

How and why actions are selected: action selection and the dark room problem

Elmarie Venter
(University of Kwa-Zulu Natal)
weventer@gmail.com

Abstract In this paper, I examine an evolutionary approach to the action selection problem and illustrate how it helps raise an objection to the predictive processing account. Clark examines the predictive processing account as a theory of brain function that aims to unify perception, action, and cognition, but – despite this aim – fails to consider action selection overtly. He offers an account of action *control* with the implication that minimizing prediction error is an imperative of living organisms because, according to the predictive processing account, action is employed to fulfill expectations and reduce prediction error. One way in which this can be achieved is by seeking out the least stimulating environment and staying there (Friston *et al.* 2012: 2). Bayesian, neuroscientific, and machine learning approaches into a single framework whose overarching principle is the minimization of surprise (or, equivalently, the maximization of expectation. But, most living organisms do not find, and stay in, surprise free environments. This paper explores this objection, also called the “dark room problem”, and examines Clark’s response to the problem. Finally, I recommend that if supplemented with an account of action selection, Clark’s account will avoid the dark room problem.

DOI 10.1515/kjps-2016-0002

1. Introduction

Selecting an appropriate action is very important to the efficiency and fitness of an animal¹ because at any time an animal has numerous available actions to choose from. This is the action selection problem and can be stated as follows: Given an agent with a repertoire of available actions, a number of goals, and knowledge about its internal states and external environment, how does an agent select the most useful or most valuable action in a given situation?² The action selection problem can be approached in at least two ways by asking the following questions. First, why are actions selected? This question would involve discussion on reward, incentive and fitness. The second question is, how are actions selected, or what mechanisms are involved in selecting an action? This question requires an account of cognitive architecture, mechanisms and processes.

2. The architecture problem

In his 2013 paper, *Whatever next? Predictive brains, situated agents, and the future of cognitive science*, Andy Clark offers an account of the mechanisms and architecture involved in cognition and action. Clark supports the predictive processing account of cognition and shares his view with physicists, such as Karl Friston, and philosophers, such as Jakob Hohwy. Though Clark does not directly consider the action selection problem, his account may offer a solution to the problem involving cognitive architecture and mechanisms.³ The architecture of cognition is not the only element of cognition to study, efficiency is also a concern. Therefore, in this paper, I propose that cognition should be studied with two objectives in mind. First, an account of cognition should consider the architecture and mechanisms involved in cognitive systems – this will involve an explanation of how the system selects an action. I call this the

1 Houston, McNamara and Steer 2007, 1531.

2 Maes, 1990, 991; Seth, 2007, 1545.

3 Clark mentions action selection only twice: first to say that minimizing prediction error is the driving force behind action selection and second to briefly raise a concern about the scope of the predictive processing account.

architecture problem. Second, a successful account will also explain why an action is selected – this is the efficiency problem. Clark's account of cognition does not provide an answer to the efficiency problem because he does not take seriously the problem of action selection but the account tells us something about the mechanisms involved in cognition and action. The lack of a solution to the efficiency problem will be discussed in future sections.

3. Predictive processing model of cognition

The predictive processing account is characterized by the key notion that the brain uses prediction–driven processes "...to acquire and deploy hierarchical generative models of the hidden causes (sometimes called latent variables) that best explain the changing patterns of sensory input that impinge upon the agent."⁴ Traditional views of perception model the brain as a passive, stimulus–driven mechanism⁵ and take the process of perception to infer from effect to cause.⁶ These traditional accounts model the brain as feature detector – a view not supported by the predictive processing account. The predictive processing account proposes that instead of simply accumulating information and building a model of the world, the brain tries to predict the sensory information and attempts to match its predictions of sensory causes with actual sensory stimuli. Higher level systems predict activity in lower level systems; errors in higher level predictions are corrected based on the lower level input received. This is to say, any errors that occur in higher level models are modified to reduce future discrepancies.⁷ The function of higher level predictions is to explain away any incoming sensory stimuli to reduce information propagated; this is because only the prediction error is propagated instead of the entire sensory stimulus.⁸

4 Clark, 2012, 760.

5 Clark, 2014, 23.

6 Hubel & Wiesel, 1965; Biederman, 1987; Riesenhuber & Poggio, 2000.

7 Clark, 2013, 182.

8 Ibid., 182.

The predictive processing account described above is traditionally used to explain the mechanisms and functions involved in perception, but has recently been extended to include action.⁹ Clark offers the “action-oriented predictive processing” account which constitutes the novel claim that action and perception follow the same predictive principles. This follows from the presumption that perception and action are computationally related and work together to reduce prediction error. Prediction error is the result of a mismatch between bottom-up input and top-down predictions. Clark’s account does not consider action selection overtly but says that action control is driven by the minimization of prediction error.¹⁰ Agents with complex cognition, however, do not only *control* actions in response to the environment but are required to *select* the best possible action in a particular situation. In the following sections, I provide a descriptive account of the nature of the cognitive system as proposed by Clark. This will provide necessary background information needed to critically examine Clark’s action-oriented predictive processing account.

3.1. The nature of the system

Clark’s general account of cognition has three properties: 1) the system is bidirectional in nature. 2) The system is hierarchical in nature. 3) The system uses Bayesian inference in the selection of predictions/top-down information. I discuss each of these properties in turn.

3.1.1. Bidirectional nature

According to the predictive processing model of the brain, information is encoded using both top-down and bottom-up processes.¹¹ Top-down information is predictive in nature, and functions to predict the causes of sensory input; bottom-up input constitutes stimuli and functions to update predictions. This view – that cognition is bidirectional in nature – differs from traditional theories of cognition which propose that top-down and bottom-up processes are mutually exclusive. According to the predictive processing account, perception is a constant interaction

9 Friston & Stephan, 2007; Brown *et al.*, 2011; Clark 2013.

10 Clark, 2013, 190.

11 *Ibid.*, 186.

between current and previous information and constitutes comparing bottom-up stimuli with top-down predictions.¹² This process is known as “explaining away” and involves explaining away the incoming stimuli by matching it with expectations.¹³

An example of this process is an experiment investigating binocular rivalry which is a visual experience that occurs when each eye is simultaneously presented with a different visual stimulus.^{14,15} What is interesting about the experiment is that instead of seeing a merged picture, only one stimulus is perceived at a time.¹⁶ The perceptual system switches between the two pictures, focusing on each for a few seconds. This experiment is important in understanding bidirectional processing because the bottom-up stimuli remain constant but the top-down information changes.¹⁷ The top-down information explain away only those elements of the incoming signal that match the current prediction¹⁸ resulting in the switching of perceptual experience. This implies that perception does not necessarily include encoding the entire stimulus but rather encoding the *difference* between the sensory stimulus and a prediction. If the prediction is good, only the discrepancy between the prediction and the incoming stimulus will be perceived as input.¹⁹ If the prediction is poor, the entire input signal is encoded as new information. This difference is known as prediction error and is reduced through updating top-down information. Even though the top-down predictions are doing much of the perceptual work, the bottom-up signals provide ongoing feedback on top-down activity because perception is only guided by expectations and not enslaved by it.²⁰ Therefore, the view put forth is not radically internalist but a hybrid view using both internal knowledge and external experience.

12 Ibid., 186.

13 Ibid., 187.

14 See Hohwy *et al.*, 2008 for an in depth discussion on binocular rivalry.

15 Clark, 2013, 184.

16 Hohwy *et al.*, 2008, 690.

17 Clark, 2013, 185.

18 Ibid., 185.

19 Hohwy *et al.*, 2008, 689.

20 Clark, 2013, 190.

3.1.2. Hierarchical nature

The bidirectional process occurs at various levels of cognition and supports the view that cognition is hierarchical in nature. A key function of the hierarchical system is to guess the next states of its neural economy – each higher level in the cognitive hierarchy predicts the informational state of the lower level.²¹ In the binocular rivalry case, “... a visual percept is predicted by a process ... operating across many different levels of a (bidirectional) processing hierarchy, each concerned with different types and scales of perceptual detail”.²² This process constitutes backward updating of higher levels to keep prediction error to a minimum. The nature of the levels in the hierarchy varies in that the basic input level encodes basic stimuli and higher levels encode more abstract and perspective invariant data.²³

3.1.3. The Bayesian brain

At each level of the cognitive hierarchy, interaction between top-down and bottom-up information takes place. Bottom-up information constitute sensory stimuli and top-down information constitute predictions about the causes of stimuli. How are these predictions generated? According to the predictive processing account, the brain performs Bayesian inference to select the models, or priors, that best predict the causes of sensory input. This means that the prediction with the highest posterior probability is the prediction that is selected and that determines the perceptual content of the system.²⁴ In the Bayesian prediction model of the brain, stable, internal hypotheses have two important constituents: prior probability and likelihood.²⁵ The hypothesis with the highest prior probability is the hypothesis that informs top-down predictions about incoming sensory stimuli.²⁶ If the prediction corresponds well to the incoming stimuli, the likelihood of the input increases. High prior probability combined with high likelihood implies high posterior probability.

21 Ibid., 183.

22 Ibid., 185.

23 Hohwy, 2010, 136.

24 Hohwy *et al.*, 2008, 688.

25 Clark, 2013, 185.

26 Hohwy, 2010, 136.

In the binocular rivalry case, the predictions of the stimuli have a roughly equal likelihood but if, for some reason, one stimulus has a higher likelihood, then it will be selected for perceptual dominance.²⁷ The conjoint prediction (of both stimuli) has a higher *likelihood* because it can predict both stimulus X and stimulus Y but it has a much lower *probability* because it is *a priori* very improbable for the two objects to occupy the same space.^{28,29} The brain has learned that only one stimulus can cause sensory input at a given place and time.³⁰ In other words, given prior probability and likelihood, only one stimulus will be perceived at that time. The stimulus that is perceived is the one with the highest likelihood given other contextual information.

The predictive processing account suggests that the probability density of information is encoded rather than the information.³¹ What this means is that when we perceive an object, the nature of the object is not represented in the brain but the relative probability of the nature of the object is encoded. This function enables the brain to deal with uncertainty, noise and ambiguity.³² For example, to encode the event TREE IN THE FOREST, the Bayesian brain will encode the conditional probability density function of this event using top-down and bottom-up processes and presents an approach where the general scene is first identified and the details are filled out.³³ The general picture is presented and initially informed by internal hypotheses, which are selected through Bayesian inference. In other words, perception constitutes a process that follows the principle: forest first, then trees.

It seems as though this is a computationally costly process but this is not the case. Neurons are not suddenly turned on when new information is encountered because the brain is already active with a large set of

27 Hohwy *et al.*, 2008, 690–691.

28 Hohwy's use of the term *a priori* is not the same as the traditional philosophical use of the term. What he means is that the brain's models are *a priori* in the sense that they encounter particular experiences after they have propagated down the hierarchical levels.

29 Hohwy *et al.*, 2008, 691.

30 Ibid., 691.

31 Clark, 2013, 188.

32 Ibid., 188.

33 Ibid., 188.

priors, which impact the processing of new information.³⁴ This means that a large amount of context information is already in place when new information is encountered; at each level of the neural economy, predictions are probabilistic and take into account uncertainty at each stage of the processing.³⁵ This illustrates the Bayesian nature of the cognitive system. When the system encounters new bottom-up information, the neuronal responses are expected to change significantly given the contextual information provided by top-down predictions.³⁶ The view that perception is a task of top-down predictions, more than it is of sensory stimulation is supported by Rauss *et al.* (2011) who claim that "...neural signals are related less to a stimulus per se than to its congruence with internal goals and predictions, calculated on the basis of previous input to the system."³⁷

3.2. The function of the system

The predictive processing account of cognition also makes an important claim about the function of cognition. It states that perception involves top-down predictions interacting with bottom-up signals from the environment. The discrepancy between top-down predictions and bottom-up signals, called prediction error, is a result of the content that is entered into the system through sensory input but that is not predicted by higher-level processes.³⁸ It is the function of the system to reduce this error by transmitting the discrepancy/prediction error backward in order to update predictions in higher levels of the hierarchy. When prediction error is minimized, the accuracy of likelihood and prior probability of predictions increase and posterior probability is increased. However, if the prediction is poor, the entire input signal is transmitted as prediction error. A process involved in prediction error minimization, is the updating of higher-level hypotheses through lower-level input.³⁹ In addition to updating the hypotheses that feed predictions, prediction error can also

34 Clark, 2013, 188.

35 Knill & Pouget, 2004, 713.

36 Clark 2013, 189.

37 Rauss *et al.*, 2011, 1249.

38 Hohwy 2010, 135.

39 Friston 2003, 1341.

be minimized through action. This supports the claim that perception and action function to reduce prediction error and expresses the view that it is the function of biological organisms to reduce prediction error.

4. Action-oriented predictive processing

Clark proposes that action also functions to minimize prediction error. One way in which action enables the minimization of prediction error is by changing how an agent samples the environment rather than changing its expectations of the environment.⁴⁰ An animal moves its sensors and body, and acts on the environment in ways that can be described as active sculpting of the environment to match expectations.⁴¹ Action, it can be said, exhibits a similar profile to perception but reduces prediction error by eliciting movements that change bottom-up information.⁴² This means that prediction error is used to adjust action to minimize the discrepancy between the consequences of action and that, which is predicted.⁴³ Clark's action-oriented predictive processing account makes two claims about perception and action. The first is that perception and action work together to move an organism through time and space.⁴⁴ The second states that perception and action follow the same principles; this is based on the premise that both perception and action are implemented through the same computational strategies.⁴⁵

Clark's first claim is that perception and action work together to reduce prediction error. In other words, prediction error is minimized through the selective sampling and active sculpting of stimuli.⁴⁶ Perception functions to update internal models and expectations about the causes of stimuli, and action works to fulfill these expectations. Friston illustrates:

⁴⁰ Hohwy *et al.*, 2008, 690.

⁴¹ Clark, 2013, 186.

⁴² Ibid., 186.

⁴³ Friston, 2003, 1349.

⁴⁴ Clark, 2013, 186.

⁴⁵ Ibid., 186.

⁴⁶ Ibid., 186.

"Perceptual learning and inference is necessary to induce prior expectations about how the sensorium unfolds. Action is engaged to resample the world to fulfill these expectations. This places perception and action in intimate relation and accounts for both with the same principle."⁴⁷

The second claim put forth by Clark's account is that perception and action are computationally similar. This claim is captured by Eliasmith (2007) who supports the notion because "The best ways of interpreting incoming information via perception, are deeply the same as the best ways of controlling out-going information via motor action...".⁴⁸ What Eliasmith has in mind is something such as a Kalman filter which functions by predicting system states and updating these predictions through incoming information. In a Kalman filter, there are mechanisms that adjust the weight of a model's predictions depending on the measurement of the prediction error. This implies that both perception and action are driven by predictions and expectations⁴⁹ and that both perception and action function to fulfill the expectations of the system.

Where does this leave us in terms of action selection? Clark does not discuss action selection in his article but he offers an account of action control by looking at the free energy principle. The free energy principle proposed that the biological imperative of living organisms is to minimize free energy.^{50,51} In terms of action, the claim is that an agent acts on the world and adapts accordingly in order to fulfill its expectations. Implied in this claim is the view that the selected action is the action that has the highest likelihood of fulfilling an agent's expectations or that can shape the environment in such a way that expectations are matched. The function of cognition can be understood in terms of prediction-error minimization because, according to the free energy account, living or-

47 Friston *et al.*, 2009, 12.

48 Eliasmith, 2007, 380.

49 Clark 2013, 186, 200.

50 I refer here to the minimization of free energy but this can be translated as minimizing prediction error. In other words, prediction error is free energy.

51 Friston, 2011; Friston *et al.*, 2012.

ganisms have a neural and biological imperative to minimize prediction error.⁵²

5. The efficiency problem

Clark's account of cognition does not consider action selection and neglects discussion on efficiency, motivation and reward. There are three features to take into account when considering action selection.⁵³ First, the outcome of an action provides information about future outcomes. The state of the environment is a source of information, which can be correlated with different variables and enables organisms to move through space and time. The information that is found in the relationship between the first and second variables can be the outcome of an action (reward or punishment). Reward or punishment, i.e. the outcome of an action, carries important information about the action that is selected and performed.

Second, there are several possible actions to choose from and some are better than others in that they result in more efficient behaviour and carry positive outcomes. Animals with flexible behaviour can learn about better and worse responses to features in the environment; this can be achieved through associative learning. Behavioural responses can be shaped over a variety of timescales and, on an evolutionary timescale, the agent may not know which is the better option but evolution selects agents that make better choices. Efficiency is not the only element that makes one action better than another positive outcomes (rewards) also affect action selection. Each outcome has a consequence, which is translated as either a reward, or a punishment; an animal would learn to associate certain rewards and punishments with selected actions.⁵⁴ If the outcome of the selected action is rewarding, the action will be repeated when the opportunity arises, and if the outcome is not rewarding, the action should be avoided. This leads to the third element that is important in action selection and particularly the efficiency problem. Each

52 Friston 2011; Friston *et al.*, 2012; Clark 2013.

53 Houston, McNamara and Steer, 2007, 1533–1534

54 Sterelny, 2003, 82.

possible action is accompanied by consequences, which are taken into account when actions are selected. Sterelny proposes that behaviour is a result of the motivation that arises from the combination of variables, both internal and external. Clark does not take into consideration the elements discussed above and as a result his account faces a serious objection called the dark room problem.

6. The dark room problem

Friston and Clark run into a serious concern regarding the behaviour of living organisms because of the claim that organisms have a sole biological imperative to fulfill expectations. The dark room problem is a *reductio ad absurdum* objection which states that accepting the claim that all living organisms have the sole biological imperative to minimize free energy leads to a peculiar, and false, consequence. This can be illustrated as follows. If the predictive processing account is correct about the function of cognition, all living organisms with cognitive capacities will seek to minimize surprise. One way of achieving this is for an organism to block out all sensory information by seeking out the least stimulating environment and staying there.⁵⁵ In such an environment, a living organism will be deprived of sensory stimuli and prediction error will be minimal, if not absent. This seems absurd since it is empirical knowledge that almost no living organisms inhabit monotonous environments or seek out situations that have no sensory stimuli.⁵⁶ Even organisms that do inhabit “dark rooms” such as caves seek out sensory stimuli by making use of other mechanisms, such as proprioception and echolocation. But this is not the primary claim of predictive processing; predictive processing is based upon the free energy principle.

According to the free energy principle, living organisms have a biological imperative to minimize prediction error.⁵⁷ This follows from the

55 Friston *et al.*, 2012, 2.

56 One exception may be Troglophiles and Troglobites (animals that complete their entire life cycle inside a cave environment). Troglophiles can also survive in above ground habitats unlike Troglobites who are often transparent (or white) and completely blind (Chapman, 1982).

57 Friston, 2011; Friston *et al.*, 2012.

principle embedded in the free energy formulation: avoid surprises and you will last longer.⁵⁸ Seeking out monotonous environments that contain the least amount of surprises enables the organism to keep prediction error to the minimum.⁵⁹ One would then expect that the predictive processing account proposes that it would make sense for organisms to seek out and stay in a dark room. As a result, the predictive processing account will also be committed to the view that organisms can successfully minimize surprise by blocking out sensory information and seeking out a monotonous environment, free from surprise. Such an environment would result in the organism being in the ultimate stable state with near perfect predictions.⁶⁰ This view is peculiar and empirically false; all living organisms with cognitive capacities do not seek out dark, monotonous corners to inhabit; the view is also nonsensical from an evolutionary and biological point of view.

What is required of Friston's account is to provide an account of free energy minimization, where organisms have an imperative to minimize surprise, but also supports the view that animals do not simply seek out a nice dark corner and stay in it.⁶¹ In the following sections, I look at Friston's response to the dark room problem and his attempt to explain what is meant by the dark room metaphor. Thereafter, I examine Clark's response to Friston's account and also critically examine Clark's response to the dark room problem. Clark and Friston's responses to the objection have important differences but they both offer solutions to the consequences of the free energy principle. Their responses exemplify two general solutions that can be offered in response to the dark room problem. First, Friston refutes the *reductio ad absurdum* argument and supports the view that the consequences of minimizing prediction error are not as absurd as they seem. Second, Clark proposes that the consequences of the account do not follow and he provides additional considerations regarding prediction and expectation of surprising activity in changing and challenging environments.

58 Friston *et al.*, 2012, 2.

59 Ibid., 2.

60 Clark, 2013, 186.

61 Ibid., 193.

6.1. Friston's response to the dark room problem

Friston responds to the dark room problem by saying that dark room agents do, in fact, exist and that we need to look no further than Troglophiles for such examples.⁶² In other words, Friston refutes the *reductio ad absurdum* argument by saying that the free energy minimization principle and its consequences do not reduce to absurdity because some organisms do seek out monotonous environments that are free from sensory stimuli. Hence what Friston means when he offers the dark room metaphor is not that all living organisms should seek out a dark corner to stay in, instead organisms should avoid surprise which is measured in relation to an organism's model of the world.⁶³ He goes on to say that solutions to minimize prediction error "...will be unique to each conspecific and its econiche".⁶⁴ For example, what is surprising to Troglophiles may not be surprising to primates and *vice versa*. According to Friston, living organisms have a biological imperative to reduce surprise – both on a short term and long term basis.

It is not obvious (though Friston might disagree) that the sole biological imperative of living organisms is to minimize surprise because all organisms with cognitive capacities do not seek out and inhabit monotonous environments. In other words, Friston's account cannot be accepted as a general account of cognition because not all organisms seek out monotonous environments. To illustrate the criticism of Friston's account more clearly: the account is not an attempt to make a prediction about *some* animals that inhabit dark caves; it makes the claim that *all* living organisms function to minimize prediction error. But animals do not seek out monotonous environments, and instead participate in challenge and exploration to satisfy needs and drives. Animals are reward-seeking and participate in activities that have positive outcomes and tend to avoid activities with negative outcomes – learning about the consequences of actions often calls for challenge, exploration and exposure to surprise, even if just initially. Friston disagrees; he claims that "...valuable states are unsurprising..." and since the agent has already learned about the

62 Friston *et al.*, 2012, 2.

63 Ibid., 2012, 1.

64 Ibid., 2012, 2.

valuable states and updated its predictions, those states are already available in the agent's environment.⁶⁵ Rewarding activities are valuable to the welfare and survival of animals because evolution favours organisms that are efficient and successful in their behaviour as response to the environment.

Friston's response to the dark room problem suggests that the consequences of the free energy minimization principle is not as bad as it seems and he attempts to rescue the account by offering an example of organisms that do in fact inhabit dark environments free from sensory stimuli. It seems as though Friston misses the point of the objection because he comes up with rationalisations for some organisms but that's not what is required. What Friston's account promises is a general account of cognition that makes predictions about the function of cognition and a rationalisation that can be generalised, but the promise is not delivered. In the next section, I discuss Clark's response to the dark room problem. Clark also refutes the *reductio ad absurdum* objection to the dark room problem; he argues against the consequences and suggests that the free energy account does not run into the dark room problem because animals should not, and will not, seek out monotonous environments free from sensory stimuli because they expect to engage in change, challenge and exploration.⁶⁶

6.2. Clark's response to the dark room problem

Clark's response is different to Friston's in that Friston confirms that the consequences follow the free energy account but he plays down the significance of the consequences; Clark, instead, claims that the consequences do not follow. Clark proposes that animals do not seek out monotonous environments free from surprise because animals expect change and exploration. In other words, in response to the dark room problem, Clark acknowledges that living organisms live in different and changing environments and proposes that organisms that live in changing and challenging environments will not seek out monotonous environments because they have expectations about change and exploration

65 Friston, 2009.

66 Clark, 2013, 193.

and furthermore expect to act out sophisticated strategies and perform cognitively complex actions.⁶⁷ The potoo bird, for example, has adapted to a complex environment such as the Amazon rainforest. The potoo's actions reflect the complexity of the environment in that it has adapted to mimic its environment in order to blend in and protect itself from predators. The potoo's feeding and reproductive behaviours are strategic and unique to its environment. The notion behind this is that organisms adapt to their environments and minimize surprise in relation to their environments, i.e. organisms in complex environments have adapted to complex environments and organisms in monotonous environments have adapted to such conditions. From this follows that the world models and priors an agent holds influence that which is surprising to an agent, and the agent's prior expectations, and that which determines surprise, are specific to different species and are unique to the individual.⁶⁸ Clark's argument implies that animals are models of their environment and their predictions are formed to be compatible with the environments they inhabit. Implied in this claim is that when an animal adapts to its environment, it learns and updates its priors. Clark supports the view that perception and action are primarily driven by the need to minimize free energy. From an evolutionary approach, this is limiting; adaptive fitness depends not only on minimizing free energy but also on reproduction and drive satisfaction – these activities are efficiently performed when the value of the outcomes are learned. Reproduction and drive satisfaction requires risky and changing behaviour, and associating the reward or punishment of an outcome with an action enables animals to discriminate between, and prioritize, actions. A serious concern raised against the predictive processing account is that the appeal to incentive and reward to explain behaviour is largely replaced by constructs of prediction and expectation.⁶⁹ Clark acknowledges the concern by asking whether the predictive processing account of cognition omits much of "...what really

67 Clark, 2013, 193.

68 Friston *et al.*, 2012, 3.

69 Clark, 2013, 200.

matters for adaptive success: things like boredom, curiosity, play exploration, foraging and the thrill of the hunt?".⁷⁰

Clark responds to the dark room problem by saying that "...change, motion, exploration and search are themselves predicted – and poised to enslave action and perception accordingly".⁷¹ He supports the notion that search and exploration are key features of life⁷² and that animals explore the environment to feed and reproduce because without exploration, the animal will perish and the species will die out. Clark's response to the dark room problem suggests that animals that have expectations about change and exploration will find monotonous environments surprising. To illustrate this informally, the animal will be surprised by its own hunger and lack of activity, and will be prompted to explore the environment. Clark's response is adequate, as far it goes, in that exploring the environment is advantageous because the animal can feed, reproduce and find information about the environment but notably missing from Clark's response is an account of action selection and motivation. Clark does not propose that animals are motivated to satisfy hunger or to reproduce and neglects discussion on incentive and reward; instead Clark proposes that an animal is motivated to explore and find information so as to not be surprised by its own inactivity. For example, an animal in a dark room will eventually get thirsty; it has two general options, relevant to this discussion, to reduce the unstable state of being thirsty. First, it can seek a waterhole or food and quench its thirst but this would involve surprise and exploration. Second, it can update its prior and expect thirst but this means the animal will perish. Without paying consideration to motivation and reward, Clark's account supports the latter (absurd) option. Therefore, the lack of discussion on incentive or reward in Clark's account is not suitable to an evolutionary or biological account of behaviour because it provides no insight into motivation to act and how to discriminate between strategies.

According to evolutionary accounts of action and cognition, action is the result of both external states of the environment and a motivation

70 Ibid., 193.

71 Ibid., 193.

72 Friston *et al.*, 2012, 1.

to change or attain a particular state in the environment.⁷³ Clark's study of predictive processing offers no account of the motivation that living organisms have to seek change, initiate motion and engage in exploration.⁷⁴ Without such an account, the neural and biological strategy stays the same – reduce surprise.⁷⁵ Clark's response (that change and exploration is expected) is therefore not satisfactory because it lacks an account of a mechanism for selecting actions that, not only promotes fitness, but that enables discrimination between actions and supports the notion of seeking reward. Such a mechanism of motivation is required especially in an environment that delivers signals, which are noisy and somewhat unreliable as well as functionally ambiguous.⁷⁶

7. Neuroeconomic model of motivation

Clark responds to the dark room problem by asserting that the consequences of accepting the free energy minimization principle do not follow and that the account does not reduce to absurdity because animals expect change and exploration and therefore do not seek out monotonous environments. The response given by Clark does not fit what is known about evolution and biology because it is suggested by evolutionary models of cognition that animals seek reward and Clark's account neglects any discussion on reward and motivation. Reward-seeking is a fundamental property of behaviour and have been recognized as such by many models of cognition.^{77,78} Importantly, reward values are not merely attached to the outcomes of actions but are predictions that are updated by making use of experience.⁷⁹ By introducing the notion of value to Clark's predictive processing account, I do not refute his account

73 Sterelny, 2003, 79.

74 Huebner, 2012.

75 Friston *et al.*, 2012, 5.

76 Sterelny, 2003, 81.

77 Barron, Søvik and Cornish, 2010.

78 Notably, Sherrington, 1906; Tinbergen, 1951; and Lorenz, 1965; and more recently Arias-Carrión, Ó., & Pöppel, E., 2007; Dickinson, A. & Balleine, B., 2002; Glimcher, P. W., & Sparks, D. L., 1992; and Glimcher, P. W., Dorris, M. C., & Bayer, H. M., 2005.

79 Spurrett, 2012.

but offer an improvement of the account because if Clark's account had taken into consideration the problem of action selection, and the role of value in action selection, the account would not run into the dark room problem.

Neuroeconomics is aimed at providing models of decision making by relating economics, psychology and neuroscience.⁸⁰ The goal of neuroeconomics (according to some) is to ground economic theory in neuroscience; this can be achieved through studying the economic theoretical framework and conducting experiments in neuroscience. Finding neuroscientific evidence that can describe choice behaviour will also enable us to make predictions about behaviour and decision making. Many traditional models used to describe action and behaviour are rooted in the classical view that behaviour is a reflex or response to the environment. This view has been supported, and studied, by many scholars, particularly behaviourists, such as Skinner and Pavlov. The starting point of developing the theory of decision making, proposed in this dissertation, is the notion that behaviour is not merely a reflex or response to the environment but that behaviour is a result of taking into consideration the expected outcome of the selected action or choice and then selecting the action with the highest expected reward. Experiments by Glimcher & Sparks (1992), Schall & Hanes (1993), Basso & Wurtz (1997) and Platt & Glimcher (1999) are some of the many experiments that show that the processes that connect sensory and motor systems involve processes other than classical reflex mechanisms. What Glimcher offers is a solution to the action selection problem that also provides an account of the neural architecture and mechanisms involved in action selection and decision making. The approach proposed by Platt & Glimcher (1999) has two classes of input: current sensory data and stored representations.⁸¹ From this follows that, decision making, according Glimcher's approach, involves a combination of current sensory data and the agent's best estimate of the outcome of an action. This view differs from Clark's in that the estimates of the outcome of an action are combined with a loss

80 Glimcher, 2011, 35.

81 Platt & Glimcher, 1999, 233.

function that specifies the value of all possible losses or gains;⁸² these are represented in the brain and influence decision making.

Glimcher supports the view that animals act to maximize a utility function and goes on to say that utility has a neural correlate, which he calls subjective utility.⁸³ Subjective utility is represented in the brain and an important element in action selection. The view that the value of an outcome is represented in the brain has been supported and tested by many scholars; of particular interest are the experiments done by Platt & Glimcher (1999), Kable & Glimcher (2007), Schultz (1998), and McClure *et al.* (2004). The scope of this paper does not allow for discussion of all the mentioned experiments but Schultz's experiment in the late nineties is of particular interest. Jacques Mirenowicz and Wolfram Schultz developed an experiment to study the activity of single dopamine neurons. The experiment involved monkeys participating in a Pavlovian conditioning task.⁸⁴ The purpose of the experiment was to explore the relation between dopamine neurons, reward and reward prediction. Dopaminergic neurons in certain brain areas, such as the ventral tegmental area, have a firing profile that corresponds to reward prediction error.⁸⁵ In Mirenowicz and Schultz's experiment, thirsty monkeys were placed in front of a spout. At random intervals, a tone was produced; the tone functioned as a conditioned stimulus (CS). After a short delay, a drop of juice was dripped from the spout; this was the reward (R). The objective of the experiment was to study the relationship between spikes in firing rates of dopamine neurons during the presence of the conditioned stimulus and during the delivery of reward. Dopamine neurons in monkeys have a baseline firing rate of three to five spikes per second. Initially, the firing rate of dopamine neurons remained constant when the tone (CS) was produced and increased when the juice (R) was dripped from the spout.⁸⁶

After a few trials, the frequency of dopamine neuron activation increased at the time the tone was presented (CS) and returned to baseline

82 Platt & Glimcher, 1999, 233.

83 Glimcher, 2011, 136.

84 Ibid., 301.

85 Shea, 2012, 6.

86 Glimcher, 2011, 301.

when the juice (R) was delivered. This indicates that the firing rates of dopamine neurons decreases in response to the reward and increased in response to the conditioned stimulus. If the juice (R) was delivered without the experimenter producing a tone (CS), the dopamine neurons responded to the reward which indicated that the reward had not lost its ability to activate dopamine neurons. In the final experiment, the tone (CS) was produced but the juice (R) was not delivered. The finding was that dopamine neuron activationdropped below baseline at the time of expected reward in the trials when the tone was produced but the juice was not delivered.⁸⁷ The experiments conducted by Mirenowicz and Schultz illustrate that reward predictability plays an important role in dopamine neuron activation. If the reward is predicted by the presence of a conditioned stimulus, dopamine neurons have a higher firing rate at the time of conditioned stimulus presentation than at the time of reward presentation. If the reward is not predicted, dopamine activation occurs at the time the reward is presented. Dopamine neurons report "...rewards according to the difference between the occurrence and the prediction of reward...".⁸⁸

Predicted rewards can be innate or learned. We learn to associate certain rewards with stimuli when the system reduces reward prediction error. Reward prediction error occurs when the reward is not fully predicted by the conditioned stimulus.⁸⁹ The process can be described as follows: the system processes an event, predictions are generated, and the difference between the prediction and actual event is computed.⁹⁰ The system makes use of prediction errors to optimize performance and predictions. This illustrates that the process of updating reward predictions is the same process as updating prediction error in Clark's predictive processing account. There is strong evidence that reward prediction error plays a role in learning and guides decision making in many

87 Glimcher, 2011, 302.

88 Schultz, 1998, 7.

89 Ibid. ,11.

90 Ibid., 12.

species.⁹¹ From this follows the inference that evolution has set up living systems to maximize the delivery of reward to the agent.⁹²

8. Conclusion

The action selection problem asks the question, how does an agent select a particular action at a particular time? In this paper, I proposed that the action selection problem can be approached in two ways. The first approach asks the question, why are actions selected? The second asks, how are actions selected and involves discussion of the architecture, mechanisms and processes involved in cognition. In response to the second approach, I proposed Clark's action oriented predictive processing account – a novel approach to the predictive processing account. The predictive processing account proposes that the brain is bidirectional and hierarchical in nature and that the brain performs Bayesian inference to select predictions of the causes of stimuli. The account also makes an important claim about the function of the brain. It proposes that the main function of the brain is to reduce prediction error. Prediction error is the discrepancy that arises as a result of the interaction between top-down and bottom-up information. One way in which prediction error can be reduced is by seeking out a surprise-free environment and staying there. This solution is biologically absurd because animals that do not explore the environment will perish.

Clark's general account of cognition lacks an account that explains how animals discriminate between strategies of action and how certain actions can be better or worse. It also lacks discussion on the motivations that prompt behaviour and does not offer an account of how the values of actions are represented. This is not to say that it is always necessary for the value of an outcome to be represented; motivation can be based on various internal drives and sensations.⁹³ This means, for example, that water will have a higher value to a thirsty animal than to a sated animal. Animals with robust and flexible cognitive mechanisms are

91 Shea, 2012, 17.

92 Ibid., 20.

93 Sterelny, 2003, 79.

able to update their goals according to the value of a reward.⁹⁴ Animals with simple response systems may act only on sensations and internal drives.⁹⁵ The predictive processing account is largely correct about the mechanisms involved in perception and action but the account requires a theory of value because learning, motivation and decision making are all processes that require valuation⁹⁶ and are important elements in the story of cognition. This means that although selecting the action that minimizes prediction error may fulfill an agent's expectations, adaptive fitness is dependent on more than just minimization of prediction error. A living organism increases the probability of adaptive fitness through generating efficient responses to the environment; this requires taking into account the value of the outcome of an action and the cost of acting to attain the outcome.⁹⁷ Evolutionary accounts of action provide insight into how cognition has evolved to enable animals to select actions that are efficient by taking into account the value of the outcome of an action and the cost of acting to attain the outcome. Though this insight is valuable, also required is an account of the mechanisms and architecture involved in the process of complex cognition because the action selection problem constitutes the efficiency problem and the architecture problem.

The response given by Clark does not fit what is known about evolution and biology because it is suggested by evolutionary models of cognition that animals seek reward and Clark's account neglects any discussion on reward and motivation. Reward-seeking is a fundamental property of behaviour and have been recognized as such by many models of cognition.⁹⁸ Importantly, reward values are not merely attached to the outcomes of actions but are predictions that are updated by making use of experience.⁹⁹

By introducing the notion of value to Clark's predictive processing account, I do not refute his account but offer an improvement of the

⁹⁴ Huebner, 2012.

⁹⁵ Sterelny, 2003, 79.

⁹⁶ Huebner, 2012.

⁹⁷ Ibid.

⁹⁸ Barron, Søvik and Cornish, 2010.

⁹⁹ Spurrett, 2012.

account because if Clark's account had taken into consideration the problem of action selection, and the role of value in action selection, the account would not run into the dark room problem. Finally, I proposed a neuroeconomic account of value and illustrated the compatibility of the model by making use of Schultz's experiment providing strong evidence that reward prediction error plays a role in learning and guides decision-making in many species.¹⁰⁰

References

- Arias-Carrión, Ó. & Pöppel, E. 2007. Dopamine, learning, and reward-seeking behavior. *Acta neurobiologiae experimentalis*, 67 (4): 481–488.
- Barron, A. B. Søvik, E. & Cornish, J. L. 2010. The roles of dopamine and related compounds in reward-seeking behavior across animal phyla. *Frontiers in behavioral neuroscience*, 4 (163).
- Basso, M. A., & Wurtz, R. H. 1997. Modulation of neuronal activity by target uncertainty. *Nature*, 389 (6646), 66–69.
- Biederman, I. 1987. Recognition-by-components: a theory of human image understanding. *Psychological review*, 94 (2), 115.
- Brown, H., Friston, K. & Bestmann, S. 2011. Active inference, attention, and motor preparation. *Frontiers in psychology*, 2 (218).
- Chapman, P. 1982. The Origins of Troglobites. *Proceedings of the University of Bristol Spelaeological Society* 16 (2): 133–141.
- Clark, A. 2012. Dreaming the Whole Cat: Generative Models, Predictive Processing, and the Enactivist Conception of Perceptual Experience. *Mind*, 121 (483): 753–771.
- Clark, A. 2013. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36 (3), 181–204.

100 Shea, 2012, 17.

- Clark, A. 2014. Perceiving as Predicting. In Mohan, M., Biggs, S., & Stokes, D. (eds.). *Perception and Its Modalities*. New York: Oxford University Press.
- Dickinson, A. & Balleine, B. 2002. The role of learning in the operation of motivational systems. In C. R. Gallistel. (ed). *Steven's handbook of experimental psychology: Learning, motivation and emotion*. Volume 3. 3rd edition. New York: John Wiley & Sons.
- Eliasmith, C. 2007. How to build a brain: From function to implementation. *Synthese*, 159 (3): 373–388.
- Friston, K. 2003. Learning and inference in the brain. *Neural Networks*, 16 (9): 1325–1352.
- Friston, K.J., Daunizeau, J., & Kiebal, S.J., 2009 Reinforcement Learning or Active Inference? *PLoS ONE* 4 (7): e6421: 1–13.
- Friston, K. 2011. Embodied inference: Or I think therefore I am, if I am what I think. In Tschacher, W. & Bergomi, C. (eds.). *The implications of embodiment: Cognition and communication*. Exeter: Imprint Academic.
- Friston, K. & Stephan, K. E. 2007. Free-energy and the brain. *Synthese*, 159 (3): 417–458.
- Friston, K., Thornton, C. & Clark, A. 2012. Free-energy minimization and the dark-room problem. *Frontiers in psychology* 3 (130): 1–7.
- Glimcher, P. W. 2011. *Foundations of neuroeconomic analysis*. New York: Oxford University Press.
- Glimcher, P. W. & Sparks, D. L. 1992. Movement selection in advance of action in the superior colliculus. *Nature*, 355 (6360): 542–545.
- Godfrey-Smith, P. 2002. Environmental complexity and the evolution of cognition. In Sternberg, R. & Kaufman, J. (eds.). *The evolution of intelligence*. Mahwah: Laurence Erlbaum Associates.
- Hohwy, J. 2010. The hypothesis testing brain: Some philosophical applications. In Christensen W., Schier, E. & Sutton, J. (eds.). ASCS09: *Proceedings of the 9th Conference of the Australasian Society for Cognitive Science*. Sydney: Macquarie Centre for Cognitive Science.
- Hohwy, J., Roepstorff, A. & Friston, K. 2008. Predictive coding explains binocular rivalry: an epistemological review. *Cognition*, 108 (3): 687–701.

- Houston, A. I., McNamara, J. M. & Steer, M. D. 2007. Do we expect natural selection to produce rational behaviour? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362 (1485): 1531–1543.
- Hubel, D. H. & Wiesel, T. N. 1965. Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *Journal of neurophysiology*, 28 (2): 229–289.
- Huebner, B. 2012. Surprisal and valuation in the predictive brain. *Frontiers in psychology*, 3 (415).
- Kable, J. W. & Glimcher, P. W. 2007. The neural correlates of subjective value during intertemporal choice. *Nature neuroscience*, 10 (12): 1625–1633.
- Knill, D. C. & Pouget, A. 2004. The Bayesian brain: the role of uncertainty in neural coding and computation. *TRENDS in Neurosciences*, 27 (12): 712–719.
- Lorenz, E. N. 1982. Atmospheric predictability experiments with a large numerical model. *Tellus*, 34(6): 505–513.
- Maes, P. 1990. Situated agents can have goals. *Robotics and autonomous systems*, 6(1): 49–70.
- McClure, S. M., Laibson, D. I., Loewenstein, G. & Cohen, J. D. 2004. Separate neural systems value immediate and delayed monetary rewards. *Science*, 306 (5695): 503–507.
- Platt, M. L. & Glimcher, P. W. 1999. Neural correlates of decision variables in parietal cortex. *Nature*, 400(6741): 233–238.
- Rauss, K., Schwartz, S. & Pourtois, G. 2011. Top-down effects on early visual processing in humans: a predictive coding framework. *Neuroscience & Biobehavioral Reviews*, 35(5): 1237–1253.
- Riesenhuber, M. & Poggio, T. 2000. Models of object recognition. *Nature neuroscience*, 3: 1199–1204.
- Schall, J. D. & Hanes, D. P. 1993. Neural basis of saccade target selection in frontal eye field during visual search. *Nature*, 366 (6454): 467–469.
- Schultz, W. 1998. Predictive reward signal of dopamine neurons. *Journal of neurophysiology*, 80 (1): 1–27.

- Shea, N. 2012. Reward Prediction Error Signals are Meta-Representational. *Noûs*, 48 (2): 314–341.
- Sherrington, C. S. 1906. *The Integrative Action of the Nervous System*. Connecticut: Yale University Press.
- Spurrett, D. 2012. What is to be Done? Why Reward is Difficult to Do Without. *Frontiers in psychology*, 3: 412.
- Sterelny, K. 2003. *Thought in a hostile world: The evolution of human cognition*. Oxford: Blackwell Publishing Ltd.
- Tinbergen, N. 1951. *The study of instinct*. Michigan: Clarendon Press.