The Priority of Preferences in the Evolution of Minds

David Spurrett

Philosophy, UKZN, South Africa, 4041 | spurrett@ukzn.ac.za

ORCID: 0000-0003-2542-358X

Abstract

More philosophical effort is spent articulating evolutionary rationales for the development of belief-like capacities than for precursors of desires or preferences. Nobody, though, seriously expects naturally evolved minds to be disinterested epistemologists. We agree that world-representing states won't pay their way without supporting capacities that prioritise from an organism's available repertoire of activities in light of stored (and occurrent) information. Some concede that desire-like states would be one way of solving this problem. Taking preferences as my starting point instead of belief-like states, I defend two conclusions. First, psychologically real preference states, which approximately token expected utilities, have a quite general evolutionary rationale. They are a solution to the problem of efficiently allocating capacities with incompatible uses. This argument is a version of the Environmental Complexity Thesis. Second, preferences can plausibly function and naturally evolve without belief-like states, even though the converse claim is incredible. Preferences, that is, can mediate between discriminations of occurrent states ('internal' or 'external') and the processes selecting activity without mediation by stored indicative representations. By tokening expected utilities of actions conditional on discriminated state, they can increase the rate at which the 'right thing' is done at appropriate times, and they can do this without the support of belief-like, world-representing states. Preferences, even incomplete and noisy sets of them, are a fuel for success that will tend to be favoured when environments are complex in ways that matter to an organism, and when the organisms have complex behavioural repertoires with heterogenous returns and costs.

Keywords: Preferences, Evolution, Folk psychology, Environmental Complexity Thesis, Reinforcement Learning

Acknowledgments: This paper was presented at the Wits Philosophy seminar in August 2023, and as a keynote at the International Society for the Philosophy of the Sciences of the Mind in December 2023, before presentation at the PSSA 2024 meeting in Stellenbosch in January 2024. I am grateful to audiences on all occasions for useful discussion.

1) Introduction

Among the questions arising when thinking about the evolution of minds are ones corresponding to the folk psychological kinds of *belief* (what an agent represents as being the case) and *desire* (what an agent wants). In this short paper I'll take for granted that folk psychology provides a credible explanandum, so that it is worth trying to provide accounts of the origins and – possibly incremental – development of beliefs and desires. I'll also adopt without defending a broadly naturalist orientation. In this first section I'll sketch key features of some existing approaches to naturalistic explanation of folk-psychological capacities, and how they tend to prioritise proto-beliefs over proto-desires. In the following section I'll provide a compressed statement of an evolutionary rationale for preferences, where preferences are understood as a minimal desire, or significant kind of proto-desire. In sections (3) and (4) my focus moves to developing an argument that proto-desires could have evolved first, and so preceded proto-beliefs in the history of cognition.

The argument that I will develop here takes as its starting point the general approach found in Godfrey Smith (1996) and Sterelny (2003). According to what Godfrey Smith called the 'Environmental Complexity Thesis' (the ECT) "…the function of cognition is to enable the agent to deal with environmental complexity" (1996, p3). The key idea of 'environmental complexity' does not refer to some agent-neutral property of reality, but rather to the fact that the world around an organism is sometimes *heterogenous in ways that matter* to that organism. When the local environment that matters to an organism is simple, it can be enough for the organism to respond in stereotyped ways to easily detectable features of it (for example by always approaching light). When the environment is significantly heterogenous, what it is appropriate for the organism to do won't be reliably or simply related to what can be cheaply detected. Sterelny calls such environments *opaque*, and argues that some opacity will be caused by *hostility* – the active production of misleading signals by other organisms. In those cases simple cue-bound responses become less effective and may be maladaptive (candles were an evolutionary novelty, leading many moths astray). The idea of the ECT is that in some cases cognition can pay its evolutionary way by making the actions of organisms more appropriate in the face of significant environmental heterogeneity. The development of cognition in the face of complexity isn't inevitable, and depends among other things on whether the behavioural returns more than cover the costs of cognition, and what cognitive mechanisms are evolvable from what starting points. The general sort of reasoning here can be applied to any cognitive capacity, but my focus here is on the folk-psychological kinds of belief and desire.

Sterelny (2003) provides an instructive illustration of the approach applied to belief. In his treatment the paradigmatic *simple* cognitive solution is called a 'detection system', which triggers a behaviour when a single external cue is detected. Sterelny calls an environment informationally *transparent* if it is "characterized by simple and reliable correspondences between sensory cues and functional properties" (Sterelny 2003, p20). Such an environment is low in environmental complexity. In it a relatively cheap and simple detection system can be a satisfactory control system, by mediating "a *specific adaptive response* to some feature (or features) of [an organism's] environment by registering *a specific environmental signal*" (Sterelny 2003, p. 14). One of Sterelny's key illustrations is the cockroach flight response. Hair cells on the heads of cockroaches respond to onward gusts of air, and trigger running away (Sterelny 2003, p14). Since enough of the gusts close to the heads of cockroaches are caused by predators, this is adaptive. Sterelny goes on to argue that cognitive elaborations of detection systems in response to increased complexity include 'robust tracking' – that is, sensitivity to multiple integrated factors instead of a single cue – and 'response breadth' – that is, having more than one possible response to a contingency. He also argues that 'belief like states' or 'decoupled representations' can evolve when response breadth and robust tracking are combined. When they are so combined there are relatively enduring cognitive states sensitive to multiple detected sources of information, and available as co-determinants of more than one behaviour.

Calling these 'belief-*like*' states isn't idle. Decoupled representations needn't be 'personal' states, or introspectively available, or have various other paradigmatic properties of beliefs when we're thinking about language-using humans. The evolutionary processes being contemplated here mostly involve non-human agents, and the cognitive capacities of interest include intermediate steps on the way to something a philosopher might count as 'full' belief. A 'belief like state' is a to some degree decoupled world-representing state that may be only shallowly conceptual, i.e. classifying as labelling, or as describing with modest inferential and no compositional capacity (Brandom 2009). Similarly proto-desires aren't the personal goal-representing states that can (with beliefs) feature in rationalisation of intentional action. In the argument that follows I'll use *preferences* as my default proto-desire, indicating a motivational state that prioritises among (some of) the available actions of an agent (see §2 below). Preferences also don't need to be person states, or available to higher cognitive processes.

It is striking that philosophers working in this general area tend to prioritse providing accounts of belief-like states over addressing preferences or motivation. The relevant parts of Godfrey Smith's (1996) *Mind and its Place in Nature* are mostly about true beliefs as a 'fuel for success'. He says himself that "the bulk" of the book is "focused on explaining the content of belief-like states in particular, as opposed to desire-like states" (1996, p. 175). The first part of Sterelny's (2003) concerns the development of folk-psychological kinds, but three quarters of it is about belief-like states, and his treatment of preference is not only shorter, but rather deflationary (for a critical treatment see Spurrett (2015)). As a final exhibit, Millikan's (2017) *Beyond Concepts* is an extended articulation and defence of speculative theoretical psychology replacements for non-psychologistic 'concepts', called unitrackers and unicepts. She allocates much more attention to 'factic' than 'affording' unitrackers and unicepts, where affording ones are more clearly related to motivation. I draw attention to this pattern not with a view to criticising it, or offering an objection to it, but in order to signal the fact that in this paper I'll invert the standard approach and prioritise preferences in the sense of giving them more attention than belief-like states.

2) The efficiency rationale for preferences

The tendency to devote more attention to the evolution of belief-like states has some exceptions. Spurrett (2021a) develops a fairly general ECT case for the evolution of preferences, which I'll adopt here. In that argument preferences are cognitive states that rank available actions conditional on detected states of the world, and the organism itself. They are explicitly intended to be candidates for minimal desires, or proto-desires, for at least some accounts of desire. With preferences thus understood, the paper argues for the 'efficiency rationale for preferences' (or ERP) according to which "Preferences enable efficient action selection." This is the basic formulation of the ERP, and although the paper goes on to develop a more qualified and hedged statement, the simple version is all we'll need here.

The applicable notion of efficiency, this being an evolutionary argument, is *biological*, and could be understood in terms of fitness. The argument for the ERP, in compressed outline, has two main moving parts. One is an argument that achieving efficient action selection can itself be demanding because of the complexity of the problem. The other is an argument that preferences can be a cognitive solution to that problem. Making action selection efficient can be difficult for an organism because "actions generally have varying (and multi-modal) costs and returns" (Spurrett 2021a, p. 491), where the costs include metabolic factors, time, risk and of course opportunity costs, and the returns include various kinds of nutritional intake, hydration, mating opportunities, shelter, rest and so forth. Action selection is a task of making allocations from an action repertoire that delivers decent returns given these varied and heterogenous bundles of consequences. Problems involving these kinds of many-many mappings are paradigmatically 'complex' in the sense relevant to the Environmental Complexity Thesis. A system of preferences could solve this problem if the preferences attached values to actions conditional on detected states in ways that were sufficiently efficient. They would provide "situation-specific rankings of available actions" (Spurrett 2021a, p. 494), in the sense that actions leading to water intake would be up-ranked in a dehydrated organism. In order to deal with trade-offs between many reward modalities, the preferences would have to be in a general or common currency, for which the existing term 'utility' is well-suited. Such a solution would not, of course, be without cost. A collection of cognitive adaptations would be needed to attach expected values to available actions, and update those values as the needs of the organism and state of the world around it changed. If such a set of adaptations returned more than it cost, though, the argument would work. And empirical discoveries in neuroeconomics and cognitive neuroscience suggest that circuits that engage in tracking the expected returns from actions across many modalities of reward are found in many taxa, and have deep evolutionary roots.

The 'Evolutionary Rationale for Preferences' has the same general structure as Sterelny's case for the evolution of belief-like states. Both characterise a class of problems which relatively simple cognitive don't solve very well, and go on to argue that a certain kind of cognitive investment could solve it better. In one case the cognitive capacity is the world-modelling of decoupled representation, in the other it is mechanisms of tracking expected returns across many modalities.

3) Chickens and Eggs

It is reasonable to wonder how proposals about the evolution of proto-beliefs, such as Sterelny's account of decoupled representation, and ones about proto-desires, such as the Evolutionary Rationale for Preferences, relate to one another. Suppose, to narrow the question somewhat, a case where we have one prima facie satisfactory version of an evolutionary rationale for proto belief and one for proto desire, and that the relationship of interest is temporal ordering. Which, if either, came first?

In the standard 'full folk psychology' case, belief and desire are *both* required to explain (or cause) action. This is part of Hume's point that "reason is, and ought only to be, the slave of the passions" (1975, p. 415). While Hume's concern was to argue that reason couldn't motivate by itself, the converse point holds: my strong desire for a snack won't send me to the pantry unless I believe there are cookies in the jar. That neither gets anything done on its own at least suggests that they may have had to evolve together if they evolved at all. But we should not be too hasty. The order of invention can be surprising in technology – a process for storing food in tin cans was patented over four decades before a general purpose can opening tool. Some inventions can be beneficial enough to get maintained on their own, and by being in use make gains from further complementary innovations possible. This can be so even when we've become used to having both at once (see, e.g. Hejnol & Martindale 2008).

In the case of the evolution of belief-like states and preferences or proto-desires, it seems as though three options regarding the order of development should be taken seriously. One is that they developed together, which might have involved the two being incrementally differentiated from an earlier capacity that combined elements of both (as, for example, Millikan's "pushmi-pullyu representations" combine indicative and imperative aspects – Millikan (1995)). The remaining two options are that one of the two developed first and, at least initially, conferred sufficient advantage on its own to be maintained before the other came along.

In the following section I argue that preferences *could* have come first, and conferred benefits without decoupled representations. So I'm not just prioritising preferences by focusing on them, I'm arguing that they *could* have preceded belief-like states in the evolution of minds.

4) Putting preferences first

My main argument for the claim that preferences could have evolved before decoupled representations starts from the fact that model-free reinforcement learning (RL) in computer science can solve *some* efficient action selection problems (see Sutton & Barto 1998, 2018). In doing so, model free reinforcement learning both involves or relies upon preferences, and operates without world-modelling or anything like decoupled representation. Consequently, it provides a kind of existence proof of preferences paying their way without decoupled representation, that is of proto-desire without, and hence potentially *before*, proto-belief. The following few paragraphs spell out the workings of this argument in a little more detail.

Reinforcement Learning is a large and rapidly moving interdisciplinary field. That doesn't matter here because I'm not trying to assert anything about the whole field. The claims I need are core results, subject to standard restrictions. In much reinforcement learning an agent is confronted with a finite Markov Decision Process (MDP) and learns a 'policy' relating states to actions, where the consequences of actions can include reward or reinforcement. (The policy, whether in the form or a lookup table or some kind of function or algorithm, is broadly analogous to a set of stimulus-response associations in behaviourist psychology, or to a control law in control theory.) Markov Decision Processes were extensively studied in optimal control theory, which provides rigorous criteria for optimality. Some techniques for solving MDPs – such as dynamic programming – require models of the dynamics of the MDP. But some don't. In 'temporal difference' (TD) learning a policy (which can initially be random) is followed, and the reward consequences of following it compared to the expected reward associated with it. Any mis-match results in an error signal (the difference between actual and expected reward at a time) which is used to drive incrementally updating the policy. Model-free TD learning can approach optimal solutions to some reinforcement learning problems. Solving them amounts to achieving efficient allocation of an action repertoire contingent on a changing environment, and so is equivalent to the explanandum in the ERP.

There are two main ways in which the implementation of TD learning, including model-free TD, involve preferences. First, the policy organises actions conditional on states into determinate relations of being better than, or as good as, one another. That is precisely what preferences do in the ERP argument: value the elements in an action repertoire conditional on the detected state of world and agent. Second, the driver of learning is an error signal based on the magnitude of reward. While it is possible to have preferences without the capacity for individual learning, the converse is not the case when the learning is by reinforcement. Reinforcement is evaluative. (The relation between reinforcement learning and Thorndike's 'law of effect' is outlined in Sutton & Barto 2018, pp. 13-21.) Reinforcement learning in general is glossed as "learning what to do […] so as to maximize a numerical reward signal" (Sutton & Barto 1998, p. 3).

Finally, what it means to be 'model free' is to lack an independent world representation (in this case a model of the finite MDP, which represents how states follow states, or how states follow actions, allowing evaluation of downstream states). Such systems lack decoupled representations, or proto-beliefs. Putting the pieces in play here together, model free reinforcement learning can solve efficient allocation problems, in ways that do require the functionality of preferences, but without decoupled representations. That is why I say that model free reinforcement learning provides a kind of existence proof of preferences *without*, and hence potentially *before*, decoupled representation.

This isn't a party trick argument based on a biologically fanciful toy model, because (as noted above) the reinforcement learning problem is quite general, and solving it is equivalent to other optimality concepts in good standing (e.g. from control theory), and the notion of a 'policy' relevantly analogous to other significant theoretical constructs. It is also increasingly clear that the brains of many real organisms engage in some kind of temporal difference learning. Schultz, Dayan and Montagu (1997) for example reported empirical work showing that activity of midbrain dopamine neurons did not make sense if dopamine was thought of as related to pleasure or reward, but did if it was regarded as a reward prediction error. (Crucially dopamine activity did not vary when expected rewards were delivered, but did spike in response to unexpected predictors of later rewards.) On the other hand, desktop reinforcement learning systems aren't generally neurally plausible. In addition, some successes of 'model free' reinforcement learning are reliant on implicit epistemic resources in ways that undermine the inference I'm drawing to the possibility of preferences without proto-belief. What I mean by this is that in much reinforcement learning research problems of perception and classification are solved by fiat in order to focus on one specific learning problem. So, for example, a reinforcement learning system learning a policy to play chess won't be expected to solve the problem of disambiguating bishops from pawns in an optic array but will occupy a world in which the complete state of the board is the current state of the MDP it faces. This isn't an objection to anything about reinforcement learning research. The point is that sometimes what gets called a 'model-free' reinforcement learning system – and is indeed model-free in the relevant technical sense – has significant implicit epistemic resources. This doesn't defeat the general point of my appeal to what model free reinforcement learning can do, it simply recommends caution about what counts as 'model free'. A further reason for caution here is that many reinforcement learning models tend to be rather *disembodied*. Unless the learning problem is itself a motor control one (like driving a walking robot) reinforcement learning systems are no more likely to be expected to deal with the physical world than AI chess systems are expected to tell where the pawns are by visual inspection. The actual evolution of cognition was, though, undoubtedly a very embodied affair, and in ways that matter to what kinds of embodied, non-representational and in other ways minimal cognition were possible. A research programme that so often abstracts away from bodies might easily lead us astray. In sum, there is some good news about neural plausibility, and some cautionary notes about being 'model free' and disembodied, so this 'existence proof' should be handled carefully.

The worries reviewed in the preceding paragraph can, I suggest, be addressed by a complementary argument that doesn't start from premises related to either preferences or proto-beliefs, but with considerations about the architecture of embodied control systems.

To begin imagine, or at least *try* to imagine, a medium sized multi-cellular creature that has a separate body part for each distinct physical behaviour or activity. The creature would be something like a living 'Swiss Army Knife', with a part or appendage for drinking liquid, and another one for putting food in its mouth, and another still for scratching its head and so on. It gets hard to imagine this quite quickly, because the accumulation of perfectly realistic functions (things you can see a real pigeon or dog do) leads to architectural problems. Where can all the appendages go? It should also be clear that if we assume that such a design is possible at all, it would have to be very inefficient compared to real animals in two ways. First, the 'Swiss Army Critter' would have to be much larger to achieve the same effect, and second it would carry around a large number of very rarely used parts.

The point of the fanciful exercise in the previous paragraph is to make vivid something so familiar that we can take it for granted, which is the fact that the real evolved multicellular bodies typically have hierarchically organised, multi-use organs. The most obvious cases of this are provided by jointed limbs which are variously used for walking, scratching, climbing, hitting, manipulating, and so forth. This overall efficient design solution isn't an unalloyed benefit, because the opportunity cost of some uses of the organs will be forgoing others. (A quadruped like a dog or cat can't walk and scratch at the same time.) And getting the benefit of such a body means having some way of solving the scheduling and control problem of timing switching between deployments of multiple-use capacities, and blocking attempts at incompatible uses. This is also, it is worth noting, an exemplary *economic* problem, involving allocating scarce means with alternative uses.

Sherrington's (1906) notion of a 'final common path' helps understand some of how real evolved nervous systems deal with this problem. He introduced the notion to refer to the last neural stage at which competition between incompatible deployments of combinations of muscles can be resolved, and recognised the need for this resolution because some different movements made conflicting demands on the same muscles. He characterised a final common path, relative to a pair of 'reflexes' (bodily movements at various scales or levels of co-ordination) that couldn't be active at the same time, as the last outbound location in the nervous system that was able to switch between them.[1] We can separate that idea from Sherrington's specific notion of reflexes and recognise, as Gallistel (1980) puts it, that arbitration relationships for control of the body must take the form of a 'lattice hierarchy' in which the level at or before which competition over deployment of degrees of freedom must be resolved is highly variable. We contract our biceps more easily if we relax our triceps at the same time, and simple and highly automatic relations of reciprocal inhibition typically facilitate this. Higher level and more sophisticated co-ordination, not to mention learning, is required to arrange our body so that we can tie a shoelace without falling over.

There are various possible ways of implementing the switching at final common paths. It could just be more reflexes in some 'hard-wired' hierarchy. Or a creature might switch between broad modes (resting, foraging, fleeing) associated with different patterns of using the body. Although biologically plausible, we can imagine a robot that used a set of lookup tables to resolve competing allocations at its final common paths. Another possibility, of course, is that the resolution of the competitions could be sensitive to the value of the options. While final common paths don't *have* to involve preferences, they provide a functional 'place' for them to be exercised. And important early research in neuroeconomics proceeded by studying some relatively localised neural circuits know to serve final common path functions. Saccadic eye movements have several advantages for neuroscientific study. Eye movements relative to the skull have a simple geometric organisation, and a simple topographic representation. Movements depend on only six specialised muscles per eye, and movements to different targets are strictly mutually exclusive. Known neural circuits represent possible saccade targets with nearby targets associated with adjacent cells in a folded up dartboard arrangement. The research found that activity in neural maps of saccade targets upstream of final common paths both predicted actual saccades, and was correlated with the learned expected value of the saccade targets (Glimcher & Sparks 1992, Platt & Glimcher 1999, Glimcher 2003, see also Glimcher 2011). Preferences, that is, are in fact neurally expressed in some final common paths.

[1] It is important that a final common path is relative to specific combinations of alternative movements. See (Spurrett 2021b).

Here's how this complements the argument from model-free reinforcement learning. First, a nearly universal solution to a body design problem is multiple-use moving and flexible (including jointed) organs with a hierarchical control system providing a convenient *location* for preferences, even if the organism has no investment in world-modelling (decoupled representation). Second, In the actual evolution of brains, it appears plausible that the difference between motor arbitration and reward learning is one of degree. Neural reinforcement learning began as, and still bears traces of being, an elaboration of motor gating (see, e.g. Barron, Søvik & Cornish 2010). The 'existence proof' argument complements the actual history of life. These considerations directly address the top two concerns about how much weight the inference from model-free reinforcement learning could bear: neural plausibility, and tendency to be disembodied.

An objection might arise at this point, going something like this: The argument above doesn't describe a scenario *without* belief-like states. The architectures being described don't actually lack proto-beliefs because the preferences rank actions conditional on discriminated world states, and the current needs of the organism. These external and internal factors are represented in the cognitive processes of the agents. And in representing possible matters of fact about the world or the agent itself, they should be counted as proto-beliefs. Perhaps they could count as *perceptual* proto-beliefs, in being closely related to perceptual relationships with the environment, and self-perception. One reason to take this worry seriously is provided by the point I made above, that care needs to be taken in assuming that every case of 'model free' reinforcement learning suggests anything about plausible embodied natural agents. But I urged taking care because I think that the inference is good in a large enough class of cases. So, unsurprisingly, I don't think the objection is decisive here. If we impose reasonable and standard demands for counting something as a proto-belief, then the objection doesn't work because the various discriminations and even classifications involved in the model free reinforcement learning scenarios don't involve proto-beliefs. They needn't be significantly decoupled, or flexibly deployable in our cognitive processes, and certainly don't need to be able to sustain inferences or exhibit any kind of compositionality. Sterelny argues that the kinds of discrimination involved in his simple 'detection systems' (which are articulated in order to be paradigmatic cognitive solutions *lacking* key features of proto-beliefs) require some level of generalisation, consistent with them being simple classifiers. That's enough to get model-free reinforcement learning into play, and well short of even the minimal test of a proto-belief that it be decoupled enough to be involved in cognitive processes in the absence of what cues it. If, on the other hand, we do count these discriminations as cases of proto-belief we intolerably trivialise the notion of proto-belief by applying it to purely occurrent states with no cognitive depth.

There's more, that is, to proto-*belief* (which involves decoupled representation) than discrimination. Counting discrimination as representation is too permissive, and dilutes the explanatory point of intermediate world-modelling cognitive states. Sea slugs are capable of some reinforcement learning (e.g. Perry, Barron & Cheng 2013) but have not been found to have abilities requiring hypothesis of decoupled representation in the sense of showing behaviour apparently conditional on non-occurent stored representations. Giant sea slugs, for example, are carnivorous and eat animal matter they encounter in fairly indiscriminate ways, "including other sea-slugs and their eggs" (Manning and Dawkins 1998: 226). Despite this, they do not eat their own eggs during egg laying. This selective disposition is realised as follows: during egg laying, they release a hormone that inhibits movement of the mouth (Davis *et al* 1977). That does the required evolutionary job without (proto) beliefs about descendants or even (in this case) preferences that over-ride eating.

In conclusion, the reinforcement learning argument supplemented by the final common path considerations encourages the thought that preferences *could* have preceded decoupled representation. It provides what some call a 'how possibly' scenario (e.g. O'Connor 2019) about the evolution of cognition that is relevant to the eventual evolution of folk-psychological capacities. Not only that, a type of agent that has preferences (which I have argued can be useful on their own) but lacks decoupled representations is poised to take advantage of decoupled representations, which enable its selections to be conditional on information that isn't locally occurrent. Such representations may be constructed over time and integrating multiple cues. Decoupled representation could, that is, have started out as the 'slave of the preferences' (following Hume 1975, II.3.3, p 415).

5) Conclusion

I've defended the thesis that there is a scenario that is both conceptually coherent and consistent with neuroscience and recorded evolution in which preferences – a kind of desire-like state – evolved *before* belief-like states. That's a significant thesis, but is also considerably weaker than the claim that they did in fact evolve first, let alone that they *necessarily* came first. The other relevant possibilities are, as noted above, that proto-beliefs evolved first, and that the two in some sense evolved together.

Of those two, the proto-belief first scenario does not seem particularly promising. The environmental complexity thesis explains cognition as contributing to 'dealing with' (environmental) complexity, which is to say that cognition pays its way by making a difference to what an organism does in what circumstances. Any investment in world-modelling has to have some route to influencing behaviour. There aren't computer science models of successful agents with world modelling capacities but whose action selections are insensitive to selection as there are of the opposite (model-free reinforcement learning). These aren't decisive considerations, and any proposal in which proto-beliefs evolved first would have to be evaluated on its merits.

What about scenarios in which the two evolved together? I referred above to Millikan's notion of "pushmi-pullyu representations" that combine indicative (i.e. discrimination) and imperative (i.e. motivating action) functionality (Millikan 1995). Such a representation has elements of proto-belief and simple motivation. But we need to be careful. An organism with a collection of 'pushmi-pullyu' representations will likely face the problem of what to do when more than one of these representations call for making use of the same bodily degrees of freedom, for example if one is calling for the limbs to be used to approach something, and another to sit down and scratch an itch. In that case some kind of selection will be needed, which needn't inevitably involve preferences, but which will be more efficient if it does. And we've already seen how networks of final common paths (where 'pushmi-pullyu' representations compete for use of the body) are both a conceptually convenient place for the exercise of preferences, and connected with their actual evolution in the history of nervous systems. The case of multiple 'pushmi-pullyu' representations is broadly analogous to the situation of an agent whose psychology is organised according to Gibsonian 'affordances' (sensorimotor registrations of possibilities for action of a body in a situation). Multiple affordances also sometimes call for incompatible uses of the body, leading to a problem of affordance selection, to which preferences can also be a solution (Spurrett 2018). So scenarios in which they evolved together still encourage the thought that elaboration of preferences could drive the process.

References

Barron, A. B., Søvik, E., & Cornish, J. L. [2010]: 'The Roles of Dopamine and Related Compounds in Reward-seeking Behaviour across Animal Phyla', *Frontiers in Behavioural Neuroscience*, **4**.

Brandom, R. (2009) How Analytic Philosophy Has Failed Cognitive Science, in: *Reason in Philosophy*, Harvard University Press.

Gallistel, C.R. (1980) *The Organisation of Action: A New Synthesis*. Lawrence Erlbaum.

Glimcher, P.W., & Sparks, D.L. (1992) Movement selection in advance of action in the superior colliculus. *Nature*, 355: 542-545.

Glimcher , P.W. (2003), 'The neurobiology of visual-saccadic decision making', *Annual Review of Neuroscience*, vol. 26, pp. 133-179.

Glimcher, P. (2011) *Foundations of Neuroeconomic Analysis*, OUP.

Godfrey-Smith, P. (1996) *Complexity and the Function of Mind in Nature*, CUP.

Hejnol, A. & Martindale, M.Q. (2008) Acoel development indicates the independent evolution of the bilaterian mouth and anus. *Nature*, 456: 382-386.

Hume, D. (1975) *A Treatise of Human Nature*, ed. L.A. Selby-Bigge, rev. P.H. Nidditch, Oxford: Clarendon Press.

Millikan, R.G. (1995) Pushmi-pullyu representations. *Philosophical Perspectives*, 9.

Millikan, R.G. (2017) *Beyond Concepts*, OUP.

O'Connor, C. (2019) *The Origins of Unfairness: Social Categories and Cultural Evolution*. Oxford University Press.

Perry, C.J., Barron, A.B., & Cheng, K. (2013) Invertebrate Learning and cognition: relating phenomena to neural substrate. *Wiley Interdisciplinary Reviews: Cognitive Science*, 4(3): 561-582.

Platt, M. L., & Glimcher, P. (1999) Neural correlates of decision variables in parietal cortex. *Nature*, 400: 233-8.

Schultz, W., Dayan, P. & Montague, R. (1997) A Neural Substrate of Prediction and Reward. *Science*, 275.

Sherrington, C.S. (1906), *The integrative action of the nervous systems*. Yale.

Spurrett, D. (2015) The Natural History of Desire, *South African Journal of Philosophy,* 34(3): 304-313..

Spurrett, D. (2018) Affording Affordances. *Teorema,* 37(3): 187-202.

Spurrett, D. (2021a) The Descent of Preferences. *British Journal for the Philosophy of Science,* 72(2): 485-510.

Spurrett, D. (2021b) Time and the Decider. *Behavioral and Brain Sciences,* 44: e133.

Sterelny, K. (2003). *Thought in a Hostile World*, Blackwell.

Sutton, R.S., & Barto, A.G. (1998, 2018 2nd edition) *Reinforcement Learning*, MIT. (Also 2018 2nd edition.)