



THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### Happily entangled

**Citation for published version:**

Miller, M & Clark, A 2018, 'Happily entangled: Prediction, emotion, and the embodied mind', *Synthese*, vol. 195, no. 6, pp. 2559–2575. <https://doi.org/10.1007/s11229-017-1399-7>

**Digital Object Identifier (DOI):**

[10.1007/s11229-017-1399-7](https://doi.org/10.1007/s11229-017-1399-7)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Peer reviewed version

**Published In:**

Synthese

**Publisher Rights Statement:**

"The final publication is available at Springer via <http://dx.doi.org/10.1007/s11229-017-1399-7>.

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# Synthese

## Happily Entangled: Prediction, Emotion, and the Embodied Mind

--Manuscript Draft--

|  |   |
|--|---|
| <b>Manuscript Number:</b>                            | SYNT-D-16-00612R3   |
| <b>Full Title:</b>                                   | Happily Entangled: Prediction, Emotion, and the Embodied Mind   |
| <b>Article Type:</b>                                 | S.I. : Predictive Brains  |
| <b>Keywords:</b>                                     | predictive processing; affective neuroscience; embodied mind; affordance competition; pulvinar  |
| <b>Corresponding Author:</b>                         | Mark Miller, Ph.D.<br>University of Edinburgh<br>Edinburgh, UNITED KINGDOM  |
| <b>Corresponding Author Secondary Information:</b>   |   |
| <b>Corresponding Author's Institution:</b>           | University of Edinburgh   |
| <b>Corresponding Author's Secondary Institution:</b> |   |
| <b>First Author:</b>                                 | Mark Miller, Ph.D.  |
| <b>First Author Secondary Information:</b>           |   |
| <b>Order of Authors:</b>                             | Mark Miller, Ph.D.<br>Andy Clark  |
| <b>Order of Authors Secondary Information:</b>       |   |
| <b>Funding Information:</b>                          |   |
| <b>Abstract:</b>                                     | Recent work in cognitive and computational neuroscience depicts the human cortex as a multi-level prediction engine. This 'predictive processing' framework shows great promise as a means of both understanding and integrating the core information processing strategies underlying perception, reasoning, and action. But how, if at all, do emotions and sub-cortical contributions fit into this emerging picture? The fit, we shall argue, is both profound and potentially transformative. Far from being additional mechanisms operating alongside the multi-level predictive core, emotion and sub-cortical processing appear to be inextricably interwoven with the construction of the multi-level predictions that determine our percepts and select our actions. The upshot is a picture of the mind as fundamentally in the business of preparing the organism for action. |

# **Happily Entangled: Prediction, Emotion, and the Embodied Mind**

Mark Miller

University of Edinburgh, School of Philosophy, Psychology, and Language Sciences, Dugald Stewart  
Building, 3 Charles Street, Edinburgh, EH8 9AD, Scotland, UK

Andy Clark

University of Edinburgh, School of Philosophy, Psychology, and Language Sciences, University of  
Edinburgh, Charles Street, Edinburgh EH8 9AD, Scotland, UK

## Happily Entangled: Prediction, Emotion, and the Embodied Mind

Mark Miller (University of Edinburgh)

Andy Clark (University of Edinburgh)

### Abstract

Recent work in cognitive and computational neuroscience depicts the human cortex as a multi-level prediction engine. This ‘predictive processing’ framework shows great promise as a means of both understanding and integrating the core information processing strategies underlying perception, reasoning, and action. But how, if at all, do emotions and sub-cortical contributions fit into this emerging picture? The fit, we shall argue, is both profound and potentially transformative. In the picture we develop, online cognitive function cannot be assigned to either the cortical or the sub-cortical component, but instead emerges from their tight co-ordination. This tight co-ordination involves processes of continuous reciprocal causation that weave together bodily information and ‘top-down’ predictions, generating a unified sense of what’s out there and why it matters. The upshot is a more truly ‘embodied’ vision of the predictive brain in action.

### 1. The Strange Architecture of Predictive Processing

In 2012 the AI pioneer Patrick Winston wrote about the “puzzling architecture” of the brain – an architecture in which “Everything is all mixed up, with information flowing bottom to top and top to bottom and sideways too.” He added, “It is a strange architecture about which we are nearly clueless” (Winston, 2012).

It is a strange architecture indeed. But that state of cluelessness is increasingly past. A wide variety of work – now spanning neuroscience, psychology, robotics and artificial intelligence is converging on the idea that one key role of that downward-flowing influence is to enable higher-levels to attempt (level-by-level, and as part of a multi-area cascade) to try to predict lower-level activity and response<sup>i</sup>. That predictive cascade leads all the way to the sensory peripheries, so that the guiding task becomes the ongoing prediction of our own evolving flows of sensory stimulation. The idea that the brain is (at least in part, and at least sometimes) acting as some form of prediction engine has a long history, stretching from early work on perception (Helmholz, 1860; MacKay, 1956; Neisser, 1967; Gregory, 1980) all the way to recent work in deep learning (Hinton, 2007, 2010).

A promising subset of such work is the emerging family of approaches known as ‘predictive processing’<sup>ii</sup>. Predictive processing plausibly represents the last and most radical step in the long retreat (see Churchland et al., 1994) from a passive, feed-forward, input-dominated view of the flow of neural processing. According to this emerging class of models biological brains are constantly active, trying to predict the streams of sensory stimulation before they arrive. Systems like that are most strongly impacted by sensed *deviations* from their predicted states. It is these deviations from predicted states (prediction errors) that now bear much of the information-processing burden, informing us of what is *newsworthy* within the dense sensory barrage. When you see that steaming

coffee-cup on the desk in front of you, your perceptual experience reflects the multi-level neural guess that best reduces visual prediction errors. To visually perceive the scene, if this story is on track, your brain attempts to *predict* the scene, allowing the ensuing error (mismatch) signals to refine its guessing until a kind of equilibrium is achieved.

To appreciate the benefits, first consider learning. Suppose you want to predict the next word in a sentence. You would be helped by knowledge of grammar. But one way to learn a surprising amount of grammar, as work on large-corpus machine learning clearly demonstrates, is to try repeatedly to predict the next word in a sentence, adjusting your future responses in the light of past patterns. You can thus use the prediction task to bootstrap your way to the world-knowledge that you can later use to perform apt prediction. Importantly, learning using *multi-level* prediction machinery delivers a *multi-scale* grip on the worldly sources of structure in the sensory signal. In such architectures, higher levels learn to specialize in predicting events and states of affairs that are – in an intuitive sense – built up from the kinds of features and properties (such as lines, shapes, and edges) targeted by lower levels. But all that lower-level response is now modulated, moment-by-moment, by top-down predictions<sup>iii</sup>.

To make the best and most flexible use of the flow of prediction error PP architectures simultaneously estimate the so-called “precision” of the prediction error signal itself. Precision is the inverse variance of a prediction error signal – in other words, it sets error bars around an error signal according to its currently estimated importance or reliability. High-precision errors enjoy greater post-synaptic gain and (hence) increased influence. Conversely, even a large prediction error signal, if it is assigned extremely low precision, may be rendered systemically impotent, unable to drive learning or further processing. This enables different circumstances to render different prediction error signals important, and may mandate different balances between processing in different brain regions and between top-down prediction and incoming sensory evidence.

Action itself is accomplished using the same resources. The core idea here (Friston et al, 2010) is that there are two ways for brains to match their predictions to the world. Either find the prediction that best accounts for the current sensory signal (perception) or alter the sensory signal to fit the predictions (action). Importantly, the flow of action can *itself* be brought about, PP suggests, by a select sub-set of predictions – prediction of the (trajectory of) proprioceptive consequences that would ensue were the desired action to be performed. This turns out to be a computationally efficient way of implementing motor commands (Friston, 2011).

A central claim of this 'active-inference' view is that top-down predictions and bodily actions co-evolve in circular and self-organizing ways. Friston and colleagues write, "Crucially, this inference or assimilation is active, in the sense that the internal states affect the causes of sensory input vicariously, through action. The resulting circular causality between perception and action fits comfortably with many formulations in embodied cognition and artificial intelligence; for example, the perception–action cycle [Fuster, 2004], active vision [Wurtz et al., 2011], the use of predictive information [Ay et al., 2008; Bialek et al., 2001; Tishby & Polani, 2010], and homeokinetic formulations [Soodak & Iberall, 1978]. Furthermore, it connects these perspectives to more general treatments of circular causality and autopoiesis in cybernetics and synergetics [Haken, 1993], [Maturana & Varela, 1980]" (2014, p.443).

The resulting picture is one in which perception and action are complementary manifestations of a single adaptive regime, whose core operating principle is the reduction of precision-weighted prediction error.

Adaptive predictions cannot, however, take shape in an organismic vacuum. What my brain predicts, moment-by-moment, needs to be delicately geared to what I need, and to what I need to be doing. And what I need, and what I need to be doing, are both matters that depend heavily upon both my current physiological states and the shape and progress of current world-engaging activity. It is here that sub-cortical structures, and especially the thalamus (and within it, the pulvinar) seem posed to play a special and crucial role. Understanding that role requires us to move beyond what Luiz Pessoa (2014, p 11) describes as the ‘cortico-centric’ image in which evolutionary older subcortical structures are dominated and controlled by the more recent cortical overlay. Instead, we will be led to endorse an ‘embedded’ view (op. cit., p.14) according to which cortical and sub-cortical states and activities change in a co-ordinated fashion characterized by ongoing patterns of mutual influence.

## **2. Continuous Reciprocal Causation in Cortico-subcortical Loops**

The term *cortical myopia* was first coined by Josef Parvizi (2009) in reference to a lingering tendency in contemporary neuroscience to under-appreciate or to ignore the rich contributions sub-cortical brain regions make to higher cognitive function and intelligent behavior. The bias comes to us in part as a hangover from 19th century experimental brain research (see LeDoux, 1987). At the time human brain evolution was imagined to be a largely linear affair, with phylogenetically newer and more advanced cortical areas overlaying and controlling older more primitive subcortical areas (Herrick, 1933; Papez, 1937). With higher cognitive functions such as language seen evolving alongside the massive expansion of the neocortex (Barton and Harvey, 2000) researchers naturally assumed higher cognition to be localized in the uppermost cortical tip of the neural axis. Together with Charles Darwin’s astute observations that basic emotions (eg. fear and rage) are shared across species, this led researchers to look for emotional/instinctual circuits in the older and highly conserved sub-cortex. Their conclusion was that human rationality emerged as the new and improved neocortex exerted increasing control over the outdated emotional-instinctual circuitry. As John Hughlings-Jackson wrote, “the higher nervous arrangements evolved out of the lower keep down those lower, just as a government evolved out of a nation controls as well as directs that nation” (Jackson, 1884, p. 662, quoted by Parvizi, 2009, p. 354). This picture of the brain has produced a long-standing tradition in cognitive neuroscience of investigating cortical and sub-cortical structures as dichotomous sets of regions with “higher” circuits controlling/inhibiting the “lower” circuits (perhaps most dramatically in MacLean’s (1990) ‘triune’ brain model).

The major issue with such cortico-centric views is not that the brain might be usefully described in hierarchical terms (see Lewis and Todd, 2005), nor is it the claim that processing ‘higher’ up the neural axis is essential for cognitive functions such as decision making and language (which they most certainly are). What needs to be doubted is the assumption of a clear division of labor between a higher ‘cognitive brain’ and subordinate ‘emotional brain’ and the assumption that higher areas influences lower ones, but not the other way around. As neuroscientific techniques improve, it is becoming increasingly clear that such a dichotomous view of neural processing is too

simplistic to capture the complex, reciprocal and self-organizing nature of human brain function.

Contrary to the Victorian view, the cortex is not a total newcomer to human brain evolution, but has in fact long been part of the basic mammalian neural floor-plan (Pessoa, 2013). Moreover both cortex *and* sub-cortex have continued to change throughout human evolution. For example sub-regions of the human amygdala are believed to be 60% larger than apes' relative to brain mass (Sherwood et al., 2012). Compare this with only a 24% increase in neocortical mass (Rilling and Insel, 1999), and surprisingly no significant increase in frontal lobe mass (Semendeferi et al., 2002). And finally there is growing evidence that cortical and sub-cortical areas evolved in a highly coordinated fashion, thereby producing rich looping interdependencies between lower emotional and higher cognitive circuits. A recent proposal by Chareyron and colleagues (2011) proposes that brain areas which are structurally or functionally interconnected evolve in tandem promoting increases in the convergence and integration of information between the areas. A good example of such coordinated cortical-sub-cortical growth comes from Barton (2012), who suggests that the increased size of the primate cerebellum should be viewed in the context of a three-way co-evolution between the diencephalon, cerebellum and the neocortex (see also Barton and Harvey, 2000). Pessoa (2014) makes a similar observation about the massive size increase in the primate amygdala and its remarkable connections (both afferent and efferent) to a wide variety of cortical and subcortical regions.

The result of this coordinated evolution has been the creation of a rich embedding of neural circuitry in which newer circuits are continually woven into older ones producing novel functional landscapes stretching across the entire brain. As Pessoa writes, this interweaving “creates a web of functional and structural couplings in a way that blurs “old” and “new”” (Pessoa, 2015, p. 49). One way to see this is to note that complex sub-cortical dynamics now continuously influence, and are continuously influenced by, complex cortical dynamics. Such processes (of ‘continuous reciprocal causation’ – see Clark (1997)) bind multiple ‘components’ into unified dynamic wholes in which the state variables of one system are the parameters of the other, and vice versa<sup>iv</sup>. Such complex couplings are characteristic of evolved systems in which neural processing, bodily action, and environmental forces are constantly and complexly combined. In the case at hand, we shall see that sub-cortical systems are themselves constantly informed by bodily changes and our own ongoing actions, thus coupling neural predictions and bodily unfoldings in deep and transformative ways.

### **3. Sampling the Coupling**

To put flesh on these teleo-functional speculations consider the huge number of subcortical structures that target cortical regions either directly or via the thalamus, including areas such as the basal forebrain (Dunnett et al., 1991), hypothalamus (Pessoa, 2014), basal ganglia (Clarke et al., 2008), amygdala (Pessoa, 2013), cerebellum (Leiner et al., 1986), and brainstem via dopaminergic and noradrenergic systems (Parvizi and Damasio, 2000; Mather et al., 2016; Markovic et al., 2015). Many of these ascending systems form important reciprocal loops with the cortex. For example the basal ganglia is connected to the cortex by at least five separate circuits, some of which form closed loops with cortex via the thalamus. This allows information flowing from cortical areas to basal ganglia to return again to the same cortical area (Parvizi, 2009). As Parvizi writes,

the richness of this looping relationship means that “in reality, there is no cortex versus basal ganglia divide. One does not exist without the other, and there is only an interlinked network of corticostriatal loops” (op. cit. p.356). From this perspective, online cognitive function cannot be assigned to either the cortical or subcortical component, but instead emerges from their tight coordination.

Work on the hypothalamus provides further examples of the rich suite of interactions weaving cortex and sub-cortex. In the past, researchers primarily focused on the hypothalamus' descending systems (connecting to brainstem and spinal cord), leading to its characterization as a homeostatic controller (Bard, 1928; Cannon, 1929). The hypothalamus also receives information from the body allowing it to finely tune affective responses to the environment (for a good discussion see Lewis and Todd, 2005). Recently however our understanding of hypothalamic connectivity has expanded to include a rich set of bidirectional pathways connecting to the whole cortical mantle (Risold, Thompson, Swanson, 1997). Contrary once again to the corticocentric vision, the hypothalamus appears to exert a large influence on cortical function both directly and indirectly via the thalamus, basal forebrain, amygdala and brainstem (all of which are themselves bidirectionally connected to the cortex and each other (Pessoa, 2014). This makes the hypothalamus the second largest contributor to cortical inputs after the thalamus (Swanson, 2000). The reciprocal connectivity to both cortical areas and the internal body would allow the hypothalamus to synchronize huge flows of information within the brain and body (Pessoa, 2013, p. 230–231). As Kiverstein and Miller have recently noted, “this coordination [facilitated by the hypothalamus] allows for cognitive and affective processes to be mobilized together allowing the animal to behave flexibly, and in ways that are adapted to the particularities of a context of activity” (Kiverstein and Miller, 2015).

Finally, consider the profound reciprocal relationship that exists between prefrontal cortex and brainstem nuclei. The brainstem (and striatal) circuits are believed to play a central role in generating rapid emotional responses (the so called ‘action tendencies’ in Frijda's work; 1986) and so have been called the “seat of emotions” (Panksepp, 1998). Meanwhile the prefrontal cortex uses highly processed information from a variety of cortical areas to directly inhibit brainstem reactions thereby allowing time for more sophisticated, and context appropriate, behaviors to emerge. But once again, this is not a one-way relationship. Instead, systems in the brainstem also modulate the frontal lobes by way of neuromodulatory chemicals. Neuromodulatory systems producing dopamine, norepinephrine and acetylcholine within the brainstem, basal forebrain, hypothalamus have terminals in a huge portion of both the sub-cortex and prefrontal areas (Arnsten & Li, 2005; Joels et al., 2006). Such neuromodulatory systems are believed to attune cortical processing to signals from the body and environment that are important for survival (Lewis and Todd, 2005). As they have recently written:

“if not for the bottom-up flow, the brain would have no energy and no direction for its activities. If not for the top-down flow, recently evolved mechanisms for perception, action and integration would have no control over bodily states and behavior. It is the reciprocity of these upward and downward flows that links sophisticated cognitive processes with basic motivational mechanisms” (op. cit. p. 20).

During complex behaviors, elements of emotion and cognition are thus so intermixed that a significant decomposition becomes impossible at the level of the brain (Pessoa,



2013). If a meaningful decomposition of emotion and cognition is indeed impossible, then processes considered to be emotional will be poised to play a much richer role than previously proposed (eg. influencing vision). Just such an account has recently been proposed, within the predictive processing framework, by Barrett and Bar (2009) – see also Barrett and Simmons (2015) and Chanes and Barrett (2016).

With massive ascending and descending circuits the PFC becomes an important center of vertical integration of information (Pessoa, 2015). PFC along with related areas such as the orbitofrontal, cingulate and insula cortex are all richly interconnected with one another and with amygdala and hypothalamus both of which have huge influence over internal (affective) processes. This collection of areas has also recently been highlighted as part of the so called 'rich club': a community of highly interconnected neural hubs that serve as the backbone for brain wide (cortical and subcortical) synchronizations (Van den Heuvel & Sporns, 2011, 2013a, 2013b; see also Chanes and Barrett (2016) for a good discussion on the relationship between rich clubs and predictive processing). This tight vertical integration of neural processing suggests that cognitive and emotional processes are strongly interdependent (Lewis, 2005; Stapleton, 2013; Pessoa, 2013; Colombetti, 2014). In contrast to the corticocentric vision, cognition and behaviour are better seen as emerging from numerous systems stretching the entire neural axis and dynamically interacting via feed-forward and feed-backward loops (see Lewis, 2005 for a richly detailed account of this ebb and flow).

#### **4. A New Look at the Thalamus**

This picture of dense cortical-sub-cortical coupling is further enriched by new understandings of the thalamus itself. While traditionally described as a byway through which information was shuttled into the cortex, today the thalamus is now being re-described as an important neural nexus point capable of orchestrating huge flows of cortical communication. As Pessoa writes, "corticothalamocortical information transfer may represent an important addition to, or even replacement of, the current dogma that corticocortical transfer of primary information exclusively involves direct corticocortical pathways" (Theyel et al., 2010)

In just this vein, Sherman and Guillery (2013) argue that large amounts of thalamic connectivity are not simple way-stations or 'first order' relays, conveying information to the cortex from some sub-cortical source such as the retina. Instead, most of the thalamus is said to be composed of 'higher-order relays': circuits that transmit information between cortical areas – specifically, from layer 5 of one cortical area to some other cortical area. This intriguing feature of the connectivity matrix is directly suggested by impressive bodies of physiological and anatomical evidence, reviewed and summarized by Sherman and Guillery. It suggests that the primary role of much thalamic connectivity may be to mediate intra-cortical communication. If so, the question naturally arises, what are the differences in the kinds of information being carried by these various pathways? Here, it is notable that "the extra-thalamic targets of drivers<sup>v</sup> to the thalamus seem to be involved in motor control" (Sherman, 2007, p.420). This opens up an intriguing possibility, which is that the information conveyed in cortico-thalamic-cortical circuits may be profoundly entangled with unfolding motoric commands and activity. This possibility has been defended and explored by Guillery (2003, 2005), and is further refined by Sherman and Guillery (2013). On this speculative account,

transthalamic cortico-cortico pathways tend to transmit information about the motor consequences of current processing in that area. This means that:

“at every level of sensory processing, perception is inextricably linked to ongoing instructions for action, prior to the action itself.” Sherman and Guillery (2011, p.1073).

Sherman and Guillery go on to suggest that the thalamus may thus play a role in establishing and maintaining transient action-specific dynamic couplings between cortical areas, and in alerting cortical areas to any unexpected motor instructions being computed by other areas (see Sherman and Guillery (2011) pp. 1074). The resulting picture is one in which “sensorimotor processing is unified throughout all levels of thalamo-cortical function” (op cit p. 1075). Processes of continuous reciprocal causation coupling cortical and thalamic sources here put higher-level prediction machinery in direct contact with unfolding bodily actions. But more importantly, they do so in ways that thus defy easy decomposition into ‘leader’ and ‘led’. Instead, bodily actions and complex top-down predictions co-evolve, delivering just the kinds of ‘circular causality’ between perception and action imagined by proponents of ‘active inference’ (see e.g. Friston, Adams et al (2012)).

This emerging vision of the densely woven cortico-sub-cortical economy is essential, we will now argue, if we are to flesh out key aspects of the predictive processing story described earlier. In particular, it will help us understand both the power, and the implementation, of a key component of that story – the variable precision-weighting of prediction error.

## **5. Precision and the Threat of Magic Modulation?**

PP accounts are distinguished, in part, by their pervasive reliance upon ‘precision estimations’ to perform a variety of key tasks and functions. Precision estimates, as mentioned earlier, track the inverse variance of the prediction error signal. In other words, they set error bars around different aspects of that signal according to their estimated reliability, given the task and context. Precise prediction error signals result in increased post-synaptic gain, driving further processing more strongly than their less precise cousins.

There are two main (but deeply related) roles played by precision within the PP architecture. First, and most generally, variable precision weighting alters the balance between top-down prediction and the incoming sensory signal, allowing us to rely on specific chunks of sensory evidence to a greater or lesser degree depending upon task and context. For example, on a clear but windy day, for many tasks, visual information remains highly accurate and reliable and should be given more weight than (say) smell. By varying the impact of specific aspects of prediction error upon further processing, PP systems achieve a remarkable degree of flexibility in their use of long-term stored knowledge about the world. In the case of the McGurk effect (see McGurk and McDonald, 1976) for example, we allow visual information from a mismatched, overdubbed video of a speaking face to overwhelm some aspects of the auditory signal, resulting in our mishearing ‘ba’ as ‘da’. This makes ecological sense since lip movements are normally an excellent cue to speech sounds, and we must often rely upon them in situations of noise or uncertainty.

Second, precision determines the nature of control. For example, Pezzulo et al. (2015) leverage the precision estimation mechanism as a means of ‘flipping’ between habitual and more reflective means of control. Habitual control emerges when sensory prediction error is rapidly resolved at lower levels of the processing hierarchy. More reflective means of control emerge when prediction error is resolved at higher levels – levels that contextualize lower-level responses. In this way:

“the ontology of behavioural paradigms in associative learning can be seen as a successive contextualisation of more elemental sensorimotor constructs, within generative models of increasing hierarchical depth” Pezzulo, Rigoli, and Friston (2015, p.18)

Higher-level representations here entrain systemic response over longer time-scales, by predicting – and hence helping to bring about - more complex sequences of responses and environmental (or bodily) states. Influential work on ‘multiple controllers’ for habitual (model-free) versus goal-directed (model-based) response is here accommodated within a single precision-modulated inferential schema in which:

“it is the precision or reliability of alternative controllers that arbitrates their relative contribution” (op cit p. 19)

High precision predictions exert greater force, and when those predictions originate from much higher levels, they entrain prospective forms of control – forms of control that anticipate and help bring about extended sequences of inputs so as to implement choices and policies concerning future actions. This is the domain that is sometimes referred to as ‘counterfactual prediction’ – prediction that is truly oriented towards the future, and concerns states of affairs that do not yet obtain. Control is thus:

“not dichotomized into two discrete systems [habitual and model-based], but viewed as distributed along a graded continuum going from the highest levels of abstract, prospective and conscious reasoning to more concrete, short-sighted unconscious levels of reasoning down to the arc reflex”. (op cit p.24)

This picture can be enriched in various ways, for example by noting that habitual control (here, the resolution of high-precision prediction errors using only lower-levels of the processing regime) may sometimes itself extend over larger time-scales, as in the case of highly skilled, over-learned sports performances. But for present purposes, what matters is simply the pervasive role of precision estimation in supporting flexible, context-sensitive responses that seamlessly negotiate a smooth continuum between more-or-less ‘automatic’ and goal-directed modes. Behaviour, if these accounts are on track, is contextualized by different hierarchical levels in ways that are arbitrated (op cit p.27) by precision dynamics. Precision here performs two distinct yet inter-related functions. It determines which areas and hierarchical levels currently exert most control. And it “reports opportunities to achieve a goal” (op cit p.28) by reflecting current confidence in those opportunities, and also by responding to signs of progress or failure.

In order to perform these functions adequately, variations in precision-estimation must be delicately responsive to an interacting medley of exteroceptive, interoceptive, and proprioceptive sensory signals. For what goals we pursue, what actions we perform, which aspects of behavior can safely be left to habitual control, and which demand

higher-level contextualization, are all matters that require the simultaneous satisfaction of multiple kinds of constraint. Are we in physical danger? Are we hungry, or likely to become hungry if we do not take precautionary action? Is there a sudden opportunity to fulfill a long-standing goal? Is our body currently able to reach some desired target? Do we have enough information to make a good call on whether or not to pursue a certain goal, or should we instead act so as to harvest additional information? Delicate waves of precision-engineered influence must reflect our brain's best task-and-context reflecting guesses about all these matters, modulating the impact of every aspect of the prediction error signal so as to soft-assemble neuronal resources into temporary webs that keep us viable and that enable us to achieve our goals.

Such spectacular fluidity might raise suspicions. It can sometimes seem as if precision-variation is playing the role of a 'magic modulator', putatively altering the balances of internal power so as to allow the PP framework to accommodate every conceivable form of adaptive behavior, from simple reflexes, to the most complex goal-driven unfoldings. Have we inadvertently imported an un-explanatory free variable into our explanatory schema? We believe that the threat of magic modulation can be averted once we better appreciate the role of sub-cortical processing in the estimation, orchestration and distribution of precision.

## **6. Sub-cortical Contributions to Precision Estimation**

Implemented by multiple means in the brain (such as neurotransmitter-based modulation, and temporal synchronies between neuronal populations) flexible precision-weighting renders these architectures spectacularly fluid and context-responsive. Sub-cortical contributions affords further opportunities to incorporate constantly updated information about the state of the body and its readiness for action, and about the uncertainties associated with the bodily information itself.

For example, Kanai et al. (2015) suggest that thalamic nuclei weight precision in the visual cortex. More specifically, their account focuses on the pulvinar. The pulvinar complex is the largest mass in primate thalamus and connects to a wide variety of cortical and subcortical areas via thalamocortical loops. It has extensive bidirectional connections with visual, temporal, parietal, cingulate, frontal and insular cortices, as well as the amygdala. As Pessoa writes, "at a gross level, it is as if the entire convoluted cortex were 'shrink-wrapped' around the pulvinar" (2014, p.11). This massive connectivity provides the pulvinar with ample opportunity to modulate the flow of information across much of the brain.

Kanai et al.'s proposal is that a key task of the pulvinar is to alter the influence (gain) of specific cortically-computed prediction errors so as to reflect their estimated precision. Such a sub-cortical contribution would be a prime instance of deep (cortico-sub-cortical) neural embedding. Core computations performed by the cortex would here be sensitively and constantly modulated by information about the state of the body and unfolding actions, as registered by the sub-cortical nexus. These gain alterations would, in turn, impact the flow of moment-by-moment cortico-cortico communication, routing and re-routing flows of information and control as task and context unfold. The pulvinar, Kanai et al suggest, is both architecturally well-suited and anatomically well-situated to perform this role.

This has quite radical implications. Because many subcortical circuits are tightly coordinated with internal bodily processes (vascular, visceral, endocrine, autonomic) information from the body turns out to play a much more important role than that assumed by the corticocentric vision of the brain. A particularly important function that requires this integration is the evaluation of (and motoric response to) sensory information. This is the bodily-sub-cortical-cortical weave that “addresses the question: given the present sensory information and the organism’s present internal state, how should it act?” (Pessoa, 2008 p.150).

Consider vision. What has been called the standard account of vision assumes a division of labor between a 'high-road' ascending from the retina through the visual cortex in a (mostly) hierarchical manner, and a 'low road' that fast-tracks affectively salient visual information from retina through the superior colliculus and pulvinar to the amygdala thereby helping to produce quick survival behaviors. The subcortical pathway is typically referred to in order to help explain the fast reaction time emotionally salient information produces in the brain and body (Pessoa, 2013). Many such dual-systems models exist, proposing two competing (or sometimes cooperating) routes: a faster, automatic, emotional, subcortical route, and a slower, controlled, cognitive, cortical route (e.g. Kahneman, 2003; Stanovich & West, 2000). These models remain both ‘cognitivist’ and corticocentric insofar as they assume emotional processing takes place entirely sub-cortically, and often in a way that is completely insulated from so ‘higher’ processes such as awareness and attention (believed to be processed cortically). They are also myopic in so far as they are blind to the wealth of subcortical contributions to perception. To help map this more complex process Pessoa describes in detail six circuits (not meant to be exhaustive) that work to fold value into cognition and behaviour by biasing cortical processing towards patterns that are important for the organism's survival. They include the amygdala, cortical valuation regions such as the OFC and insula, fronto-parietal attentional networks, basal forebrain and reticular nucleus, and the pulvinar.

Of particular importance for present purposes is the novel role Pessoa and colleagues propose for the pulvinar nucleus. Citing a wealth of experimental evidence and neuro-anatomical data, Pessoa suggests abandoning the standard view of pulvinar as a simple subcortical by-way by which affectively salient signals are fast tracked from retina to amygdala (Pessoa, 2013). He offers in its place a detailed account of how the pulvinar biases visual processing towards signals that have emotional or motivational significance in part by way of its rich looping relationship with multiple cortical and sub-cortical areas.

Importantly, the pulvinar is connected with the entire cortical mantle (Shipp, 2003). Kanai and colleagues focus primarily on the inferior portion of the pulvinar which is connected with striate and extrastriate cortex (targeting all 20-30 visual areas). However the lateral and medial pulvinar are richly connected to many other cortical and subcortical regions. The lateral pulvinar connects to temporal and parietal lobes (as well as portions of extrastriate cortex), and the medial pulvinar connects to the parietal, frontal, orbital, cingulate and insular cortex and amygdala. Many of these areas in turn have rich bidirectional relationships with visual cortex as we saw above (including the OFC and amygdala). The medial pulvinar here modulates the flow of multimodal information between a huge collection of cortical and sub-cortical areas including OFC, AIC, ACC, and amygdala believed to be of central importance in determining the affective value of signals and preparing the organism to respond appropriately (Pessoa and Adolphs 2010). In this way the medial pulvinar is proposed to amplify weak or

fleeting signals that have biological value thereby producing greater behavioural responses (Pessoa 2014, p.72). This optimal positioning and rich interconnectivity allows the pulvinar to fold value into the action-perception system in ways that respect these affective valuations.

This may be a good moment to respond briefly to an important pair of issues raised by an anonymous referee. The first is: why do we take predictive processing (rather than some other theory) to be good candidate for a theoretical account of sub-cortical-cortical connectivity? The second – closely related – is whether the complex dynamical story we favor, replete as it is with complex looping influence and couplings, is actually consistent with the fundamental tenets of predictive processing anyway. Both these issues resolve once it is appreciated that our fundamental claim is that sub-cortical processing plays a major role in delivering the evolving flow of precision estimation essential to fluid and task-optimized cortical processing. Such precision estimations lie at the very heart of the predictive processing machine, sculpting the moment-by-moment flow of information in the brain. Their role, recall, is to enable contextual information to reconfigure the impact of any area or level upon any other area or level according to the changing details of task and of inner and outer context. Our suggestion, in line with Kanai et al (2015) is that these precision estimations are, to a surprising extent, sub-cortically mediated. This is what keeps them in touch (as they need to be) with both the ongoing flow of embodied action and the changing physiological state of the organism. It was not our aim, however, to present evidence that predictive processing is the best story about cortico-sub-cortical connectivity. Rather, we assume (for the purposes of the paper) that the predictive processing story is worth pursuing in general, and ask how sub-cortical influence might fit into the story. The fit, we argued, is excellent – such connectivity is in fact ideally placed to carry out the important task of embodied-action-based precision modulation. The upshot is that we should expect to find subtle looping dynamics whereby precision-weighted prediction error both reflects and enables ongoing action – the kinds of circular dynamics rightly foregrounded in traditional dynamical systems approaches. To fully appreciate the potential significance of this complex interaction, we next locate affect where it belongs – as a reflection of changing states of organismic readiness for action.

## **7. Affect and Action-Readiness**

Cisek (2007) was the flagship treatment of the so-called ‘affordance competition’ hypothesis, according to which:

“the brain processes sensory information to specify, in parallel, several potential actions that are currently available. These potential actions compete against each other for further processing, while information is collected to bias this competition until a single response is selected” Cisek (2007) p.1585.

The brain, so the story goes, is constantly computing, or starting to compute, a large set of possible actions. These possible actions (which are essentially states of partial ‘action readiness’) are computed constantly and in parallel. They are also, as Cisek and Kalaska (2011, p. 279) put it, ‘pragmatic’ insofar as “they are adapted to produce good control as opposed to producing accurate descriptions of the sensory environment or a motor plan”. All this makes good ecological sense, allowing time-pressed animals to partially

‘pre-compute’ multiple possible actions, any one of which can then be selected, completed, and deployed at short notice and with minimal further processing.

In line with such a view, Hoshi and Tanji (2007) found activity in monkey premotor cortex correlated with the potential movements of either hand in a bimanual reaching response task in which the monkey had to wait upon a cue signaling which hand to use. Similar results have been obtained for the preparation of visual saccades (Powell and Goldberg, 2000) and using behavioral and lesion studies of reaching behavior in human subjects (Humphreys and Riddoch, 2000; Castiello, 1999). Decision-making seems to be folded right into these densely interanimated loops so that, to a certain degree at least, “decisions about actions emerge within the same populations of cells that define the physical properties of those actions and guide their execution” (Cisek and Kalaska, 2011, p. 282).

Emotion belongs at the very heart of this embodied nexus. As processing proceeds, affect and content must be co-computed: intertwined (Barrett & Bar (2009)) within the process of settling upon a coherent, action-guiding interpretation of the scene. Sub-cortical mechanisms that assign precisions to cortically computed quantities seem ideally suited to the implementation of such affectively-informed affordance competition in the brain. The sub-cortical loops here keep ‘higher-level’ prediction systems constantly in touch with our evolving actions, Consistent with such a picture, Frijda (1986, 2007) proposes that affect itself reports on embodied action-readiness programs that simultaneously indicate the relationship between the organism and the environment, and motivate the organism to improve that relationship. As Frijda writes, "emotion, by its very nature, is change in action readiness to maintain or change one’s relationship to an object or event” (Frijda, 2004, p. 158). Frijda’s account thus slots neatly into place with the work of Barrett and Bar (2009), Lewis and Todd (2005) and Pessoa (2015) discussed earlier. The common theme is that affect reports on action-readiness, revealing cognition, emotion, and action as inextricably entwined .

Our speculative story – or better, story sketch - is now complete. The broad connectivity of the medial and lateral pulvinar allows it to integrate various streams of information including affect (insula), action (cingulate), value (OFC) and cognition/attention (frontoparietal). Such thalamocortical loops work to amplify weak or fleeting signals that have biological value thereby producing greater behavioral responses (Pessoa 2014, p.72). Sub-cortically orchestrated precision weighting thus emerges as a potent (and notably non-magical) tool for modulating bodily response, affect, and action. By fully weaving in sub-cortical contributions, we arrive at a vision of a brain permeated by affect, constantly preparing the organism for action<sup>vi</sup>.

## **8. Conclusions: Coupling the Active Body and the Predictive Brain**

If cortico- sub-cortical weave plays the roles we are suggesting, the consequences for our understanding of prediction, perception, and action are profound. On the one hand, attention to this delicate multi-dimensional weave should help allay a major worry about the PP approach – the worry that too many puzzles and problems are being solved by the blanket appeal to context- variable precision assignments. For that blanket appeal, it may now be hoped, may be cashed out in many different ways, that make the most of these (relatively newly- discovered) properties of multiple interacting cortico- sub-cortical loops. In particular, we argued that reflection of the role of the medial pulvinar highlights

the role thalamocortical loops play in directing various flows of information towards what is affectively salient.

The resulting picture is profoundly 'embodied' insofar as mutual couplings (with the full dynamical signature of continuous reciprocal causation) obtain between body, sub-cortex, and cortex, with sub-cortical (specifically thalamic) nuclei serving to bring bodily information constantly into the mix. These sub-cortical loops help influence precision estimations in ways that reflect bodily states and unfolding actions, allowing value (to the organism) and affect (relating to interocepted bodily states) to exert a continuous influence on high-level predictions, that themselves exert a continuous influence on bodily states and unfolding actions.

What begins to emerge is a richer vision of the predictive brain itself. Our neural prediction engines are fundamentally in the business of preparing the organism for action, courtesy of constant sub-cortically mediated two-way communication with bodily unfoldings. In this way we construct an affect-laden experiential world that is as much about our own changing needs as it is about the state of a mind-independent reality<sup>vii</sup>. Mind like these are thoroughly permeated by emotion and by readiness for action. Emotion, cognition, and preparation for action here form a single whole, self-organized around precision weighted, prediction-error minimizing interactions that span cortical and sub-cortical circuits. These interactions couple the active body to the predictive brain.

## References

- Akins, K (2006). Of sensory systems and the "aboutness" of mental states. *Journal of Philosophy*, 93 (7), 337-372.
- Arnsten, A. F., and Li, B. M. (2005). Neurobiology of executive functions: Catecholamine influences on prefrontal cortical functions. *Biological Psychiatry*, 57(11), 1377–1384.
- Ay, N., Bertschinger, N., Der, R., Güttler, F., & Olbrich, E. (2008). Predictive information and explorative behavior of autonomous robots. *The European Physical Journal B-Condensed Matter and Complex Systems*, 63(3), 329-339.
- Bard P. (1928). A diencephalic mechanism for the expression of rage with special reference to the sympathetic nervous system. *Am J Physiol.*, 84:490–515.
- Barrett, L. F., Bliss-Moreau, E. (2009). Affect as a Psychological Primitive. In Mark P. Zanna, editor: *Advances in Experimental Social Psychology*, Vol. 41:167-218.
- Barrett, L. F., Bar, M. (2009). See it with feeling: Affective predictions during object perception. Theme issue: Predictions in the brain: Using our past to generate a future (M. Bar Ed.) *Philosophical Transactions of the Royal Society B*, 364: 1325-1334.
- Barrett, L. F., and Simmons, W. K. (2015). Interoceptive predictions in the brain. *Nature Reviews Neuroscience*, 16, 419–429.
- Barton, R., and Harvey, P. H. (2000). Mosaic evolution of brain structure in mammals.



*Nature*, 405, 1055–1058.

Barton, R. (2012). Embodied cognitive evolution and the cerebellum. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367 (1599): 2097-2107.

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

Bialek, W., Nemenman, I., & Tishby, N. (2001). Predictability, complexity, and learning. *Neural computation*, 13(11), 2409-2463.

Bruineberg J., and Rietveld E. (2014). Self-organization, free energy minimization, and optimal grip on a field of affordances. *Frontiers of Human Neuroscience*, Vol. 8.

Cannon, W. B. (1929). *Bodily Changes in Pain, Hunger, Fear and Rage* 2nd ed. (New York: Appleton).

Castiello U. (1999). Mechanisms of selection for the control of hand action. *Trends. Cogn. Sci.* 3(7):264–71.

Chanes and Barrett (2016). Refining the role of limbic areas in cortical processing. *Trends in Cognitive Sciences*, 20:2: 96-106.

Chareyron L. J., Banta Lavenex P., Amaral D. G., Lavenex P. (2011). Stereological analysis of the rat and monkey amygdala. *J Comp Neurol.*, 519: 3218–39.

Churchland, P. S., Ramachandran, V., et al. (1994). ‘A Critique of Pure Vision’, In C. Koch and J. Davis, eds., *Large-Scale Neuronal Theories of the Brain*, (Cambridge, MA: MIT Press).

Cisek, P., Kalaska, J. F. (2010). Neural mechanisms for interacting with a world full of action choices. *Annu Rev Neurosci.*, 2010;33:269-98.

Cisek, P. (2007). Cortical mechanisms of action selection: the affordance competition hypothesis. *Philos Trans R Soc Lond B Biol Sci.*, 362(1485):1585-99.

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

Clark, A (2014). *Mindware: An Introduction to the Philosophy of Cognitive Science (Second Edition)* (Oxford University Press, NY).

Clark, A (2016). *Surfing Uncertainty: Prediction, Action, and the Embodied Mind* (Oxford University Press, NY).

Clarke, H. F. et al. (2008). Lesions of the medial striatum in monkeys produce perseverative impairments during reversal learning similar to those produced by lesions of the orbitofrontal cortex. *J. Neurosci.*, 28: 10972–10982.

Colombetti, G. (2013). *The Feeling Body: Affective Science Meets the Enactive Mind*. (MIT Press).

- Dunnett, S. B. et al. (1991). The basal forebrain-cortical cholinergic system: interpreting the functional consequences of excitotoxic lesions. *Trends Neurosci.*, 14: 494–501.
- Frijda, N. H. (1986). *The Emotions*. (Cambridge, UK: Cambridge University Press).
- Friston K. (2010). The free-energy principle: a unified brain theory? *Nat Rev Neurosci.* 11(2):127-38.
- Friston, K (2011). What is optimal about motor control? *Neuron* 72: 488-498.
- Friston K, Adams R. A, Perrinet L and Breakspear M (2012). Perceptions as hypotheses: saccades as experiments. *Frontiers in Psychology* 3:151.
- Friston K, Daunizeau J, Kilner J, Kiebel S. J. (2010). Action and behavior: a free-energy formulation. *Biol Cybern.* 102(3):227-260.
- Fuster, J. M. (2004). Upper processing stages of the perception–action cycle. *Trends in cognitive sciences*, 8(4), 143-145.
- Fuster, J. M. (2004). Upper processing stages of the perception–action cycle. *Trends in cognitive sciences*, 8(4), 143-145.
- Gregory, R. L. (1980). Perceptions as hypotheses. *Phil. Trans. R. Soc. Lond., Series B, Biological Sciences* 290(1038): 181-197.
- Guillery R. W. (2003). Branching thalamic afferents link action and perception. *Journal of Neurophysiology*, 90: 539-48.
- Guillery R. W. (2005). Anatomical pathways that link perception and action. *Progress in Brain Research*, 149: 235-56.
- Haken, H. (1983). Synergetics: An Introduction. *Non-Equilibrium Phase Transition and Self-Organisation in Physics, Chemistry and Biology*, 3rd ed. Berlin, Germany: Springer-Verlag.
- Helmholtz, H. (1860/1962). *Handbuch der physiologischen optik* (Southall, J. P. C. (Ed.), English trans.), Vol. 3. (New York: Dover).
- Herrick C. J. (1933). The functions of the olfactory parts of the cerebral cortex. *Proceedings of the National Academy of Sciences, USA*, 19:7–14.
- Hinton, G. E. (2007). Learning Multiple Layers of Representation. *Trends in Cognitive Sciences*, 11, 428-434.
- Hinton, G. E. (2010). Learning to represent visual input. *Philosophical Transactions of the Royal Society, B*. Vol 365: 177-184.
- Hohwy, J (2013). *The Predictive Mind* (Oxford University press, NY).
- Humphreys GW, Riddoch JM. (2000). One more cup of coffee for the road: object-action assemblies, response blocking and response capture after frontal lobe damage. *Exp. Brain Res.* 133: 81–93.

- Joels, M., Pu, Z., Wiegert, O., Oitzl, M. S., and Krugers, H. J. (2006). Learning under stress: How does it work? *Trends in Cognitive Sciences*, 10(4), 152–158.
- Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, 58: 697–720.
- Kanai, R., Komura, Y., Shipp, S., Friston, K. (2015). Cerebral hierarchies: predictive processing, precision and the pulvinar. *Philos Trans R Soc Lond B Biol Sci.*, 370(1668).
- Kiverstein, J., and Miller, M. (2015). The embodied brain: towards a radical embodied cognitive neuroscience. *Front Hum Neurosci.*, 9: 237.
- LeDoux, J. (1987). Emotion. In: Plum, F., editor. *Handbook of Physiology. 1: The Nervous System. Vol V, Higher Functions of the Brain*. Bethesda: *American Physiological Society*, 419-460.
- Leiner, H. C. et al. (1986). Does the cerebellum contribute to mental skills? *Behav. Neurosci.*, 100, 443–454.
- Lewis, M. (2005). Bridging emotion theory and neurobiology through dynamic systems modelling. *Behavioral and Brain Sciences*, 28:169-245.
- Lewis, M., Todd, R. (2005). Getting emotional: A neural perspective on emotion, intention, and consciousness. *Journal of Consciousness Studies*, 12 (8-10):210-235.
- MacKay, D (1956). The epistemological problem for automata. In CE Shannon & J McCarthy, editors, *Automata Studies*, Princeton, NJ: Princeton University Press, 235-251.
- Maturana, H. R., & Varela, F. (1980). Autopoiesis: the organization of the living. In V. F. Maturana HR (Ed.), *Autopoiesis and Cognition*. Dordrecht, Netherlands: Reidel.
- McGurk, H., and MacDonald, J. (1976). Hearing lips and seeing voices. *Nature* 264 (5588): 746–748.
- MacLean, P. D. (1990). *The Triune Brain in Evolution: Role in Paleocerebral Functions*. New York: Plenum Press.
- Markovic, J., Anderson, A. K., Todd, R. M. (2015). Tuning to the significant: neural and genetic processes underlying affective enhancement of visual perception and memory. *Behav. Brain Res.* 259:229-41.
- Mather, M., Clewett, D., Sakaki, M., Harley, C. W. (2015). Norepinephrine ignites local hot spots of neuronal excitation: How arousal amplifies selectivity in perception and memory. *Behavioral and Brain Sciences*, 1:1-100.
- Panksepp, J. (1998). *Affective Neuroscience: The Foundations of Human and Animal Emotions* (New York: Oxford University Press).
- Papez J. (1937). A proposed mechanism of emotion. *Archives of Neurology and Psychiatry*, 79:217– 224.

- Parvizi, J. (2009). Corticocentric myopia: old bias in new cognitive sciences. *Trends Cogn Sci.*, 13 (8): 354-9.
- Parvizi, J., Damasio, A. R. (2000). Consciousness and the brainstem. *Cognition*, 79, 135–160.
- Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience*, 9(2), 148–158.
- Pessoa, L. (2013). *The Cognitive Emotional Brain: from Interactions to Integration* (MIT Press).
- Pessoa, L. (2014). Understanding brain networks and brain organization. *Physics of Life Reviews*, 11: 400-35.
- Pessoa, L. & Adolphs, R. (2010). Emotion processing and the amygdala: From a “low road” to “many roads” of evaluating biological significance. *Nature Reviews. Neurosciences* 11(11):773–83.
- Petro L.S. & Muckli L. (2016). The brain's predictive prowess revealed in primary visual cortex. *PNAS*.
- Petro L.S., Vizioli L. & Muckli L. (2014). Contributions of cortical feedback to sensory processing in primary visual cortex. *Frontiers in Psychology*, 5:1223.
- Pezzulo, G., Rigoli, F., & Friston, K. (2015). Active Inference, homeostatic regulation and adaptive behavioural control. *Progress in Neurobiology*, 134, 17–35.
- Powell K. D, Goldberg M. E. (2000). Response of neurons in the lateral intraparietal area to a distractor flashed during the delay period of a memory-guided saccade. *J. Neurophysiol.* 84(1):301–10.
- Rilling J. K., Insel T. R. (1999). The primate neocortex in comparative perspective using magnetic resonance imaging. *J Hum Evol.*, 37:191–223.
- Risold, P. Y., Thompson, R. H., and Swanson, L. W. (1997). The structural organization of connections between hypothalamus and cerebral cortex. *Brain Research. Brain Research Reviews*, 24(2–3), 197–254.
- Semendeferi K., Lu A., Schenker N., Damasio H. (2002). Humans and great apes share a large frontal cortex. *Nat Neurosci.*, 5:272–6.
- Seth, A. K. (2015). The Cybernetic Bayesian Brain - From Interoceptive Inference to Sensorimotor Contingencies. In T. Metzinger & J. M. Windt (Eds). *Open MIND*: 35(T). Frankfurt am Main: MIND Group.
- Sherman, S. M. (2007). The thalamus is more than just a relay. *Current Opinion in Neurobiology*, 17:417-422.
- Sherman S. M. (2016). Thalamus plays a central role in ongoing cortical functioning. *Nature Reviews: Neuroscience* 16(4):533-41

- Sherman S. M., and Guillery RW (2011). Distinct functions for direct and transthalamic corticocortical connections. *J Neurophysiol* 106:1068-1077.
- Sherman, S. M. & Guillery, R. W. (2013). *Thalamocortical Processing: Understanding the Messages that Link the Cortex to the World* (MIT Press).
- Sherwood, C., Bauernfein, A. L., Bianchi, S., Raghanti, M. A., and Hof, P. R. (2012). “Human brain evolution writ large and small,” in *Evolution of the Primate Brain: From Neuron to Behavior*, eds M. A. Hofman and D. Falk (Oxford: Elsevier), 237–257.
- Shipp S. (2003). The functional logic of cortico-pulvinar connections. *Philos Trans R Soc Lond B, Biol Sci*, 358:1605–24.
- Soodak, H., & Iberall, A. (1978). Homeokinetics: A physical Science of Complex Systems. *Science*, 201, 18.
- Stanovich, K. E., West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, 23:645–726.
- Stapleton, M. (2013). Steps to a “Properly Embodied” cognitive science. *Cognitive Systems Research*, 22–23:1–11.
- Swanson L. W. (2000). Cerebral hemisphere regulation of motivated behavior. *Brain Res*, 886:113–64.
- Theyel B., Llano D., Sherman S. (2010). The corticothalamocortical circuit drives higher-order cortex in the mouse. *Nat Neuroscience*, 13:84–8.
- Tishby, N., & Polani, D. (2011). Information theory of decisions and actions. In *Perception-action cycle* (pp. 601-636). Springer New York.
- van den Heuvel MP, Sporns O. (2011). Rich-club organization of the human connectome. *J. Neurosci.* 31, 15 775 – 15 786.
- van den Heuvel MP, Sporns O. (2013a). Network hubs in the human brain. *Trends Cogn. Sci.* 17, 683 – 696.
- van den Heuvel M. P., Sporns O. (2013b). An anatomical substrate for integration among functional networks in human cortex. *J. Neurosci.* 33, 14 489 – 14 500.
- Winston, P. H. (2012) The Next 50 Years: A Personal View *Biologically Inspired Cognitive Architectures* 1: 92–99.
- Wurtz, R. H., McAlonan, K., Cavanaugh, J., & Berman, R. A. (2011). Thalamic pathways for active vision. *Trends in cognitive sciences*, 15(4), 177-184.

---

<sup>i</sup> There is a large and growing literature here. Good places to start include Friston (2005), (2010), Clark (2013), Hohwy (2013), Clark (2016).

---

<sup>ii</sup> For this usage, see Clark (2013).

<sup>iii</sup> This helps make sense of recent work showing that top-down effects (expectation and context) impact processing even in early visual processing areas such as V1- see Petro et al (2014) Petro and Muckli (2016). Recent work in cognitive neuroscience has begun to suggest some of the detailed ways in which biological brains might implement such multi-level prediction machines – see Bastos et al (2012).

<sup>iv</sup> See Clark (2014) chapter 7.

<sup>v</sup> A driver is traditionally distinguished from a modulator. Drivers, as the name suggests, are seen as primary transmitters of information whereas modulators alter the impact of that information. Driver inputs to a thalamic relay are thus diagnostic of the function of that relay, whereas modulator inputs are not - see Sherman and Guillery (2011). Within PP, precision-weighting acts a kind of universal modulator.

<sup>vi</sup> This is a contemporary version of the profoundly ‘motocentric’ vision of the brain suggested in the classic work by Churchland et al (1994).

<sup>vii</sup> Such a perceptual realm is constructed in a fashion that is deeply ‘narcissistic’ in exactly the sense of Akins (2006).