Edinburgh Research Explorer

Radical predictive processing

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
The Southern Journal of Philosophy

Radical Predictive Processing

Andy Clark

Abstract

Recent work in computational and cognitive neuroscience depicts the brain as an ever-active prediction machine: an inner engine continuously striving to anticipate the incoming sensory barrage. I briefly introduce this class of models before contrasting two ways of understanding the implied vision of mind. One way (Conservative Predictive Processing) depicts the predictive mind as an insulated inner arena populated by representations so rich and reconstructive as to enable the organism to 'throw away the world'. The other (Radical Predictive Processing) stresses the use of fast and frugal, action-involving solutions of the kind highlighted by much work in robotics and embodied cognition. But it goes further, by showing how predictive schemes can combine frugal and more knowledge-intensive strategies, switching between them fluently and continuously as task and context dictate. I end by exploring some parallels with work in enactivism, and by noting a certain ambivalence concerning internal representations and their role in the predictive mind.

1. The Backdrop: Brains, Bodies, and World

Embodied cognitive science is no longer in its infancy. It is now over twenty years since the first waves of work on mobile, autonomous robots combined with increasing dissatisfaction with the prevailing 'classical computationalist' framework to suggest a new approach to the sciences of mind – an approach that recognized the many ways in which bodily form, action, and the canny use of environmental structure could simplify and transform the tasks facing the biological brain (for reviews, see Clark (1997), Pfeifer and Bongard (2006)). Since then, an explosion of work on embodied cognition has revealed the many ways in which biological intelligence relies upon the complex interplay between brain, body, and world. Neural representations, this work has suggested, are not action-neutral mirrors of the world. Instead they are in some deep sense 'action-oriented' (Clark (1997), Engel et al (2013)). They are geared to promoting successful, fast, fluent actions and engagements for a creature with specific needs and bodily form. Such

representations will be as minimal as possible, neither encoding nor processing information in costly ways when simpler routines, combined with world-exploiting actions, can do the job.

One of the abiding questions, which this research left open, was how best to understand the biological brain – the organ at the center of many of these webs of embodied and situated action. Much of the early impact hereabouts was negative – we should not understand the brain as a kind of micro-managing controller but rather as a lazy orchestrator, making the most of whatever body and world have to offer. In recent years, however, work on the 'predictive brain' has begun to suggest a richer, more positive story. Brains like ours, this work suggests, are multilevel prediction engines forever trying to anticipate the incoming flows of sensory information.

Such accounts have roots in familiar (but importantly incomplete) models of 'passive perception'. The brain, on these accounts, combines prior knowledge or expectations (including knowledge about the present context) with the incoming sensory evidence to yield a percept that reflects its best available 'hypothesis' concerning the most probable state of the world. This view of 'perception-as-inference' originates with Helmholz (1860) and has had many more recent champions, including Neisser (1967) and Gregory (1980). It is only in recent years, however, that these broad visions have been given effective computational flesh, shown to be (roughly speaking) neurally plausible, and seen to converge with compelling bodies of work in psychophysics and cognitive psychology showing that much of perception conforms to optimal (Bayesian) ways of combining sensory evidence with prior knowledge (see Bubic et al (2010), Kveraga et al (2007), Friston (2005) (2009), Hohwy (2013), Clark (2013)). Huang and Rao (2011), Bastos et al (2012), Knill and Pouget (2004), Yuille and Kersten (2006), Weiss et al (2002)).

A promising account that combines all these elements is work on 'hierarchical predictive coding' or 'predictive processing' (PP) models of perception (Friston (2005), (2009), Hohwy (2013), Clark (2013)). These approaches make extensive use of probabilistic 'internal models', but they locate them within a framework that is very different from the passive, input-output dominated approaches that characterized much previous work in cognitive and computational neuroscience. Recently, such accounts have been extended to make greater contact with issues concerning action, agency and choice, and with the nature and origins of various disturbances to subjective experience (Friston, Adams et al (2012), Fletcher and Frith (2009)). Such accounts are highly compatible with key insights from work on embodied cognition. The stress on neural prediction binds perception, cognition, and the control of action together in a single overarching framework in which they

conspire (at varying timescales) to reduce 'prediction errors' – the discrepancies between what is currently predicted and the evolving shape of the sensory barrage. In the next section, I put some flesh on the PP story, before turning (in sections 3 to 7) to the 'fit' with embodied, extended, and enactivist approaches to understanding the mind.

2. Predictive Processing

Two core features of the predictive processing (PP) account stand out as especially relevant to our current concerns. First, perception involves the use of a unified body of acquired knowledge (a multi-level 'generative model') to predict the incoming sensory barrage. Second, the use of that knowledge is subject to a constant kind of second-order assessment (known as 'precision estimation') that determines the weighting assigned to specific predictions at all levels of processing, and to different aspects of the incoming sensory signal. These weightings reflect the varying reliability, in context, of differing aspects of the generative model and of the sensory inputs currently available.

This basic schema is best introduced by work on simple sensory perception. A familiar view depicts perception as essentially a process of 'bottom-up' feature detection. Visual cortex, to take the most-studied example, is thus "traditionally viewed as a hierarchy of neural feature detectors, with neural population responses being driven by bottom-up stimulus features" (Egner et al 2010 p. 16601). This is a view of the perceiving brain as passive and stimulus-driven, taking energetic inputs from the senses and turning them into a coherent percept by a kind of step-wise build-up moving from the simplest features to the more complex. From pixel intensities up to lines and edges and on to complex meaningful shapes (like teacups), accumulating structure and complexity along the way in a kind of Lego-block fashion.

PP performs a kind of 'Bayesian flip' upon this standard (passive, feedforward) image of sensory processing. Instead of trying to build a model of what's out there on the basis of a panoply of low-level sensory cues, these models aim, in effect, to predict the current suite of low-level sensory cues from their best models of what's likely to be out there (for this formulation, see Hohwy (2007)). In these models (see Rao and Ballard (1999), Lee and Mumford (2003), Friston (2005) (2009)) percepts emerge via a recurrent cascade of 'top-down' predictions that involve (mostly sub-personal) expectations, spanning multiple spatial and temporal scales. These expectations are probabilistic and concern the *present* nature and state of the world as presented via the driving sensory signal. It is this key role for downwards-flowing prediction that allows these systems to cope with noisy and ambiguous

sensory inputs. The downwards predictions reflect what the system expects given what it already 'knows' about the world and about the current context. These predictions are combined with incoming sensory data to arrive at better guesses about the signal source. The driving sensory signal is here compared, at multiple levels, to a cascade of downwards predictions and mismatches send forward 'prediction error signals' that nuance or alter the prediction until a match is found and the sensory data is accommodated. This process runs concurrently and continuously across multiple levels of a processing hierarchy. As a very simple example, consider the figure below.

Reading left to right, the predictive brain meets the raw sensory stimulations from the central inscription with a strong expectation of the numeral 13. Reading top to bottom, the raw sensory stimulations are met with a different prediction – the letter B. In Bayesian terms, in the context of reading the 12, the 13 hypothesis makes the raw visual data most probable, while in the context of reading the A, the B hypothesis makes the raw visual data most probable (for more on this example, see Lupyan and Clark (In Press)).

How is the knowledge used to drive the predictions acquired in the first place? Some may be innate, inherent in the basic shape of the neural economy. But a major attraction of the multi-layer predictive processing approach is that it lends itself very naturally to a form of unsupervised (or, if you prefer, self-supervised) learning in which the attempt to predict powers the learning that makes better predictions possible. In these models, each 'higher' neural population is constantly trying to predict the rolling (ongoing) state of the neural population below it. During learning, that prediction is compared to the state that actually occurs and the neural population generating the prediction gently (automatically) self- tweaked,

via standard gradient descent or other means[1], so as to progressively reduce the error. The prediction task is thus a kind of 'bootstrap heaven'. To predict the next word in a sentence, it helps to know stuff about grammar (and lots more too). But one way to learn a surprising amount about grammar is just to look for the best ways to predict the next words in sentences. So you can use the prediction task to bootstrap your way to the grammar, that you then use in the prediction task in future.

Such learning is nicely demonstrated in Rao and Ballard's early (1999) model of predictive coding in the visual cortex. Rao and Ballard implemented a multilayer neural network whose input was samples (image patches) from pictures of natural scenes. The network had no concrete guiding task or purpose. Instead, visual signals were processed via a hierarchical system in which each level tried (in the way just sketched) to predict the activity at the level below it using recurrent (feedback) connections. If the feedback successfully predicted the lower level activity, no further action needed to ensue. Failures to predict enabled tuning and revision of the model (initially, just a random set of connection weights) generating the predictions thus slowly delivering knowledge of the regularities governing the domain. After exposure to thousands of image patches, the Rao and Ballard network had learnt to use responses in the first layer to extract features such as oriented edges and bars, while the second layer captured combinations of such features corresponding to patterns involving larger spatial configurations. Using the predictive coding strategy, and given only the statistical properties of the signals derived from the natural images, the network was thus able to induce a simple multi-layered model of the structure of the data source (images of natural scenes).

Now imagine a much richer, more multi-layered, version of the same procedure. Here, top-level predictions concern matters that are increasingly discrete and abstract (typically, more temporally extended and spatially diffuse) such as that the Memphis Tigers are playing baseball, now, in the stadium where you are sitting. Lower level predictions track states whose spatial or temporal signatures are continuous, local, and more fine-grained, such as the texture of a surface (that of a plastic beer glass perhaps). Between these extremes, interpose as many levels and kinds of prediction as you can imagine. The task of perception, given such a multi-layered prediction machine, is to match the incoming multi-modal sensory signal with apt top-down predictions at every level. That means finding the best set of predictions given the data, the context, and prior knowledge. The role of ongoing sensory input here is to constrain the brain's chronic ongoing guessing.

An important feature of the predictive processing account is that the weight that is given to the driving sensory signal (hence the value of prediction errors concerning

that signal) can be varied according to its degree of certainty or uncertainty. This is achieved by altering the gain (the 'volume' to use the standard auditory analogy) on the error-units accordingly. The effect of this is to allow the brain to vary the balance between sensory inputs and prior expectations at different levels (see Friston (2009) p. 299). This means that the weighting of sensory prediction errors (hence the relative influence of sensory inputs and prior expectations) at any level of processing within the whole hierarchical cascade may itself be flexibly modulated. This is sometimes described as optimizing "the relative precision of empirical (top-down) priors and (bottom-up) sensory evidence" (Friston (2009) p. 299). This feature will prove pivotal when we start to explore the links with low-cost (more 'embodied') solutions below.

What (finally) about action: the lynchpin of embodied cognitive science? An intriguing speculation is that:

> " the best ways of interpreting incoming information via perception, are deeply the same as the best ways of controlling outgoing information via motor action….so the notion that there are a few specifiable computational principles governing neural function seems plausible. " Eliasmith (2007) p. 380

Just such a parallelism emerges in recent work on 'active inference' (see e.g. Friston (2009), Friston, Daunizeau et al (2010)). Action, it is suggested, involves a kind of self-fulfilling prophecy in which neural circuitry first predicts the sensory (especially proprioceptive) consequences of the selected action. Those consequences do not immediately obtain, however, so prediction error ensues. The resulting cascade of prediction error is then quashed – ultimately at the level of spinal reflexes (Adams et al (2013)) – by moving the bodily plant to bring the action about. Predictions of the sensory consequences of actions thus play the role more normally assigned to motor instructions, in what amounts to a contemporary version of the 'ideomotor' theory of James (1890) and Lotze (1852). The difference between motor and visual cortex, on this account, lies in what kind of thing (for example, a trajectory of motion) is predicted rather than how it is predicted.

Putting all this together reveals perception and action as locked in a continuous circular causal embrace. Perceptual hypotheses here inform actions that elicit, from the world, the very streams of multi-modal stimulation that they predict (see Friston, Adams et al (2012)). In this way we continuously bring into being the flows of salient sensory stimulation that our brains predict. This, as we shall later see in more detail, builds a robust bridge between PP and some of the core ideas behind 'enactivism' – for example, the idea (Varela, Thompson and Rosch (1991))) that we bring forth our worlds through action. Dynamically speaking, the idea is

that the whole embodied, active system self-organizes around the organismically-computable quantity 'prediction error'. This delivers a multi-level, multi-area, grip on the evolving sensory barrage – a grip spanning multiple spatial and temporal scales. That grip simultaneously determines perception and action, and thus selects (enacts) the ongoing stream of sensory bombardment itself. The 'generative model' that here issues sensory predictions is thus nothing but that multi-level, multi-area, multi-scale, body-and-action involving grip on the unfolding sensory stream. To achieve that grip, if PP is on track, is to know the structured and meaningful world that we encounter in experience and action.

We now rehearse some core themes from the literature on embodied cognition before returning (section 5) to the 'fit' with work on the predictive brain.

3. Productive Laziness

A recurrent theme in work on the embodied, environmentally situated mind has been the value of 'productive laziness'. I owe this phrase to Aaron Sloman (personal communication), but the general idea goes back at least to Herbert Simon's (1956) explorations of economical but effective strategies and heuristics: problem-solving recipes that are not (in any absolute sense) optimal or guaranteed to work under all conditions, but that are 'good enough' to meet a need while respecting limitations of time and processing power. For example, rather than attempt a full examination of reviews and menus for every restaurant within a 5 mile radius, we might very well choose one that a trustworthy friend mentioned yesterday instead. We do so reasonably confident that it will be good enough, and thereby save the temporal and energetic costs of taking further information into account.

Fast, heuristically-governed strategies for reasoning are, however, only one part of the rich mosaic of 'productive laziness'. Another part (the focus of much of my own previous work in this area – see Clark (1997) (2008)) involves what might be thought of as ecologically efficient uses of sensing, and the distribution of labor between brain, body, and world. For example, there are circumstances (as Sloman (2013) points out) in which the best way to get through an open door is to rely upon a simple servo-control, or bump-and-swerve, mechanism. Or consider the task of two-legged locomotion. Some bipedal robots (Honda's flagship 'Asimo' is perhaps the best-known example) walk by means of very precise, and energy-intensive, joint-angle control systems. Biological walking agents, by contrast, make maximal use of the mass properties and bio-mechanical couplings present in the overall musculoskeletal system and walking apparatus itself. Nature's own bi-pedal walkers thus make extensive use of so-called "passive dynamics", the kinematics and organization inhering in the physical device alone (McGeer 1990). It is such

passive dynamics that enable some quite simple toys, that have no onboard power source, to stroll fluently down a gentle incline. Such toys have minimal actuation and no control system. Their walking is a consequence not of complex joint movement planning and actuating, but of their basic morphology (the shape of the body, the distribution of linkages and weights of components, etc). Locomotion, as nicely noted by Collins et al (2001 p.608), is thus "a natural motion of legged mechanisms, just as swinging is a natural motion of pendulums".

Passive walkers (and their elegant powered counterparts – see Collins et al (2001)) conform to what Pfeifer and Bongard (2007) describe as a 'Principle of Ecological Balance'. This principle states:

> "first…that given a certain task environment there has to be a match between the complexities of the agent's sensory, motor, and neural systems…second….that there is a certain balance or task-distribution between morphology, materials, control, and environment" Pfeifer and Bongard (2007) p 123

This principle reflects one of the big lessons of contemporary robotics, which is that the co-evolution of morphology (which can include sensor placement, body-plan, and even the choice of basic building materials etc) and control yields a golden opportunity to spread the problem-solving load between brain, body and world. Robotics thus rediscovers many ideas explicit in the continuing tradition of J.J. Gibson and of 'ecological psychology' see Gibson (1979), Turvey and Carello (1986)). William Warren, commenting on a quote from Gibson (1979), suggests that:

> "biology capitalizes on the regularities of the entire system as a means of ordering behavior. Specifically, the structure and physics of the environment, the biomechanics of the body, perceptual information about the state of the agent-environment system, and the demands of the task all serve to constrain the behavioral outcome" Warren (2006) p.358

Another Gibsonian theme concerns the role of sensing in action. According to a familiar (more classical) vision, the role of sensing is to get as much information into the system as is needed to solve the problem. For example, a planning agent might scan the environment so as to build up a problem-sufficient model of what's out there and where it is located, at which point the reasoning engine can effectively throw away the world and operate instead upon the inner model, planning and then executing a response (perhaps checking now and then during execution to be sure that nothing has changed). Alternative approaches (see e.g. Gibson (1979), Lee and Reddish (1981), Beer (2000) (2003), Chemero (2009))

depict sensing as a channel productively coupling agent and environment, sidestepping where possible the need to convert world-originating signals into a persisting inner model of the external scene.

Thus consider the 'outfielder's problem': running to catch a fly ball in baseball. Giving perception its standard role, we might have assumed that the job of the visual system is to transduce information about the current position of the ball so as to allow a distinct 'reasoning system' to project its future trajectory. Nature, however, looks to have found a more elegant and efficient solution. The solution, a version of which was first proposed in Chapman (1968), involves running in a way that seems to keep the ball moving at a constant speed through the visual field. As long as the fielder's own movements cancel any apparent changes in the ball's optical acceleration, she will end up in the location where the ball is going to hit the ground. This solution, Optical Acceleration Cancellation (OAC), explains why fielders, when asked to stand still and simply predict where the ball will land, typically do rather badly. They are unable to predict the landing spot because OAC is a strategy that works by means of moment-by-moment self-corrections that crucially involve the agent's own movements. The suggestion that we rely on such a strategy is also confirmed by some interesting virtual reality experiments in which the ball's trajectory is suddenly altered in flight, in ways that could not happen in the real world – see Fink, Foo, and Warren (2009)). OAC is a nice case of fast, economical problem-solving. The canny use of data available in the optic flow enables the catcher to sidestep the need to deploy a rich inner model to calculate the forward trajectory of the ball[2].

Such strategies are suggestive (see also Maturana 1980) of a very different role for the perceptual coupling itself. Instead of using sensing to get enough information inside, past the visual bottleneck, so as to allow the reasoning system to 'throw away the world' and solve the problem wholly internally, they use the sensor as an open conduit allowing environmental magnitudes to exert a constant influence on behavior. Sensing is here depicted as the opening of a channel, with successful whole-system behavior emerging when activity in this channel is kept within a certain range. In such cases, as Randall Beer puts it:

> "the focus shifts from accurately representing an environment to continuously engaging that environment with a body so as to stabilize appropriate co-ordinated patterns of behavior" (Beer 2000, p.97).

Finally, embodied agents are also able to act on their worlds in ways that actively generate cognitively and computationally potent time-locked patterns of sensory stimulation. In human infants, grasping, poking, pulling, sucking and shoving

creates a rich flow of time-locked *multi-modal* sensory stimulation. Such multi-modal input streams have been shown (Lungarella and Sporns (2005)) to aid category learning and concept formation. The key to such capabilities is infant's capacity to maintain co-ordinated sensorimotor engagement with its environment. Self-generated motor activity, such work suggests, acts as a "complement to neural information-processing" (Lungarella and Sporns (2005) p.25). One major strand of work in robotics and artificial life thus stresses the importance of the distribution of the problem-solving load across the brain, the active body, and the manipulable structures of the local environment. This distribution allows the productively lazy brain to do as little as possible while still solving (or rather, while the whole embodied, environmentally-located system solves) the problem.

4. Interactive Flow

Work on embodied cognition also calls into question the idea that there is a sequential flow of processing whose stages neatly correspond to perceiving, thinking, and acting. When we engage the world in daily behavior, we often don't do it by first passively taking in lots of information, then making a full plan, then implementing the plan courtesy of some sequence of motor commands. Instead, sensing, thinking, and acting conspire, overlap, and start to merge together as whole perceptuo-motor systems engage the world.

Examples of such merging and interweaving include work on interactive vision (Churchland et al (1994)), dynamic field theory (Thelen et al (2001)), and 'deictic pointers' (Ballard et al (1997)) – for some reviews, see Clark (1997) (2008). For present purposes, two illustrative examples will suffice. The first and simplest (Mataric (1990) (1992)) is from the early days of situated robotics. As a robot moves around a simple maze, it detects landmarks registered as a combination of sensory input and current motion. A narrow corridor thus registers as a combination of forward motion and short lateral distance readings from sonar sensors. Later, if the robot is required to find its way back to a remembered location, it activates an interlinked body of such combined sensory and motor readings. The stored 'map' of the environment is thus immediately fit to act as a recipe for action, since the motor signals are part of the stored knowledge. In other words, the relation between two locations is directly encoded as the set of motor signals that moved the robot from one to the other. The inner map is thus *itself* the recipe for the necessary motor actions. By contrast, a more classical approach might first generate a more objective map which would then need to be *reasoned over* in order to plan the route. The Mataric robot thus deploys what Clark (1997) describes as 'action-oriented representations' - representations that describe the world by depicting it in terms of possible actions[3].

As a second illustration, consider the task studied by Ballard et al (1997). In this task, a subject is given a model pattern of colored blocks and asked to copy the pattern by moving similar blocks, one at a time, from a reserve area to a new workspace. The task is performed by drag and drop using a mouse and monitor, and as you perform, eye tracker technology monitors exactly where and when you are looking as you tackle the problem. What subjects did *not* do, Ballard et al discovered, was to look at the target, decide on the color and position of the next block to be added, then execute their mini- plan by moving a block from the reserve area. Instead, repeated rapid saccades to the model were used during the performance of the task - many more saccades than you might expect. For example, the model is consulted *both before and after* picking up a block, suggesting that when glancing at the model, the subject stores only one piece of information: either the color or the position of the next block to be copied, but not both. Even when repeated saccades are made to the same site, very minimal information looked to be retained. Instead, repeated fixations seem to be providing specific items of information 'just in time' for use[4]. Repeated saccades to the physical model thus allowed the subjects to deploy what Ballard et al dub 'minimal memory strategies' to solve the problem. The idea is that the brain creates its programs so as to minimize the amount of working memory that is required, and that eye motions are here recruited to place a new piece of information into memory. By altering the task demands, Ballard et al were also able to systematically alter the particular mixes of biological memory and active, embodied retrieval recruited to solve different versions of the problem, concluding that in this task "eye movements, head movements, and memory load trade off against each other in a flexible way" (op cit p.732). This is another now-familiar (but still important) lesson from embodied cognition. Eye movements here allow the subject to use the external world itself, where appropriate, as a kind of storage buffer (for lots more on this kind of strategy, see Wilson (2004), Clark (2008)).

Putting all this together already suggests a much more integrated model of perception, cognition and action. Perception is here tangled up with possibilities for action and is continuously influenced by cognitive, contextual, and motor factors. This is, indeed, exactly the picture suggested by Pfeifer et al's (2007) notion of the 'self-structuring of information flows'. Action serves to deliver fragments of information 'just in time', and that information guides action, in an ongoing circular causal embrace. Perception thus construed need not yield a rich, detailed and action-neutral inner model awaiting the services of 'central cognition' to deduce appropriate actions. In fact, these distinctions (between perception, cognition and action) now seem to obscure, rather than illuminate, the true flow. In a certain sense, the brain is revealed not as (primarily) an engine of reason or quiet deliberation, but as an organ for the environmentally situated control of

action. Cheap, fast, world-exploiting action, rather than the pursuit of truth, optimality, or deductive inference, is now the key organizing principle.

Embodied, situated agents, all this suggests, are masters of 'soft assembly', building, dissolving, and rebuilding temporary ensembles that exploit whatever is available, creating shifting problem-solving wholes that effortlessly span brain, body, and world.

5. Frugal Prediction Engines

The PP model, with its relentless stress on the role of top-down, generative-model based prediction may seem an unlikely choice as the theory of the brain best suited to partner the kinds of alternative model of mind just sketched. And there is, indeed, a way of understanding PP that would be quite inimical to the various "lessons from embodiment" mentioned above. Such an understanding (let's call it 'conservative predictive processing') would depict our cognitive contact with the world as rooted in a kind of neuronally-encoded rich inner recapitulation of an observer-independent reality. This would offer a prediction-based version of what Anderson (2014) describes as the 'reconstructive' approach to perception. Such an approach uses sensing to build up (in the classical feedforward version) or activate (in the Conservative Predictive Processing model) an inner model that recapitulates the structure and richness of the real-world. That rich inner model is then able to stand-in for the external world for the purposes of planning, reasoning, and the guidance of action. The upshot is that the model-rich cognizer is able to 'throw away the world' and select her actions and responses by manipulating the inner model instead.

I shall not further rehearse the many reasons why we should be suspicious of such approaches (for such assessments, see Churchland et al (1994), Clark (1997), Pfeifer and Bongard (xx)). But we have already seen one practical (in fact, flagship) example of an alternative, in the form of the OAC approach to solving the outfielder's problem. Here, sensing delivers an action-based *grip* upon the world, rather than a rich reconstruction apt for detached reasoning. Such a grip may intrinsically involve organismic action, as when the outfielder runs so as to keep the image of the ball stationary on the retina. By thus acting she ensures that she will be in a position to catch the ball when it descends towards the pitch. In such cases, behavioral success is not the outcome of reasoning defined over a kind of inner replica of the external world. Rather, it is the outcome of perception/action cycles that operate by keeping sensory stimulations within certain bounds. This is the same kind of strategy celebrated by other work in ecological psychology showing, for example, how diving seabirds (gannets) predict time-to-impact according to the relative rate of expansion of the image in the optic array - see Lee

and Reddish (1981), and discussion in Tresilian (1999). Such strategies are deeply *non-reconstructive*.

PP, however, is in no way committed to the conservative (richly reconstructive) reading that would render it incompatible with such solutions. On the contrary, one of the fundamental principles of PP renders it opposed to that reading. This is because the goodness of a predictive model is determined by accuracy minus complexity[5]. Accuracy here names the ability to predict the task-salient sensory flux, while complexity varies according to the number of parameters in the model. Thus Fitzgerald et al note that:

> "Bayes optimal agents seek both to maximize the accuracy of their predictions and to minimize the complexity of the models they use to generate those predictions" Fitzgerald et al (2014, p.1)

Minimizing complexity requires reducing computational costs as far as possible, consistent with performing the task at hand. Formally, this can be achieved by incorporating a complexity-penalizing factor - sometimes called an Occam factor, after the 13th century philosopher William of Occam who famously cautioned us not to "multiply entities beyond necessity". Overall 'model evidence' is then a kind of composite quantity reflecting a delicate (and context variable) accuracy/complexity trade-off. PP systems are thus driven to learn and deploy the least complex (fewest parameters) solutions that will serve our needs. Crucially, PP accounts are also diametrically opposed to 'passive perceiver' stories. Instead, as we saw in section 2, the prediction machinery simultaneously drives perception and action so as to reduce salient prediction error. That means action gets called upon to serve simplicity too.

To see how this might all work in practice, recall the role of 'precision weighting' in sculpting patterns of causal influence among neuronal populations. Precision weighting alters the 'gain' on specific prediction error units, and thus provides a means of systematically varying the relative influence of different neural populations. The most familiar role of such manipulations is to vary the balance of influence between bottom-up sensory information and top-down model-based expectation. But another important role is to implement fluid and flexible forms of large-scale 'gating' among neural populations. This works because very low-precision prediction errors will have little or no influence upon ongoing processing, and will fail to recruit or nuance higher-level representations. Altering the distribution of precision weightings thus amounts to altering patterns of 'effective neuronal connectivity' – in other words, altering the 'simplest circuit diagram' (Aertsen and Preissl, 1991) underlying current processing. This delivers an inner processing economy that is pervasively context-sensitive.

Now let's return to the outfielder's problem described earlier. Here too, already-active neural predictions and simple, rapidly-processed perceptual cues must work together (if PP is correct) to determine a pattern of precision-weightings for different prediction error signals. This creates a transient web of effective connectivity (a temporary distributed circuit) and, within that circuit, it sets the balance between top-down and bottom-up modes of influence. In the case at hand, however, efficiency demands selecting a circuit in which sensing plays the non-reconstructive role described above. The temporary task of visual sensing, in this context, becomes that of cancelling the optical acceleration of the fly ball. That means giving high weighting to the prediction errors associated with cancelling the vertical acceleration of the ball's optical projection, and (to put it bluntly) not caring very much about anything else. Apt precision weightings thus select a pre-learnt, fast, low-cost strategy for solving the problem. Contextually recruited patterns of precision weighting thus accomplish a form of set-selection or strategy switching[6]. This assumes that slower processes of learning and adaptive plasticity have already sculpted patterns of neural connectivity in ways that make the low-cost strategy available. But this is unproblematic. It can be motivated in general terms by the drive towards minimizing complexity (which is indistinguishable, under plausible constraints, from the drive towards 'satisficing'). The required learning can thus be accomplished using prediction error minimization operating at many time-scales. Such processes range all the way from the slow learning of the child baseball player, to the faster online adaptation of the pro-player factoring in (during a match) changing specifics of the wind conditions and the play of opposing batters.

The upshot is a complex but rewarding picture in which bedrock processes of predictive learning slowly install models that include precision expectations allowing patterns of effective connectivity to be built and re-built 'on the fly'. This enables fast, knowledge-sparse modes of response to be recruited and nuanced according to current context. The resulting compatibility of 'productively lazy' and model-based approaches should come as no surprise. To see this, we need only reflect that the model or model fragment that underlies any given behaviour can be a simple, easily computed, heuristic (a simplified 'rule-of-thumb') just as easily as something with a more complex causal structure. Such low-cost models will in many cases rely upon action, exploiting patterns of circular causal commerce (between perceptual inputs and motor actions) to deliver task-relevant information 'just in time' for use. Dealing with a complex time-pressured world demands the use of many strategies, ranging from very simple heuristics to more complex structures of interacting approximations. Within PP, that diverse landscape exists within an overarching eco-system in which many strategies emerge, dissolve, and interact according to changing estimations of precision[7]. Such an eco-system is

continuously dynamically self-reconfiguring, constantly engaging actions that yield new inputs that recruit new strategies in a potent rolling cycle.

Notice, finally, that known external (e.g. environmental) operations are entirely apt for precision-based selection. For example, when performing the block-placing task (Ballard et al (1997)) described above, the brain must assign high precision to the predictions that underlie the various actions that are allowing us to 'use the world as it's own best model' while performing the task. Such world-engaging actions are determined, in turn, by the acquired estimation that reliable, salient (task-relevant) information is available at such-and-such a location and at such-and-such a time. The strategies thus selected are, just as Ballard et al suggested, minimal-internal-memory strategies whose success conditions require both organismic action and the co-operation of the external environment. The transient task-specific assemblies selected by varying estimations of precision are thus not restricted to the brain or even the organism[8]. Instead, transient neural coalitions select actions that may fold in the use of any amount of environmental structure and scaffolding[9]. Radical PP thus offers a systematic way of combining deep, model-based flexibility with the use of multiple, fast, efficient, environmentally-exploitative, routes to action and response[10].

6. Enacting Our Worlds

An immediate implication of this larger story is that there is a very real sense in which human agents help construct the very worlds they encounter in perception. This process of construction corresponds rather closely to the mysterious-sounding notion of 'enacting a world', at least as that notion appears in Varela et al (1991)[11].

Varela et al write that:

> "The overall concern of an enactive approach to perception is not to determine how some perceiver-independent world is to be recovered; it is, rather, to determine the common principles or lawful linkages between sensory and motor systems that explain how action can be perceptually-guided in a perceiver-dependent world" Varela et al (1991) p. 173

Such an approach to perception is prefigured, Varela et al report, in the work of Merleau-Ponty (1945/1962). There, Merleau-Ponty stresses the important degree to which perception itself is structured by human action. Thus we often think of perception as simply the source of information that is then used for the guidance of action. But expand the temporal window a little and it becomes clear that we

might equally well think of action as the selector of the perceptual stimulations themselves. In the words of Merleau-Ponty:

> "…since all the stimulations which the organism receives have in turn been possible only by its preceding movements which have culminated in exposing the receptor organ to external influences, one could also say that behavior is the first cause of all the stimulations" Merleau-Ponty (1945/1962) p.13

In a striking image, Merleau-Ponty then compares the active organism to a keyboard which moves itself around so as to offer different keys to the "in itself monotonous action of an external hammer" (op cit)[12]. The message that the world 'types onto the perceiver' is thus largely created (or so the image suggests) by the nature and action of the perceiver herself: the way she offers herself to the world. The upshot, according to Varela et al (1991, p. 174) is that "the organism and environment [are] bound together in reciprocal specification and selection".

This kind of relation is described by Varela et al as one of 'structural coupling' in which" the species brings forth and specifies its own domain of problems" (op cit p.198) and in that sense 'enacts' or brings forth (op cit p.205) its own world. In discussing these matters, Varela et al are also concerned to stress that the relevant histories of structural coupling may select what they describe as 'non-optimal' features, traits, and behaviors: ones that involve 'satisficing' (see Simon (1956)) in the sense introduced (without using that label) in section 3 above, i.e. settling for whatever 'good enough' solution or structure "has sufficient integrity to persist" (Varela et al (op cit) p.196). Radical PP has the resources to cash all these enactivist cheques, depicting the organism and the organism-salient world as bound together in a process of mutual specification in which the simplest approximations apt to support a history of viable interaction are the ones that are learnt, selected, and maintained.

The simplest way in which a PP-style organism might be said to actively construct its world is by sampling. Action here serves perception by moving the body and sense-organs around in ways that aim to 'serve up' predicted patterns of stimulation. In particular, they aim to serve up predicted sequences of high-reliability, task-relevant information. This is a very clear case, it seems to me, of the kind of 'active keyboard' effect imagined by Merleau-Ponty - the organism selectively moves its body and receptors to try to discover the very stimuli that it predicts. In this way, different organisms and individuals may selectively sample in ways that both actively construct and continuously confirm the existence of different 'worlds'. It is in this sense that, as Friston, Adams, and Montague (2012 p.

22) comment, our implicit and explicit models might be said to "create their own data".

Such a process repeats at several organizational scales. Thus we humans do not merely sample some natural environment. We also structure that environment by building material artifacts (from homes to highways), creating cultural practices and institutions, and trading in all manner of symbolic and notational props, aids, and scaffoldings. Some of our practices and institutions are also designed to train us to sample our human-built environment more effectively – examples would include sports practice, training in the use of specific tools and software, learning to speed-read, and many, many more. Finally, some of our technological infrastructure is now self-altering in ways that are designed to reduce the load on the predictive agent, learning from our past behaviors and searches so as to serve up the right options at the right time. In all these ways, and at all these interacting scales of space and time, we build and selectively sample the very worlds that - in iterated bouts of statistically-sensitive interaction - install the generative models that we bring to bear upon them.

The task of the generative model in all these settings is to capture the simplest approximations that will support the actions required to do the job – that means taking into account whatever work can be done by a creature's morphology, physical actions, and socio-technological surroundings. Such approximations are constrained to "provide the simplest (most parsimonious) explanations for sampled outcomes" (Friston, Adams, and Montague (2012) p.22). There is thus no conflict with work that stresses biological frugality, satisficing, or the ubiquity of simple but adequate solutions that make the most of brain, body, and world.

7. Revisiting Representation

There remains, however, at least one famously vexed issue upon which PP and the enactivist (at least if history is any guide) seem doomed to disagree. That is the issue of 'internal representation'. Thus Varela et al are explicit that, on the enactivist conception "cognition is no longer seen as problem solving on the basis of representations" (op cit p.205). PP, however, deals extensively in internal models – rich, frugal, and all points in-between - whose role is to control action by predicting complex plays of sensory data. This, the enactivist might fear, is where our promising story about neural processing breaks bad. Why not simply ditch the talk of inner models and internal representations and stay on the true path of enactivist virtue?

This issue requires a lot more discussion than I shall (perhaps mercifully) attempt here.[13] Nonetheless, the remaining distance between PP and the enactivist may not

be as great as that bald opposition suggests. We can begin by noticing that PP, although it openly trades in talk of inner models and representations, invokes representations that are probabilistic and action-oriented through and through. These are representations that are fundamentally in the business of serving up actions within the context of rolling sensorimotor cycles. Such representations aim to *engage* the world, rather than to depict it in some action-neutral fashion, and they are firmly rooted in the patterns of organism-environment interaction that served up the sensory stimulations that installed the probabilistic generative model.

The inner models thus constructed involve efficient encodings that deliver a grip upon a world of multiple competing affordances for action. The shape of that grip is well captured by Itay Shani who writes that:

> "Actual sensory systems are not concerned with truth and accuracy as such but rather, with action and the need to maintain the functional stability of the organisms in which they are embedded. They do not report, or register, what is where like an idealized scientific observer but, rather, help organisms to cope with changing conditions in their external, and internal (somatic), environments." Shani (2006) p. 90

This is exactly the role played, if PP is correct, by the internal multi-level probabilistic generative models that guide perception and action. It is also the role more broadly played by what Engel et al (2013) describe as 'dynamic directives' – dispositions towards action that are rooted in emergent ensembles that can include multiple neural and bodily structures.

What are the *contents* of the states governed by these multi-level action-oriented probabilistic generative models? The generative model issues predictions that estimate various identifiable worldly states (including states of the body. But it is also necessary, as we have repeatedly seen, to estimate the context-variable reliability (precision) of the neural estimations themselves. It is these precision-weighted estimates that drive action, and it is action that then samples the scene, delivering percepts that select more actions. Such looping complexities will make it hard (perhaps impossible) adequately to capture the contents or the cognitive roles of many key inner states and processes using the terms and vocabulary of ordinary daily speech. That vocabulary is 'designed' for communication, and (perhaps) for various forms of cognitive self-stimulation. The probabilistic generative model, by contrast, is designed to engage the world in rolling, uncertainty-modulated, cycles of perception and action. The representations thus constructed are:

> "Not actual re-presentations or duplicates of objects in the world but….incomplete, abstract code that makes predictions about the world and

revises its predictions on the basis of interaction with the world" Lauwereyns (2012) p.74

Within PP high-level states (of the generative model) target large-scale, increasingly invariant patterns in space and time. Such states help us to keep track of specific individuals, properties, and events despite large moment-by-moment variations in the stream of sensory stimulation. Unpacked via cascades of descending prediction, such higher-level states simultaneously inform both perception and action, locking them into continuous circular causal flows. Instead of simply describing 'how the world is', these models - even when considered at the 'higher' more abstract levels - are geared to engaging those aspects of the world that matter to us. They are delivering a grip on the *patterns that matter* for the *interactions that matter*.

Varela at al (1991) strongly reject appeals to 'internal representation'. But for them, this notion implies the 'action-neutral' capture of what they call a 'pregiven world'. Organism and world, they argue, are instead co-defined by a history of structural coupling: a kind of active 'fitting' of each to the other, rather than a passive 'mirroring'. Radical PP, I have tried to show, fully respects this intuition. It posits a hierarchical generative model that helps maintain the integrity and viability of a system by enabling it to minimize prediction errors and thus avoid compromising (possibly fatal) encounters with the environment. That distributed inner model is itself the result of self-organizing dynamics operating at multiple temporal scales, and it functions selectively to expose the agent to the patterns of stimulation that it predicts. The generative model thus functions – just as an enactivist might insist - to enable and maintain a structural coupling in which the viability of the organism is preserved.

8. Conclusions: Embodied, Embedded, Enactive Prediction Machines

Creatures like us are built, it seems, to be persistently active and productively lazy. We are built to maximize success while minimizing effort, both intellectual and physical. We do this, Radical PP suggests, by deploying webs of inner resources that are both prediction-based and fundamentally action-oriented. Such resources are not in the business of representing the world in some passive, descriptive manner. Instead, they are built to engage it in complex rolling cycles in which actions determine percepts that select actions, evoking and exploiting all manner of environmental structures and opportunities along the way.

The probabilistic, prediction-driven inner economy also selects, moment-by-moment, the very strategies that we deploy. Those strategies range from the quick and dirty to the slow and accurate, from those dominated by bottom-up sensory

flow to those more reliant upon top-down contextual modulation, and all points and admixtures in between. Such strategies readily include the use of available bodily actions and the exploitation of local environmental resources. The worry that predictive processing organizations might over-emphasize computationally expensive, representation-heavy strategies over other (quicker, dirtier, more 'embodied') ones is thus fully and satisfyingly resolved. The ever-active predictive brain stands revealed as a lazy brain - a brain vigilant for any opportunity to do less, while achieving more.

Acknowledgement

References

Adams, F. and Aizawa, K. (2001). The Bounds of Cognition. *Philosophical Psychology* 14:1: 43-64.

Adams RA, Shipp S, Friston KJ (2013) Predictions not commands: active inference in the motor system *Brain Struct Funct.* 218(3):611-43

Aertsen A, Preißl H. (1991). Dynamics of activityand connectivity in physiological neuronal networks. In *Non Linear Dynamics and Neuronal Networks*, ed. HG Schuster, pp. 281– 302. New York.

Anderson, M (2014) *After Phrenology* (MIT Press)

Anderson, M.L., Richardson, M. & Chemero, A. (2012). Eroding the boundaries of cognition: Implications of embodiment. *Topics in Cognitive Science,* 4(4): 717-30.

Ballard, D., Hayhoe, M., Pook, P., and Rao, R. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences,* 20, 4.

Bastos AM, Usrey WM, Adams RA, Mangun GR, Fries P, Friston KJ. (2012) Canonical microcircuits for predictive coding. *Neuron* 76:695-711.

Beer, R (2003) The Dynamics of Active Categorical Perception in an Evolved Model Agent *Adaptive Behavior* 11: 209-243

Beer, R (2000) Dynamical approaches to cognitive science. *Trends in Cognitive Sciences* 4(3):91-99

Bubic A, von Cramon DY and Schubotz RI (2010) Prediction, cognition and the brain. *Front. Hum. Neurosci.* 4:25: 1-15

Chapman S (1968) Catching a baseball. *American Journal of Physics*. 36:868–870.

Chemero A. (2009). *Radical Embodied Cognitive Science*. Cambridge, MA: MIT Press.

Churchland, P.S., Ramachandran, V., & Sejnowski, T. (1994). A Critique of Pure Vision. In C. Koch & J. Davis (Eds.), *Large-Scale Neuronal Theories of the Brain* . Cambridge, MA: MIT Press.23-61

Clark, A (1989) *Microcognition: Philosophy, Cognitive Science and Parallel Distributed Processing*, MIT Press/Bradford Books, Cambridge: MA)

Clark, A (1997) *Being There: Putting Brain, Body and World Together Again* (MIT Press, Camb. MA).

Clark, A (2008) *Supersizing the Mind: Action, Embodiment, and Cognitive Extension* (Oxford University Press, NY)

Clark, A (2013) Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognitive Science *Behavioral and Brain Sciences* 36: 3: p. 181-204

Clark, A and Chalmers, D. (1998). The Extended Mind. *Analysis* 58:1:7-19

Clark, A (In Press) *Surfing Uncertainty: Prediction, Action, and the Embodied Mind* (Oxford University Press)

Collins, S.H., Wisse, M., Ruina, A. (2001) A 3-D Passive Dynamic Walking Robot with Two Legs and Knees, *International Journal of Robotics Research*, 20 (7):607-615.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977) Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society series B*, 39:1-38.

Egner, T., Monti, J. M., Summerfield, C. (2010). Expectation and surprise determine neural population responses in the ventral visual stream. *Journal of Neuroscience,* 30(49): 16601-16608.

Eliasmith, C. (2007). How to build a brain: From function to implementation. *Synthese*. 153(3): 373-388

Fink, P. W., Foo, P. S., & Warren, W. H. (2009). Catching fly balls in virtual reality: A critical test of the outfielder problem. *Journal of Vision*, 9(13):14, 1–8

FitzGerald, T., Dolan, R., and Friston, K. (2014) Model Averaging, Optimal Inference, and Habit Formation. *Frontiers in Human Neuroscience* 8: 1–11. doi:10.3389/fnhum.2014.00457.

Fletcher, P and Frith, C (2009) Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. *Nature Reviews: Neuroscience* 10: 48-58

Friston K. (2005). A theory of cortical responses. *Philos Trans R Soc Lond B Biol*

*Sci.*29;360(1456):815-36.

Friston K. (2009). The free-energy principle: a rough guide to the brain? *Trends Cogn Sci.* 13: 293–301

Friston K, Adams RA, Perrinet L and Breakspear M (2012) Perceptions as hypotheses: saccades as experiments. *Front. Psychology* 3:151. doi: 10.3389/fpsyg.2012.00151

Friston K, Daunizeau J, Kilner J, Kiebel SJ. (2010).   Action and behavior: a free-energy formulation. *Biol Cybern.* 102(3):227-260

Friston K, Shiner T, FitzGerald T, Galea JM, Adams R, et al. (2012) Dopamine, Affordance and Active Inference. *PLoS Comput Biol* 8(1): e1002327. doi:10.1371/journal.pcbi.1002327

Friston, K., Adams, R., and Montague, R (2012) What Is Value-Accumulated Reward or Evidence? *Frontiers in Neurorobotics* 6: 11.
doi:10.3389/fnbot.2012.00011.

Froese, T. and Di Paolo, E. A. (2011). The enactive approach: Theoretical sketches from cell to society. *Pragmatics and Cognition*, 19, 1-36.

Gallagher, S, Hutto, D, Slaby, J & Cole, J (2013) The brain as part of an enactive system *Behavioral and Brain Sciences*, vol 36, no. 4, pp. 421-422.

Gibson, J.J (1979) *The Ecological Approach To Visual Perception* Boston, MA Houghton-Mifflin

Gregory, R. L. (1980). Perceptions as hypotheses. *Phil. Trans. R. Soc. Lond., Series B, Biological Sciences* 290(1038): 181-197.

Haruno, M., Wolpert, D. M., & Kawato, M. (2003). Hierarchical mosaic for movement generation. *International Congress Series*, 1250, 575–590.

Helmholtz, H. (1860/1962). *Handbuch der physiologischen optik* (Southall, J. P. C. (Ed.), English trans.),Vol. 3. New York: Dover.

Hohwy, J (2013) *The Predictive Mind* (Oxford University press, NY)

Huang, Y and Rao, R (2011) Predictive Coding *Wiley Interdisciplinary Reviews: Cognitive Science*, 2: 580–593

Hutto, D. D., & Myin, E. (2013). *Radicalizing Enactivism: Basic Minds without Content.*

James, W. (1890/1950). *The Principles of Psychology*, Vol. I, II. Cambridge, MA: Harvard University Press.

Knill, D., and Pouget, A. (2004) The Bayesian brain: the role of uncertainty in neural coding and computation *Trends in Neuroscience.* 27(12):712-9

Kveraga, K., Ghuman, A., and Bar, M. (2007) Top-down predictions in the cognitive brain. *Brain and Cognition* 65: 145-168

Lauwereyns, J (2012) *Brain and the Gaze: On the Active Boundaries of Vision.* Cambridge, MA: The MIT Press

Lee, D. and Reddish, P. (1981) Plummeting Gannets: A Paradigm of Ecological Optics *Nature* 293:293-294

Lee, T.S., and Mumford, D. (2003). Hierarchical Bayesian inference in the visual cortex. *Journal of Optical Society of America, A. .* 20(7): 1434-1448.

Lotze, H. (1852) *Medicinische Psychologie oder Physiologie der Seele.* Leipzig, Germany: Weidmannsche Buchhandlung.

Lungarella, M. and Sporns, O. (2005). Information self-structuring: key principles for learning and development. *Proceedings 2005 IEEE Intern. Conf. Development and Learning,* pp. 25-30.

Lupyan, G and Clark, A (In Press) Words and the World: Predictive coding and the language-perception- cognition interface. *Current Directions in Psychological Science*

Matarić, M (1990) Navigating With a Rat Brain: A Neurobiologically-Inspired Model for Robot Spatial Representation, in *Proceedings, From Animals to Animats: First International Conference on Simulation of Adaptive Behavior (SAB-90),* J-A. Meyer and S. Wilson, eds., MIT Press 169-175.

Matarić, M (1992) Integration of Representation Into Goal-Driven Behavior-Based Robots, in *IEEE Transactions on Robotics and Automation,* 8(3) 304-312.

Maturana, H (1980) Biology of Cognition. in Maturana, H, Humberto, R and Varela, F *Autopoiesis and Cognition* (Dordrecht: Reidel) p.2-62

Maturana, H., & Varela. F. (1980). *Autopoiesis and cognition: The realization of the living.* Boston: Reidel.

McGeer, T. (1990) Passive dynamic walking. *International Journal of Robotics Research.* 9(2): 68-82.

Menary, R. (ed) (2010) *The Extended Mind* (MIT Press, Camb. MA).

Merleau-Ponty, M. (1945/1962) *The phenomenology of perception.* Trans. Colin Smith. (London: Routledge and Kegan Paul)

Millikan, R.G. 1996. Pushmi-pullyu representations. In J. Tomberlin (ed.) *Philosophical Perspectives* IX, 185-200.

Neisser, U., (1967). *Cognitive Psychology.* Appleton-Century-Crofts, New York.

Noë, A. (2004) *Action in Perception.* Cambridge, MA: The MIT Press.

Noë, A (2010) *Out of Our Heads: Why You Are Not Your Brain, and Other Lessons from the Biology of Consciousness* Farrar, Straus and Giroux: NY

Pfeifer R. and Bongard J. (2006) *How the Body Shapes the Way We Think: A New View of Intelligence* (MIT Press, Camb. MA)

Pfeifer, R., Lungarella, M., Sporns, O., Kuniyoshi, Y.(2007) On the information theoretic implications of embodiment - principles and methods. *Lecture Notes in Computer Science (LNCS)* vol. 4850 (Springer, Heidelberg).

Ramsey W.M. (2007) *Representation Reconsidered.* Cambridge: Cambridge University

Rao, R and Ballard, D. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects, *Nature Neuroscience* 2, 1, 79

Rupert, R (2004) Challenges to the Hypothesis of Extended Cognition *Journal of Philosophy* 101:8: 389-428

Rupert, R (2009) *Cognitive Systems and the Extended Mind* (Oxford University Press, New York)

Shaffer, D. M., Krauchunas, S. M., Eddy, M., & McBeath, M. K. (2004). How dogs navigate to catch Frisbees. *Psychological Science, 15,* 437– 441.

Simon, H. A. (1956) Rational choice and the structure of the environment. *Psychological Review*, Vol. 63 No. 2, 129-138.

Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A (2002) Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 64: 583–639.

Thelen, E., Schöner, G., Scheier, C., & Smith, L. (2001). The dynamics of embodiment: A field theory of infant perseverative reaching. *Behavioral and Brain Sciences* 24, 1-33.

Thompson, E (2010) *Mind in Life: Biology, Phenomenology, and the Sciences of Mind* (Harvard University Press)

Tresilian, J. R. (1999) Visually timed action: time-out for 'tau'? *Trends in Cognitive Sciences,* 3, 301-310.

Turvey, M and Carello, , C (1986) The ecological approach to perceiving-acting: A pictorial essay. *Acta Psychologica* 63, 133-155

Varela, F., Thompson, E., and Rosch.,E (1991) *The Embodied Mind* (MIT Press, Camb. MA)

Warren, W (2006) The Dynamics of Action and Perception. Psychological Review 113:2:358-389
Weiss Y., Simoncelli E.P. and Adelson E.H. (2002) Motion Illusions as Optimal Percepts *Nature Neuroscience* 5: 6: 598 – 604

Wilson, R. A. (2004) *Boundaries of the Mind: The Individual in the Fragile Sciences--Cognition* (Cambridge

University Press, Cambridge, UK

Yuille and Kersten (2006) Vision as Bayesian inference: analysis by synthesis? *Trends in Cognitive Science*: 10: 7: 301-308

---

[1] A popular choice is the EM (Expectation-Maximization) algorithm described by Dempster et al (1977).

[2] There are related accounts of how dogs catch Frisbees, a rather more demanding task due to occasional dramatic fluctuations in the flight path (see Shaffer et al (2004)).

[3] Such representations bear some resemblance to what Gibson (1979) called "affordances," where these are the potentials of use and activity that the local environment offers to a specific kind of being: chairs afford sitting (to humans), and so on. See also Millikan's (1996) account of 'pushmipullyu' representations, that simultaneously describe how things are and prescribe how to act.

[4] To test this hypothesis, Ballard et al used a computer program to alter the color of a block while the subject was looking elsewhere. For most of these interventions, subjects did not notice the changes even for blocks and locations that had been visited many times before, or that were the focus of the current action.

[5] See Spiegelhalter et al (2002), Fitzgerald, Dolan, and Friston (2014), Hohwy (2013).

[6] Such an effect has been demonstrated in some simple simulations of cued reaching under the influence of changing tonic levels of dopamine firing – see Friston, Shiner et al (2012).

[7] This kind of approach, in which simple, efficient models are gated and enabled within the larger context of a hierarchical Bayesian system, is by no means unique to PP. It is present, for example, in the MOSAIC framework due to (Haruno, Wolpert, & Kawato, 2001). In general, this kind of strategy will be available wherever estimations of our own uncertainty are available to gate and nuance online response.

[8] See Anderson et al (2012), Anderson (2014).

[9] This also means that Radical PP is compatible with (though it does not in itself establish or lend additional support to) yet another major 'alternative model of mind' – the idea (Clark and Chalmers (1998), Clark (2008)) that bio-external structures and operations may sometimes form such integral parts of an agent's cognitive routines as to be best thought of as 'extending the mind. For a thorough rehearsal of the positive arguments, see Clark (2008). For critiques, see Rupert (2004) (2009), Adams and Aizawa (2008)). For a rich sampling of the ongoing debate, see the essays in Menary (2010).

¹⁰ For lots more on this 'embodied prediction engine' story, see Clark (In Press).

¹¹ There is now a large, and not altogether unified, literature on enaction. For our purposes, however, it will suffice to consider only the classic statement by Varela et al (1991). Important contributions to the larger space of enactivist, and enactivist-inspired, theorizing include Noe (2004) (2010), Thompson (2010), and Froese and Di Paulo (2011). The edited volume by Stewart et al (2010) provides an excellent window onto much of this larger space.

¹² Part of this image is misleading, insofar as it suggests that the external world is merely a source of undifferentiated perturbations (the repeated striking of a monotonous hammer). What seems correct is that the agent, by exposing herself to the varied stimulations predicted by the generative model, actively contributes to the world as sampled. Since it is only the world as sampled that the model needs to accommodate and explain, this delivers a very real sense in which (subject to the overarching constraint of structural self-maintenance i.e. persistence and survival) we do indeed build or 'enact' our individual and species-specific worlds.

¹³ I have engaged such arguments at great length elsewhere – see Clark (1989) (1997) (2008) (2012). For sustained arguments *against* the explanatory appeal to internal representation, see Ramsey (2007), Chemero (2009), Hutto and Myin (2013). For some useful discussion, see Sprevak (2010) (2013), Gallagher, Hutto, Slaby, and Cole (2013).