounts indeterminism at the material level as the basis for any free will worth wanting). This would be, in effect, an argument from ignorance – an explanatory gap. And indeed, Tallis says that human agency, based in intentionality and consciousness, *must* remain mysterious since these are, he claims, immune to scientific explanation.⁸ But of course it needs demonstration that these capacities necessarily transcend what scientific investigation might reveal about what the brain, in its bodily and social context, makes possible. Such science is well underway, and further research into the nature of our mental capacities will likely tighten the connection between them and their physical basis. Tallis suggests such progress is unlikely given what he sees as a basic ontological divide: “. . . propositional attitudes cannot be identified with material events in or material states of the brain since there is nothing in neural activity to sustain the intentionality of items of

⁸ “Agency is real and so, too, is the mystery of agency and the intentionality that makes it possible... The dark heart of that mystery is our *unfathomable ability* to see the light, to see or know how things are, and face a world revealed, by literal light or the light of knowledge, to a self-revealing self... There is *nothing in natural science that gives us any insight into how it comes about* that there is something in nature that transforms the anterior surface of the head into something that confronts it, describes it, gives it a name, and puts that name in inverted commas thus: ‘Nature’.” (Tallis, 2021, pp. 169–70), emphasis added.

consciousness of any sort.” (Tallis, 2021, p. 105). Whether or not propositional attitudes are literally *identical* with brain states or events, as opposed to involving the representational *content* carried by those states and events, is, I think, an open question. But why believe that nothing in neural activity sustains intentionality, when clearly the brain is the material basis for all our mental functioning? The science of representation, of mental content as carried by neural and artificial systems, aims to understand the material organization of a system’s internal states that are plausibly construed as representing states of the world and the system within it.⁹ Propositional attitudes are not obviously forever beyond reach of such understanding.

Even if intentionality escapes full naturalization in a formal theoretical framework – only those of future generations satisfied that mind science is complete will know whether it does or not – the exercise of even our highest, most abstract, mental capacities will nevertheless remain a natural phenomenon. We *evolved* to be as smart as we are, and the role of culture in leveraging intentionality doesn’t escape its physical basis, but simply builds upon it. It’s only because the mental content of intentionality is carried by neurally realized processes that connect to and govern speech and action that it could get a grip on physical behavior; the material brain is thus not a prison we need escape.¹⁰

Doing otherwise

We can clarify what’s at issue by posing a question raised earlier: in an actual situation, all conditions and factors, internal and external, as they were, could you have done (thought, simulated, deliberated, chosen, acted) otherwise? Tallis, believing we have metaphysically real alternative possibilities (see note 3), would claim that yes, you could have, and not in a way owing to any sort of non-self, sub-personal randomness, but in a way that’s *up to you* – the basic commonsense requirement of having free will. Having this *unconditional ability to do otherwise* (do otherwise under the exact same conditions) is, I’d suggest, what’s really at stake in our purported distance from nature.¹¹ Tallis would of course want the agent’s exercise of their mental capacities, especially the consideration of counterfactuals, to *determine* (account for, explain) the outcome in an actual choice scenario. Such *agential determinism* (as distinct from libertarian “agent causation,” which is premised on indeterminism)¹² seems necessary for the assignment of responsibility, credit, and blame. But Tallis doesn’t think the agent’s exercise of her mental capacity for simulating possibilities is *itself* deterministic or determined; rather, she could have simulated otherwise given the same initial conditions. Only having metaphysically real alternative possibilities, one of which is selected by the agent, can put us sufficiently outside nature to give us free will. The mere *epistemic* openness of the future for us will not suffice.¹³

⁹ See for example Churchland (2013) and Shea (2018).

¹⁰ “The prison of the universe is personalized in the prison of the brain.” (Tallis, 2021, p. 15)

¹¹ On the unconditional ability to do otherwise, contrasted with the conditional or counterfactual ability, see Nadelhoffer et al. (2020).

¹² On libertarian agent causation and indeterminism, see for example Chisholm (1966).

¹³ Epistemic openness – not knowing what the future holds – might naturally give rise to the intuition of a metaphysically open, unfixed future.

The difficulty here is to explain how, given the particular state of the agent in an actual situation, where such a state might be described at the physical, psychological, and intentional (belief-desire) levels, a different outcome would have resulted in a way that wasn't arbitrary and thus non-agential – not *up to me*. Given my beliefs and desires, which presumably determined my reasons for doing X and not Y, why *wouldn't* I have done X (and so might have done Y), given that my reasons and deliberations were presumptively determinative? Why *would* I have done otherwise, given all factors and conditions as they were? I want the process of weighing considerations for and against possible courses of action to be a reliable reflection of my character and my relevant standing and immediate priorities, even if that process sometimes fails to deliver what in hindsight would have been better. I want my character and priorities, along with impersonal reason and evidence, to reliably *determine* my deliberations in concert with the specific variables in play. Given this requirement for genuine and effective agency, I couldn't have done otherwise in an actual situation: quite fortunately I *don't* have the capacity to transcend the motives, deliberations, and intentions in play, although of course I might wish they had been otherwise or played out differently. Nor can I transcend their neural basis that partially constitutes me as a material being, nor would I rationally want to. All told, we shouldn't hanker after the unconditional ability to do, and in particular, *simulate* otherwise in actual situations. We don't need it to be active agents – not mere conduits of causation – whose mental processing adds causal value to the world. Rather we need, and usually have, given our behavioral and cognitive flexibility, the *conditional* or *counterfactual* ability to do otherwise in relevantly different situations, e.g., in imagined or future situations, in response to the different conditions in play (Nadelhoffer et al., 2020). Had different arguments been presented to me in opposition to a UBI, I might have concluded otherwise about its desirability.

In citing the agential determinants of a choice, whether at the physical, psychological, or reasons-based levels (all might apply), determinists have, in principle, a clear explanation of it. To have been otherwise, something in its explanatory antecedents would have had to have been otherwise, but, crucially, this sort of traceability doesn't undermine the central role of the agent's internal mental processing in the explanation – Tallis's main concern. If, as an anti-determinist might claim, the agent herself isn't fully traceable to sufficient physical, psychological, and reasons-based factors that explain her current status (and thus her choice), that lack of traceability means something *outside the agent's own prior characteristics* figures in that explanation. Trying to escape traceability to current, historical, personal, or sub-personal factors and conditions when accounting for agents and their choices simply mystifies agency (section 4 above). The person becomes an unintelligible black box of choice (Clark 2005), and it renders the person's contribution to self-formation and current behavior *less*, not more, determinative. Such freedom, were it real, would be disempowering. Instead, we rationally should want our deliberations to be reliably determined by formal and informal rules of inference given reliable capacities for remembering, perceiving, and anticipating circumstances, all of which we rationally want to be subserved by reliable sub-personal sensory, perceptual, and information processing mechanisms.

Should a random or indeterministic factor, say at the quantum level, play a role in behavior, we can say the agent might have done otherwise in an actual situation, but not in a way that affords more control and responsibility than under a multi-level explanatory

determinism, and likely less. For all practical purposes, therefore, we can usually ignore such factors (they are likely minimal in any case) and focus our efforts at explanation and control of action on the physical, psychological, and reasons-based determinants we know exist, while continuing to investigate the complexities of human behavior. Such *pragmatic* (not universal) determinism acknowledges the full range of our capacities, including conscious intentionality, but it also acknowledges their material basis with which the exercise of our mental abilities must, as obviously natural endowments, be consistent. Pragmatic determinism situates us entirely *within* nature, a nature that, through us and our precocious ingenuity, has given birth to a culture and technology that may generate artificial agents that surpass us in their abilities to know, simulate, predict, and control.

The viability of inclusive naturalism

The supposition that intentionality and the contents of consciousness somehow float free of their multi-level determinants, for which there is little or no evidence, puts us at a practical disadvantage. Such a supposition might make us less motivated to understand the actual conditions necessary for putting reason to use in service of human and planetary flourishing. And to suppose we could have done otherwise in actual situations, but simply chose not to in a way that's ultimately and untraceably up to us, allows for assigning a very strong sort of libertarian credit and blame. It illicitly singles out the agent as a first cause, one deserving, for instance, of harsh retribution for crimes or unlimited rewards for achievements. I'm not suggesting that Tallis endorses such attitudes and policies, only that belief in human causal exceptionalism – our untraceability – can have practical and ethical pitfalls, what some have called the “dark side” of free will (Nadelhoffer & Tocchetto, 2013, Caruso, 2014, Caruso, 2019).

All told, I see no good reason to want to break free of the pragmatic, multi-level, explanatory determinism I've advocated here, one that can incorporate intentionality as a real but fully natural capacity. Our abilities to consciously reason and imagine don't have to be causally disconnected from material reality to be real and effective determinants themselves; rather they need their own reliable material platform upon which to build virtual worlds to explore and sometimes make real. Were we not fully subject to physical, psychological, rational, and social determinants, we'd be loose cannons, rebels without causes, little local gods acting for no good reasons and beyond reach of social control. Fortunately, we are not free in this respect, even as we remain real, active agents who can articulate and defend the actual personal and political freedoms we hold so dear.

The possible negative consequences of belief in a human distance from nature versus the possible upsides of accepting a naturalistic pragmatic determinism, for example when preventing and treating addiction (Clark, 2021, 2022), do not of course bear on the truth about human agency. Questions about the type and range of our physical and mental agential capacities are matters for empirical investigation, and the project of our scientific self-understanding proceeds apace. Tallis's claim that conscious intentionality will never be encompassed within a science-based, naturalistic explanatory framework, and that we are therefore exempt from naturalization in that respect, is vulnerable to disconfirmation. Moreover, it seems to me unwarranted to even want such exemption, given that successful

embodied human agency depends on reliable physical mechanisms that underpin our (reasonably) reliable reasoning and simulational capacities, which we normally exercise cognizant of what we rightly suppose are robust regularities in animal and material behavior. If, as I think the evidence in hand already strongly suggests, we aren't in any sense outside nature and the multilevel determinants that we embody, that explain us, and that we act in light of, this will not constitute any sort of imprisonment, but rather a liberation: a mature realization of our true place in the natural world that will, I have argued elsewhere, confer many practical and ethical advantages (Clark, 2007, 2016).

Acknowledgements

I'm grateful to James Tartaglia for helpful comments on this paper.

References

- Caruso, G. (2014). The Dark Side of Free Will. Talk for TEDx Chemung River. <https://www.youtube.com/watch?v=rfOMqehl-ZA>
- Caruso, G. (2019). Free will skepticism and its implications: An argument for optimism. In E. Shaw, D. Pereboom, & G. Caruso (Eds.), *Free will skepticism in law and society* (pp. 43–72). Cambridge University Press, <https://philpapers.org/archive/CARFWS-2.pdf>
- Chisholm, R. (1966). Freedom and action. In K. Lehrer (Ed.), *Freedom and determinism* (pp. 11–44). New York: Random House.
- Churchland, P. M. (2013). *Plato's camera: How the physical brain captures a landscape of abstract universals*. Cambridge, MA: MIT Press.
- Clark T.W. (2005). Hodgson's Black Box. *Journal of Consciousness Studies*, 12(1), 20–75. <https://naturalism.org/philosophy/free-will/hodgson%E2%80%99s-black-box>
- Clark T. W. (2007). *Encountering naturalism: A worldview and its uses*. Center for Naturalism. <https://naturalism.org/worldview-naturalism/encountering-naturalism-a-worldview-and-its-uses>
- Clark, T. W. (2016). Naturalism and well-being. In P. Zuckerman (Ed.), *Religion: Beyond Religion*. New York: Macmillan Reference USA. <http://www.naturalism.org/sites/naturalism.org/files/Naturalism%20and%20well%20being.pdf>
- Clark, T.W. (2019). Locating consciousness: Why experience can't be objectified. *Journal of Consciousness Studies*, 26(11–12). https://naturalism.org/sites/naturalism.org/files/Locating%20Consciousness_0.pdf
- Clark, T. W. (2021). Determinism and destigmatization: Mitigating blame for addiction. *Neuroethics*, 14, 219–230. <https://doi.org/10.1007/s12152-020-09440-w>
- Clark, T. W. (2022). Causal responsibility for addiction. *Addictive Behaviors*, 130, 107292. <https://doi.org/10.1016/j.addbeh.2022.107292>
- Dennett, D. C. (1987). *The intentional stance*. Cambridge, MA: M.I.T. Press.
- Metzinger, T. (2003). *Being no one: The self-model theory of consciousness*. Cambridge, MA: MIT Press.
- Nadelhoffer, T. & Tocchetto, D. G. (2013). The potential dark side of believing in free will (and related concepts). In G. Caruso (Ed.), *Exploring the illusion of free will and moral responsibility*. New York: Lexington Books.
- Nadelhoffer, T., Yin, S., & Graves, R. (2020). Folk intuitions and the conditional ability to do otherwise. *Philosophical Psychology* 33(7), 968–996. <https://philpapers.org/rec/NADFIA-2>.
- Shea, N. (2018). *Representation in cognitive science*. Oxford: Oxford University Press.

Tartaglia, J. (2020). *Philosophy in a technological world: Gods and Titans*. London: Bloomsbury Academic.

The Heller School for Social Policy and Management
Brandeis University
415 South Street MS 035
Waltham, MA 02453
USA
Email: twclark@brandeis.edu