Intelligence as Accurate Prediction

Trond A. Tjøstheim¹ · Andreas Stephens¹

Accepted: 7 March 2021 / Published online: 16 March 2021
© The Author(s) 2021

Abstract

This paper argues that intelligence can be approximated by the ability to produce accurate predictions. It is further argued that general intelligence can be approximated by context dependent predictive abilities combined with the ability to use working memory to abstract away contextual information. The flexibility associated with general intelligence can be understood as the ability to use selective attention to focus on specific aspects of sensory impressions to identify patterns, which can then be used to predict events in novel situations and environments. The argumentation synthesizes Godfrey-Smith's environmental complexity theory, adding the notion of niche broadness as well as changes concerning the view of cognition and control, and Hohwy's predictive mind theory, making explicit the significance of accuracy as a composite of trueness and precision where the nervous system acts as a distributed controller motivating actions that keep the body in homeostasis.

Keywords Intelligence · General intelligence · Prediction · Prediction error minimization · Free energy · Context

1 Introduction

Although 'intelligence' is a widely used concept in everyday language, researchers have struggled to unite around a common definition. One obvious reason might be that intelligence often is interpreted as a broad concept which includes a number of different abilities that are realized by a variety of neural mechanisms. In this paper we aim to examine a specific interpretation of intelligence, focusing more broadly on animals than humans specifically. This implies foregoing treatment of notions associated primarily with human intelligence research, such as IQ tests and Spearman's *g* factor (Spearman 1904, 1928).

✉ Andreas Stephens
andreas.stephens@fil.lu.se

¹ Department of Philosophy and Cognitive Sciences, Lund University, Lund, Sweden

In order to get an overview of the plethora of definitions of ‘intelligence’ that exist, Legg and Hutter (2007a; see also Legg and Hutter 2007b; Yousefian et al. 2016) review 70 definitions and conceive that they, plausibly, can be placed in three main categories: collective (encyclopedic) definitions; psychologist definitions; and artificial intelligence (AI) researcher definitions. From this material, Legg and Hutter distill three commonly held features where intelligence: “Is a property that an individual agent has as it interacts with its environment or environments; Is related to the agent’s ability to succeed or profit with respect to some goal or objective; Depends on how able the agent is to adapt to different objectives and environments” (Legg and Hutter 2007a, p. 9, bullet list in original). They proceed to synthesize their findings into a definition: “Intelligence measures an agent’s ability to achieve goals in a wide range of environments” (Legg and Hutter 2007a, p. 9).

We consider Legg and Hutter’s (2007a) review and discussion to be very illuminating. Nevertheless, we find it – as well as most discussions of intelligence – lacking an explanation regarding just *how* biological systems actually achieve their goals (in a wide range of environments). Hence, in this work, we hope to elucidate the natural phenomenon of intelligence, i.e. the cognitive processes underlying goal-achievements.

First steps in this direction can be found in investigations of intelligence as predictive ability where “the core of intelligence is the ability to anticipate and predict variation and novelty and to devise strategies to cope with this novelty” (Geary 2009, p. 22). We will in this article investigate intelligence, and general intelligence, by defending two theses:

- 1) The natural phenomenon intelligence can be approximated as accurate prediction.
- 2) The natural phenomenon general intelligence can be approximated as the ability to abstract away context information, identify patterns, and transfer accurate predictions across contexts, as well as the ability to perform mental transformation and comparison.

We believe that this is a fruitful focus since organisms,’ including humans,’ ability to distinguish between patterns in its environment that are unstable, short-term stable, or long-term stable will have an effect on its ability to predict an uncertain future, in turn affecting its survival (see e.g. Plotkin 1993; Cellucci 2017). Intelligence offers a way for an organism to do this. Above we make a distinction between intelligence and general intelligence.¹ This distinction may be made more concrete by considering e.g. differences in problem-solving and reversal learning ability in urban and rural finches, as reported by Audet et al. (2016). Reversal learning is the ability to inhibit and update a previously learned rule to accommodate changes in a task’s reward structure (Lissek et al. 2002). They found that urban-dwelling birds were better at both problem-solving and reversal learning than were country-dwelling birds. This is interesting since what characterizes an urban environment compared to a rural one is typically its disturbance

¹ We recognize that ‘intelligence’ and ‘general intelligence’ can be a somewhat confusing use of terminology. Particularly since we intend ‘intelligence’ to be a broader term than ‘general intelligence.’ It would perhaps be less confusing if we had used ‘context general intelligence’ and ‘context specific intelligence’ vs the more abstract ‘intelligence.’ The reason for the current choice is however that both ‘intelligence’ and ‘general intelligence’ are typically more widely used terms, and hence attempting to elucidate their semantic differences and commonality appeared to us more fruitful.

levels by traffic, human pedestrians, and industrial noise, as well as its variance over time (Lowry et al. 2013). To prosper in such an environment, species hence have to display cognitive flexibility. Contexts may change rapidly, e.g. when buildings are replaced or areas refurbished, but food-gathering skills can be adapted. In the countryside, environments tend to be more stable. Country-dwelling birds are also able to successfully predict their environment, and hence display intelligence. But the predictive skills displayed by urban-dwelling species are less context specific, or more general, than those of their country-dwelling cousins. Hence, we use the term ‘intelligence’ to be a more abstract term, covering accurate prediction as such, be it context-specific, or context general. In other words, both finch types exhibit intelligence (accurate prediction), although urban-dwelling birds exhibit more generalized intelligence (an ability to accurately predict, spanning a broader variety of contexts).

2 Background

Related to the ability of an organism to perform accurate predictions is the concept of ‘prediction error’ minimization (Hohwy 2013) which we will discuss further below. In the context presented here, prediction error minimization can be thought of as the principle underlying the learning mechanism of an organism. In Hohwy’s (2013, p. 1) words, the brain can then be seen as “a sophisticated hypothesis-testing mechanism, which is constantly involved in minimizing the error of its predictions of the sensory input it receives from the world.” Given our theses, this means that intelligence is at the very center of both how we perceive and act. On evolutionary time scales, the environment will then pressure the information processing capabilities of an organism to take on sufficient structure and complexity to track and predict patterns pertinent to survival and reproduction.

This might be reformulated into the maxim that to stay alive, organisms need to maintain homeostasis. While the external world imposes numerous signals on our senses, only some of these signals are of use and translatable into action in the service of regulation (Bruineberg et al. 2018). Prior beliefs are (mostly unconsciously and through multiple hierarchical levels) used to make top-down predictions that then get tuned by the world’s bottom-up feedback reaching the system as error signals. Such tuning is possible to accomplish by way of error minimization in two ways; as a change of the internal model or by changing the world (through action):

Perceptual inference is a matter of hierarchical prediction error minimization. Top-down predictions are compared to the actual sensory signal, or with the inference of the level below, and the difference, the prediction error, is used as a feedback signal for the internal models generating the predictions. (Hohwy 2013, p. 59)

The way the brain accomplishes these feats have both faster and slower elements:

Prediction error driven supervision by the world can happen via representations of quickly changing states of the world and via representations of more slowly changing parameters of the world. The former is thought to occur in the brain’s

synaptic activity (the way neurons interact in the short term), and the latter in synaptic efficacy (the way connection strengths are set over the longer term) [...]. (Hohwy 2013, p. 49)

According to Godfrey-Smith, cognition has the main function of enabling a system to deal with the complexity of the environment (Godfrey-Smith 1996, 2017a, b). Interestingly this also means that there is a vast cognitive complexity in regard to how different animals' cognitive faculties have evolved, highlighting how there is a varying ability to use reflective cognitive abilities to circumvent reflexive behavioral responses. He elaborates on this point by claiming that dealing with the environment is done by the notion of organisms controlling their actions in response to their environment, and further that the point of cognition is to supply a means for exercising this control (Godfrey-Smith 2002). Cognition is seen as a form of information processing.

To make 'information' more specific Godfrey-Smith draws on Dretske (1981) and his understanding of information as a resource. Organisms can then use this resource for making their way around their niche, and going about their business. However, it may be unclear in precisely which way information can be regarded as a resource. The environment may present various stimuli to the organism in the form of light, vibrations in air, or shifting concentrations of molecules. If the organism has sensors that can accommodate these stimuli in the form of light sensitive proteins, or structures that are sensitive to mechanical vibration, and transform them into action potentials in the case of organisms with nervous systems, or intracellular molecular cascades in the case of bacteria, then these stimuli can be used to change behaviour in a way that keeps the organism in homeostasis. In this sense, the stimuli, or information that the organism can extract from its environment is a kind of resource, albeit one that is not necessarily spent, like energy is spent. Rather, this can be framed as the, novelty, or surprise of the environment, being consumed by the organism (see e.g. Friston 2009). The difference between what the organism predicts and what the environment serves up is thus being reduced in the process of prediction error minimization (Hohwy 2013). As an example, consider the case of bacterial chemotaxis (see e.g. Falke et al. 1997). To a bacterium in the search for nutrients, the environment is filled with varying concentrations of different kinds of molecules. Some of these molecules are nutrients, and bacteria have evolved to move towards higher concentrations of them. This means that the bacterial niche can be seen as a gradient field, where the gradient is the spatially varying concentration of chemicals. However, the bacterium may be in a place where no gradient can be found, and nutrient concentrations are random. Connecting back to the notion of information as a resource, in a place of random concentrations, informational content is low. The environment, in other words, does not give the bacterium any clue as to in which direction to move to maintain its homeostasis. Bacteria get out of such places by 'tumbling,' or initiating random movement (Falke et al. 1997). When a nutritional gradient (i.e. information about where to go) is found, the bacterium stops tumbling and instead smoothly moves in the direction of increasing nutrient concentration (Falke et al. 1997). In this way, the bacterium can be viewed as 'consuming' the informational resource of its environment in the service of finding nutrients and maintaining homeostasis.

Godfrey-Smith finally specifies complexity as 'heterogeneity.' This might be interpreted in more mathematical terms as (statistical) variance. Now, intelligence, in

most creatures, is reasonably seen as being present in a matter of degree; in other words, intelligence ought to be seen as a gradual, rather than an all-or-nothing, affair (see e.g. Godfrey-Smith 2017b; see also Klein and Barron 2016). In order to predict its environment, an organism thus has a limited amount of information from its sensors, but also limited information processing capacity to exploit the information it does get. To allow prediction further ahead in time generally requires higher memory capacity, since predictions often are a form of remembering, or matching a currently experienced sensory sequence to one experienced in the past. To allow more broad matching requires that memory becomes less specific, or more abstract. It may also require that the organism can direct attention to different sensory streams and match those. That is, matching may be partial across the sensorium. For example, the sound of something approaching may spell bad news for a grazing animal even if there is no smell of predators in the air, and no predator is seen.

We will mainly use the term ‘accuracy’ in relation to intelligence in this work. In accordance with Menditto et al. (2007), this term will be taken to mean a composite of ‘trueness’ and ‘precision,’ where the former refers to how close an average observation is to an accepted reference value, and the latter refers to how close repeated independent measures are to each other. In the context of cognitive systems making predictions, we interpret these terms to mean in the case of trueness the correctness of categorization, and in the case of precision the level of correct detail, or the narrowness of a statistical distribution. For example, a predator might predict the appearance of prey animals by a watering hole on a given time of day. The prediction might be true in the sense that prey does appear, in contrast with non-prey. It might also be precise, in that a particular species of prey appears (e.g. an antelope vs. a zebra), and at the particular predicted time of day (e.g. in the morning vs. the afternoon).

3 Accurate Prediction

We will support our first thesis (The natural phenomenon intelligence can be approximated as accurate prediction) by two main arguments: the first considers adaptive processes in nature; the second considers the necessity of organisms to negotiate obstacles and solve problems.

3.1 Argument 1.1: Adaptation

In this argument we aim to show that organisms must predict, and the accuracy of predictions affects organisms’ ability to survive and procreate.

The totality of biological processes, along with rocks, soil, water, and the atmosphere, make up the earth’s biosphere. The biosphere is made up of a myriad of individual organisms, both plants and animals, and microorganisms, all interacting with each other. From the perspective of an individual organism, its environment is a “landscape of affordance” (Bruineberg and Rietveld 2014), which can be called an econiche (Bruineberg and Rietveld 2014). This implies that an organism’s econiche affords ways of interaction, dependent on that organism’s body and information processing system. Interaction implies also that

organism and econiche affect each other. For example, animals can affect the course of a river by means of eating from trees at the river side; the river will affect rocks and soil of the landscape by means of erosion, but will also affect the behaviour of animals who drink from it, and their walking trails. Through natural selection organisms adapt within their ecological niche over generational time, where the selection process, by acting on the genetic code of the organisms, shapes both somatic and mental processes of the organisms in question (Geary 2005).

The econiche contains a wealth of information that can be described as gradient fields. A gradient field is a field with changing values, akin to landscape with hills and valleys (McMullin 2002). For example, the photons of different wavelengths that bounce around, can be described as an electromagnetic field, while the magnetic photons generated by the earth's spinning metal core can be described as a magnetic field. Even the mass of the earth's crust can be described as a field. For brevity, we will forthwith use the term 'field' as a metaphor for the densities of photons, heat, or mass that can be interpreted and used by an organism. The point here is that this information is available for an organism to exploit, and take advantage of to improve its chances of survival. However, the sense in which we use the word 'information' needs to be made more specific here. We assume that an organism is coupled to its environment through a finite number of sensors, each of which has a finite number of stable states it can be in. Hence we use Shannon's (1948) notion of information here, and its complement Shannon 'entropy' for a measure of the uncertainty faced by an organism.

For an organism to take advantage of a niche's information it must possess sensory apparati suitable for converting the varying field gradients into a form amenable for internal processing. Through natural selection, organisms adapt to sense the fields that are of use to them by expressing proteins that can act as transducers (Cohen et al. 1995). Examples include the proteins that make up photon receptors for light, proteins that change when exposed to pressure, and proteins that change with vibrational frequency.

One defining property of organisms is that they are bounded; although they are part of their niche, they are also somehow distinct from it. They have an outside, and an inside as it were, and the biological processes that go on inside the organism must be maintained. This process of continuous maintenance, of keeping the internal temperature at the necessary value, of keeping an adequate supply of energy and nutrients, is called homeostasis (Mason et al. 2017, pp. 864–887). Survival implies maintaining homeostasis. This has several implications for how the organism processes information from its environment. We will look at three such implications.

Given that an organism has evolved the ability to express proteins that can track changes in some gradient fields that surround it, the signals produced by those proteins must somehow be connected to the organism's motor system. In this discussion, we take the meaning of sensations to be understood as the kind of behaviour it elicits (Atlan and Cohen 1998; Powell and Still 1979). This is motivated by the findings that associate percepts with behaviour through perception-motor loops in organisms' central nervous systems. This can be compared to the notion of meaningful features in an environment as those that are perceived as affordances, where affordances can be thought of as "opportunities for action" such that the sensations that are meaningful

are precisely those that convey affordances. Hence, organisms tend to evolve reflexive connections between sensory signals and motor behaviour. Certain signals, or combinations of signals will automatically trigger particular motor sequences. Coarsely, such motor patterns can be classified as approach or avoidance behaviour. A shark smelling blood in the water will change its course and swim along the concentration gradient of the blood, towards its source; an antelope smelling a lion on the wind will run along the concentration gradient of lion odor chemicals, but away from its source.

To be able to act on signals from the environment implies that organisms have evolved some form of discrimination, or comparison mechanisms. There are an infinite number of patterns available in an organisms' environment, but it will only approach or avoid a certain subset of them. Although the shark has evolved a keen appreciation of blood in the water, it is likely less interested in chemicals that signal rotting seaweed. Similarly for the antelope; it can recognize the smell of lions, but will likely be indifferent to the smell of fresh fish. Sensory signals are constantly filtered and compared to innate and learned patterns in the process of motivating behaviour. But organisms primarily recognize patterns that signal threat, food, or mates.

Complementary to the ability to sense and perceive patterns relating to survival and procreation is a behavioural repertoire that can take advantage of those patterns. The shark must be able to swim towards, and intercept prey. It must also be able to eat the prey once it gets to it. To survive, an organism must be able to evade threats; it must have evasive behaviour. The antelope can run very fast. It is also agile, and can change direction rapidly. Its behaviour in escape is not easily predicted. This is an important point for our thesis, and we will return to predator-prey interactions below.

So far, we have established that organisms live in a world of physical gradient fields, and have evolved sensory apparati to track those fields, but also behaviour to take advantage of them. Next, we will explore what the above implies for intelligence, and how rates of change play a part.

Although an organism's sensors may track the physical fields in its environment, those sensors still have a frequency range over which they function best. Outside of this range, both slower and quicker, the sensor's accuracy will attenuate. Sensing alone is not adequate for survival, the organism must also act on the basis of what it perceives. Behavioural activation is associated with a time delay in which signals flow around the organism's nervous system to settle on which of several possible behaviours to engage in. The motor system then has its own delays which are dependent on the mass of the organism and the amount of power its motor system can generate, and how quickly it can go about such activation. Hence, if an organism's internal processes, including sensory-, information processing-, and motor processes, have a slower sensory-motor cycle time than the rates of change of its surrounding environment, the organism is forced to resort to prediction to be able to maintain homeostasis. A classic example from cybernetics is the predictor for anti aircraft gunfire (Ashby 1961). The gun cannot aim at where the aircraft is at the time of fire, but at the place where the aircraft is predicted to be when the shell intersects with the aircrafts movement vector. Likewise, a lion must predict as best it can the apparently random jumps of its prey, or it will tire out before securing its meal. In other words, purely reactive processes will in many cases be too slow for an organism to adequately keep itself fed, or evade predation.

Predicting sensory signals implies that organisms can also predict the effect on the world of their own actions. According to Stepp and Turvey (2010), this prediction may

be in the form of model simulation, but may also be the result of the organism being coupled to its environment in the sense of being driven by it. They call the former ‘weak anticipation,’ and the latter ‘strong anticipation,’ where the “strength” is due to oscillator coupling and physical law, rather than learned statistics. The mechanism responsible for the coupling is feedback delay between a driver- and a driven system. Here we are primarily concerned with weak anticipation, since oscillator coupling according to Stepp and Turvey becomes weaker as timescales lengthen. Hence we are concerned with situations that are too noisy, uncertain, or too long term to allow coupling. In such situations, a memory that allows storing sensory sequences may be more important, as well as the ability to somehow play back those sequences at a rate faster than the sensory input. But this playback version may not always exactly match signals from the real world. In other words, predictions have accuracy associated with them. From this follows that the accuracy of an organism’s predictions affects how well it can choose the action that optimally changes the world to its benefit. Regarding behaviour, this may have two implications. The first is that organisms should evolve to store behavioural sequences adequate for changing the environment to the organisms’ benefit. Behaviour can be stored in the genome, either directly as hardwired structures of the organism, or indirectly as signal pathways that are biased such that the behaviour tends to be expressed. An example of the former is bacterial flagella (Barrett 2015, p. 20), while the evolution of aversive responses to likenesses of predator eyes in birds (e.g. De Bona et al. 2015) is an example of the latter. A second implication is that behaviour may be stored in the organism’s nervous system, and that the organism is able to learn adequate behavioural sequences. Behaviour stored in the genome is adequate if the environment is approximately constant during the lifetime of the organism. If the environment is not stable, the organism must be able to learn and change. In this case, the flexibility of a plastic nervous system will tend to be selected for. Requirements for learning rises sharply with organism complexity, such that most multicellular organisms, including plants (Baluška and Mancuso 2018) have some form of system in place for behavioural adaptation.

The above notions can also be related to behaviour as yielding epistemic value through active inference as described by Friston et al. (2015). It is also related to the notion mentioned above, of an econiche as a landscape of affordances (Bruineberg and Rietveld 2014) as well as of resources. Evolutionary processes may tune the sensitivity of an organism’s reward system to the reward for novelty, i.e. epistemic value. If tuned up, the organism will tend to be more motivated to expand its roaming area. If tuned down, the organism will tend to stay in a more limited spatial area and exploit the resources that are available there. Hence the resource abundance of a particular place over time may dynamically affect sensitivity to epistemic value and its complement, risk. These changes may happen in the organism’s lifetime by physiological tuning in the reward system, or over evolutionary time by natural selection and epigenetic processes that turn relevant genes on or off.

Summing up, organisms in environmental niches that have processes that change faster than the organism’s internal processes can adequately track, will tend to evolve nervous systems that can predict those processes. The accuracy of the predictions directly affect probabilities of successful behaviour, and hence survival and procreation rates.

3.2 Argument 1.2: Problem Solving and Obstacle Negotiation

In this section we argue that the process of problem solving, of which negotiating obstacles, can be regarded as a special case, is central to survival and procreation for organisms in non-stable environmental niches. We also argue that accurate prediction is necessary for successful obstacle negotiation.

An ecological niche to which an organism is an integral part, can be described, from that organism's point of view as patterns of resource availability, and patterns of obstacles to those resources. Here, we consider threats like predators also as obstacles. In an environment with rich resources and few patterns of obstacles, there will be little evolutionary pressure to develop behavioural flexibility beyond direct approach. The environment around hydrothermal vents is an example of such a niche. Here, particularly adapted bacteria convert abundant CO₂ and Hydrogen emitted from vents on the ocean floor, into methane (Sleep and Bird 2007). For most multicellular organisms, however, getting resources takes effort, and obstacles abound. Negotiating obstacles becomes central to survival and procreation. Evolution tends to facilitate processes central to survival and procreation by making them intrinsically motivating. Eating when hungry and drinking when thirsty are two examples for such processes. For organisms in shifting niches, the process of negotiating obstacles can thus itself become rewarding (Špinka and Wemelsfelder 2011).

Anything that hinders an organism in approaching rewarding stimuli like food or mates can be considered an obstacle. Likewise, obstacles can also hinder escape from a threat. One salient property of obstacles is that they require effort and energy expenditure to overcome, which, had the obstacle not been there, could be saved. What counts as obstacles are specific to context and individuals. For an antelope with an injured foot, a felled tree can be a difficult obstacle to overcome, whereas a healthy antelope can easily jump over it. In a more artificial setting, a crow attempting to get food by figuring out a mechanism set up by human experimenters, is too facing an obstacle. But since the effort involved is mental rather than physical, such obstacles tend to be called by the more abstract term problems. Refining the concept of obstacles somewhat, we now argue that obstacles can usefully be considered as being on a continuum from adverse to non-adverse. Non-adverse obstacles cannot intentionally² hinder or threaten an organism, nor can they make predictions of the organism to hinder it. Examples of non-adverse obstacles are the felled tree mentioned above, features of geography like mountains, or atmospheric phenomena like storms and hurricanes. Adverse obstacles on the other hand are capable of prediction and behavioural flexibility. Most predators are adverse obstacles, but conspecifics like rivaling males in a chimpanzee troop, or crows that hide and move food around to keep other crows from stealing it, are also examples of adverse obstacles. We claim that there is a difference in the mental effort and capacity required to negotiate adverse and non-adverse obstacles. This is due to their difference in complexity, and predictability.

Predictive abilities in organisms are subject to evolutionary pressures, in the same way as are features of the body, like the length of the legs, or the size of the teeth. Prediction can increase the probability of successful negotiation compared to trial and error. From the perspective of energy expenditure it will tend to be more efficient, since engaging the motor

² We use 'intentionally' in a non-strict way here that does not necessarily imply subjective consciousness.

system tends to cost more energy than engaging neural populations (Mink et al. 1981). But properties of a particular ecological niche will determine whether the cost of producing the necessary brain structures are worth it. In an energy dense niche with few complex obstacles, the evolutionary pressures to improve prediction will be less than in a niche where energy is sparse, and must be kept from meddling neighbors. However, and more fundamentally, the energy density of a niche will determine whether an organism can at all develop and sustain the neural structures necessary for sufficient prediction. If obstacle negotiation also requires cooperation, coalition building, and strategizing, the pressure to expand predictive abilities can become very strong, since successful prediction of conspecifics can have outside rewards. This appears to have been the case with humans (Brüne and Brüne-Cohrs 2006).

In a predator-prey interaction between two organisms, both organisms pose adverse obstacles for each other. The predator will actively pursue, and attempt to stop and kill the prey. The prey on the other hand will actively avoid and try to sabotage the predator's path to a meal. In this situation, the prey species will be pushed towards evolving behaviour, particularly escape behaviour, that is harder and harder to predict. The predator species, to survive must therefore compensate by either developing strategies where the unpredictable behaviour of the prey is less important, or it must become better at predicting. The former may be a product of the latter, as may be seen in pack hunting situations, such as when lions surround and collectively stalk antelopes, obviating their agility (Stander 1992), or dolphins chase prey into a barrier of conspecifics thus limiting the use of speed and random behaviour as escape tactics (Gazda et al. 2005). Such ratcheting, or arms-race dynamics is common for coupled predator-prey species, and may also include the efficiency of lethal venom in the predator versus the resistance of the prey for that venom (Holding et al. 2016), or development of spikes in plant species versus the toughness of the mouth in herbivores (Hawley et al. 2007). This is often referred to as a 'Red Queen' dynamic, after the Lewis Carroll character in *Alice in Wonderland*, which had to "run just to stay in place."

Given that obstacles display complexity, and that a continuous scale of complexity exist where the simplest non-adverse obstacles are at one extreme, and the most cunning adverse obstacles are at the other extreme, we propose that accuracy of predictions follow a similar scale. If negotiating an obstacle can be achieved by a wide selection of behaviours from an organism's behavioural repertoire, the precision of predictions is less important. A felled tree in the path can be jumped over, walked over, or perhaps walked around. As the complexity of obstacles increase, however, finer distinctions must be made, and perhaps only a very few behaviours are adequate for success. To take an example from the human world, a banker might wonder: will the financial deal go through? Only if the right documents have been carefully prepared, and the customer is treated with exactly enough friendliness, respect, and authority.

In this section we have argued that ecological niches differ in the demands they place on prediction, depending on the availability of energy, and the complexity of obstacles that hinder survival and procreation in that niche. Regarding obstacle complexity, we have made a distinction between whether obstacles can be considered intentionally adverse or not. The former place heavier requirements on predictions than the latter. We have also argued that obstacle complexity affects how accurate predictions are required to be to ensure successful negotiation of those obstacles. Higher complexity requires higher accuracy in terms of both correct categorization and precision of that categorization.

So far we have considered context dependent, or specialized intelligence and predictive abilities. We will now turn to our second thesis, that of general intelligence.

4 Abstraction Allows for Transfer between Contexts

Our second thesis stated that: The natural phenomenon general intelligence can be approximated as the ability to abstract away context information and transfer accurate predictions across contexts, and further approximated by the ability to perform mental transformation and comparison.

The first part of this thesis will be supported by the argument that organisms that gather energy from more than a single niche will increase their probability of survival by being able to carry over patterns learned in one niche to a different niche. Since learned patterns are necessarily bound to specific contextual information due to the structure of the nervous system (Smith and Mizumori 2006), freeing patterns from their context requires abstraction (Behrens et al. 2018). The cost of context independence is less precision. But as long as the generalized pattern confers predictive abilities that save energy compared to trial and error, we propose that the development cost of brain structures that do abstraction is paid for, and those structures will therefore be positively selected for.

The second part will be supported by the argument that the ability to mentally try out behaviour and simulate situations without physically carrying them out, confers advantage in terms of energy conservation.

4.1 Argument 2.1: Broad Environmental Niche

Context can be understood as everything in an organism's complete sensory input, that is not the focus of attention. By default, behaviour is associated with context. This is due to the fundamentally associative nature of nervous systems that support Hebbian learning where coincidentally firing neurons become associated through synapse formation and strengthening (Hebb 1949). From the arguments presented above, it follows also that predictions likewise are associated with particular contexts.

Although the term 'environmental niche' is difficult to precisely define, we will borrow from statistics, and take it to mean the distribution of sensory signals that an organism encounters. This is equivalent to Godfrey-Smith's notion of environmental complexity mentioned above. Niches then exist on a continuum from narrow to broad (or low to high complexity), with narrow niches displaying distributions with less variance than do broad niches. To make this more concrete, let us consider again bacteria living in hydrothermal vents on the ocean floor. The variance across their sensors are much less than the variance across the sensors of a crow living in a human city. The bacteria sense heat in the water, and concentrations of nutrients; both are relatively stable qualities. The crow, on the other hand, can experience very large variance, particular across its optical sensors. Depending on its habits, it can see human dwellings both close up and far away, as well as parks, and perhaps forests outside the city. Hence, the crow's niche is at the broad end of the scale. As tends to be the case, humans can experience the most variance, living anywhere from the equator to the arctic.

Taking an evolutionary perspective, the niches occupied by species have been growing broader. The earliest organisms were likely similar to the hydrothermal bacteria, while at the present time, many species are forced to broaden their niches due to human behaviour (Luniak 2004; Birnie-Gauvin et al. 2016). We suggest that an important driver for niche broadening is resource depletion. To maintain their energy requirements, organisms may be forced to move. This movement exposes them to different opportunities, but also different obstacles.

There are organisms that occupy broad niches, but have separate life phases associated with different contexts. A butterfly starts life as a caterpillar eating leaves off the tree on which it was born. But it metamorphoses, and changes niche irreversibly. Across its lifespan, the two niches are stable and narrow enough that behaviour and context can be stored genetically, without incurring the costs of a flexible nervous system. In contrast, for organisms that employ the same basic body structure throughout their lives, such flexibility is often required. Given that behaviour and context patterns are stored in a distributed fashion in the nervous system (Manns and Eichenbaum 2006), an implication is that the two may be disassociated.

Disassociating behaviour from context requires abstraction (Munakata et al. 2011; Brass et al. 2005). This means that the brain is structured in a hierarchical way, allowing general patterns to be maintained such that they become independent of a particular context (Behrens et al. 2018; Christoff and Keramiatian 2007).

Context information may be important for a prediction's accuracy. Removing this information will necessarily reduce the accuracy of the prediction. In the following we will argue that 1) an abstract, generalized prediction is more accurate than a random guess, and 2) that behaviour elicited from a generalized prediction is more allostatically optimal³ than is uninformed trial-and-error behaviour. First, research on memory indicates that inaccuracy may be adaptive. Semantic memory as defined by Klein et al. (2002) is inherently inaccurate and has a summary nature. This inaccuracy is adaptive in terms of being generalized and accommodating of (robust to) variations (Klein et al. 2002). Schacter et al. (2007) review evidence that memory is prospective⁴; its function appears to be to support decision-making in the present by allowing simulation and prediction of future states. There are indications that several species, including birds (Bobrowicz et al. 2019; Kabadayi and Osvath 2017) and great apes (Osvath and Osvath 2008; Bobrowicz et al. 2020) can make predictions not only in contextual similar situations, but also based on functional similarities. That is, they can make decontextualized, or generalized predictions (Bobrowicz 2019, p. 102). From this we propose that an abstract, generalized prediction can be more accurate than a random guess.

Second, behaviour both physical and mental carries metabolic cost. In the context of foraging, trial and error can be interpreted as investing metabolic energy (a trial) in order to harvest energy. Trials are risky and can fail (error). Environments differ in terms of the metabolic costs they impose on behaviour, as well as their richness in terms of resources available for foraging (Molles and Sher 2018, p. 175). Now,

³ 'Allostasis' is defined as the maintenance of homeostasis. By 'allostatically optimal' we intend behaviour that rewards maximal energy per energy invested and hence can optimally maintain allostatic processes.

⁴ 'Prospection' can be thought of as a kind of mental behaviour that consists in experiencing, or simulating the future (Gilbert and Wilson 2007)

organisms tend to optimize the rate of energy harvested from their environments (Stephens and Krebs 1986). One avenue for optimization is using experience, simulation, and prospection to bias foraging behaviour towards targets that have been successful in the past (Schacter et al. 2007), and to utilize prospection to minimize costly (Redish 2016) or risky (Castegnetti et al. 2020) behaviour. A complementary avenue for optimization is the ability to inhibit behaviour that could be dangerous or disadvantageous in the current situation (Kabadayi et al. 2016; van Horik et al. 2019). From this we propose that behaviour elicited from a generalized prediction can be more allostatically optimal than is uninformed trial-and-error behaviour. Given 1) and 2), being able to disassociate and abstract away context from predictions is likely to be evolutionarily advantageous for an organism.

To sum up, in this section we have argued firstly that, organisms can be forced to broaden their niche due to resource depletion, and that such niche broadening also exposes the organism to a broader range of obstacles. Secondly we have argued that as long as context and behaviour are associated, but stored in a distributed fashion across an organism's nervous system, neural populations that represent contextual information can be abstracted away. This abstraction frees both behaviour and predictions to be used in novel contexts that still have some similarity with the original one, albeit with some accuracy penalty. This abstractive process, we suggest, is the mechanism mediating the generality of general intelligence. Thirdly we have argued that as long as the generalized predictions are accurate enough to allow behaviour to be more cost effective than random trial and error, the ability to abstract and generalize will be advantageous for the organism, and hence will tend to be selected for.

Next, we will look closer at how selective attention becomes important as simulations stretch longer into the future. We will also talk about how manipulating information in working memory can be used for prediction, to simulate manipulation of objects and carrying out actions before carrying them out physically.

4.2 Argument 2.2: Working Memory Transformation and Comparison

Flexibility in assigning attention is necessary for effective working memory transformations. There are several ways attention can be used in the context of mental simulation. One is focusing on a particular property to vary it, like size, colour, or location (Olivers et al. 2006). Another is the selection of one particular goal, among many competing ones (Corbetta and Shulman 2002). An organism may at the same time be hungry, thirsty, and bored. But it may not be possible to satisfy all these wants at the same time. Often a particular one will be strong enough to dominate, but at other times one goal must be selected by means of top down attention. The longer ahead an organism can predict and plan, the more necessary attentional selection of intermediate goals becomes, to satisfy longer term goals. Coupled to this ability is complementary upregulation of structures that can inhibit impulse behaviour and enforce self control (Figner et al. 2010). This is necessary to maintain approach behaviour towards a goal for extended periods of time.

Sometimes, it is necessary for an organism to simulate behaviour that involves manipulating objects in the environment, i.e. to engage in a kind of prospection as mentioned above.. For problem solving situations which particularly involves moving and rotating objects, working memory appears to confer a set of virtual effectors.

Combined with the ability to remember and prospect, these virtual effectors confer extended problem solving abilities by constructing ad-hoc models of the environment and obstacles within it; these models can be used to coarsely try out behaviour to simulate effects, without physically carrying out the actions. It would hence likely carry evolutionary benefits for an organism to have access to such a mental world (Craig 1943). We suggest that such models likely synthesize patterns learned from experience such that e.g. physical laws like gravity can be included. However, experiments in humans (Sohn et al. 2002; Strauß et al. 2014), but also in animals (Dias et al. 1996, 1997; Birrell and Brown 2000) indicate that inhibitive populations can also be controlled volitionally. Inhibitive mechanisms allow particular aspects of experience to be isolated and reapplied to arbitrary novel objects. In practice this means that an organism with a suitably equipped working memory can learn to e.g. mentally rotate arbitrary objects after having watched a particular object being rotated. Depending on the complexity of the working memory networks, it is likely that arbitrary transformations can be learned in this way. For example, Lara et al. (2009) showed that rhesus monkeys appear to have gustatory working memory, which indicates that transformation can be done on taste representations. Aside from the well studied phenomenon of mental rotation, Resnick and Shipley (2013) observe that geologists learn to do brittle transformations in working memory, while Atit et al. (2013) studied mental twisting transformations.

A central ability when doing mental transformations like those mentioned above, is that of comparison. This allows matching transformations to a target, and simulation towards a specific valued outcome. Typically such outcomes are either the reduction of a homeostatic drive like hunger or thirst, or accomplishing an intermediary goal towards such a drive reduction (Reeve and Lee 2012).

Working memory manipulation can be tied to prediction error minimization in two ways. First as a means of prediction generation, the error of which will be minimized by means of active inference (Pezzulo et al. 2016; Friston et al. 2015). Second, as part of the allostatic process of minimizing interoceptive prediction errors by means of foraging. We propose that working memory transformations here works as a kind of vicarious trial and error (Redish 2016), but for manipulation tasks, figuring out mechanisms, and tool use (e.g. Haidle 2010; Seed and Byrne 2010; Orban and Caruana 2014). This is in contrast to vicarious trial and error in ambulation towards a place. But like in path finding we expect use of working memory for simulation to be limited to situations where physical trial and error is comparatively costly (Potts et al. 2018).

To sum up, we have argued that working memory can be used to assemble ad-hoc models for testing out behaviour and simulating situations, and for comparing simulated outcomes with specific valued outcomes. These abilities can be used to save energy and avoid risky situations, and hence confer evolutionary advantage in niches with adequate energy density. As organisms become able to predict further ahead in time, judicious use of selective attention becomes increasingly necessary. Attention thus allows organisms to control impulses and work towards longer term goals.

5 Implications and Discussion

In this section, we first compare our theory to Godfrey-Smith's environmental complexity theory, and second to Hohwy's predictive mind theory. In the third subsection,

we discuss the difference between magnitude and generality of intelligence. Then, in the fourth subsection we discuss implications of our theory with regard to how intelligence is commonly measured. We acknowledge that our theory does not at present cover all aspects of intelligence, and we discuss some of these aspects in the fifth subsection.

5.1 Intelligence as Prediction vs the Environmental Complexity Thesis

Our work builds on Godfrey-Smith's environmental complexity thesis in several ways. Firstly, we assume, like him, that intelligence and cognition is for controlling action, whether that action be physical by means of the motor system, or mental by means of working memory and associated systems. Secondly, we adopt his evolutionary view that whatever functions mechanisms for processing information have, those functions are subject to natural selection depending on their effects. That is, information processing is functional in terms of increasing probability of survival or of procreation. Related to this is the concept of cognition, which we agree with Godfrey-Smith can be understood as abilities for coordinating behavioural with environmental patterns (Godfrey-Smith 2002). This includes learning the reward structure of an environment. Thirdly, our notion of niche broadness concurs with Godfrey-Smith's concept of environmental complexity as (statistical) variance. However, we chose 'niche' as a way of emphasizing that an environment may consist of many niches, each with their own energy densities and obstacles.

We attempt in this work to refine the environmental complexity thesis via our conception of intelligence as prediction, and detailing of context specific versus context independent prediction. It might be opportune at this point to clarify how intelligence as prediction relates to the concept of cognition. If cognitive abilities exist on a scale that mirrors niche complexity, then both different kinds of abilities may be called for, as well as different resolution of those abilities. This is perhaps clearest when considering perception. This ability helps an organism discern categories of patterns in an environment, and hence choose suitable behaviour to match. The nature of sensors developed by the organism allow tracking of different gradient fields, as explained above. But organisms may differ in how they are pressured to develop perceptive resolution, or visual acuity, based on those sensors. The sparrow and the eagle can both see adequately, but what counts as adequate is different in each case. The acuity of the eagle is likely several times that of the sparrow. In a similar sense, what is required to predict patterns in one niche is insufficient in another. We would hence refine Godfrey-Smith's argument that cognition is for controlling action, and argue that, choice of optimal behaviour is what counts, predictions facilitate this choice, and some cognitive abilities like perception, attention, and memory, facilitates prediction. Others abilities, like executive control, become helpful as goals extend beyond the present and inhibition of immediate gratification confers advantage.

5.2 Intelligence as Prediction vs the Predictive Mind Thesis

Hohwy's thesis of the predictive mind (Hohwy 2013) is another building block used in this work. Particularly, we adopt Hohwy's view on the brain as a prediction making organ. We consider his explanation of how mental models of the world are updated, as

dependent on the difference between what is predicted by the models, and what evidence is presented by sensors as plausible. Hohwy proposes also details about how the central nervous system is hierarchical in its organization, and that precision estimates are being performed at the various levels to determine to which degree models are updated based on incoming sensory signals..

We agree that the models generated by the cognitive system interacting with the world can be used for causal inference, as the predictive mind thesis proposes. There is an interesting way in which both causal inference and prospection might be seen as outcomes of the same process. In both cases, this involves the generation of a path between two positions in time and states of the world. But in the former case, the target point is the present, while in the latter case the target is some future time. However, both cases can be served by using a generative model of the world and the self, which may include synthesizing memory traces with manipulations in working memory. The difference is that for causal inference, moving along the path implies moving backward in time, while for prospection, the direction is moving forward. Predictive ability should serve each case.

This work complements the predictive mind thesis by interpreting predictive abilities specifically as approximating the intelligence concept. The distinction between context specificity and context independence implies that the predictive mind must include particular mechanisms for encoding context, as well as mechanisms for abstracting away that context. To complement Hohwy's inclusion of precision on predictions, we make explicit the significance of accuracy in terms of both trueness and precision. This is implicit in the notion of prediction error minimization, but individual organisms may differ in their ability to cope with prediction errors, impacting their predictive abilities. Further, the issue of precision is implicit in our thesis concerning general intelligence as an ability to abstract and transfer predictions to new contexts, but with a cost of reduced precision.

The notion of hierarchical perception referred to by Hohwy can be integrated with notions of working memory and executive control in the following way. According to Chanes and Barrett (2016) the central nervous system can be viewed as a hierarchy where limbic structures can be conceptualized as being on top, and receive prediction errors from nuclei in the brain stem that regulate and predict homeostatic variables like e.g. energy and nutrient availability, and body temperature. Limbic structures involving the hippocampus, can associate rewarding nutrients with places (e.g. Kanoski and Grill 2017). These structures may then induce images of such places into working memory (Chanes and Barrett 2016), effectively turning them into goal-sites for approach. As approach behaviour commences, executive functions contribute by maintaining the goal in working memory (e.g. Bastos et al. 2018), as well as inhibition of impulsive behaviours that would lead the organism away from the decided on goal (Carr and Dweck 2011). Further down the hierarchy are various motor and sensory structures that continuously predict the environment and correct behaviour to maintain approach until consumption of the goal is completed.

5.3 Obstacle Complexity vs Generality

A distinction needs to be made between accurate prediction of complex obstacles, and generality of intelligence. That is, the *level* or *magnitude* of intelligence versus how *broad* it is. In this section we will propose that magnitude of intelligence is related to the complexity of typical obstacles within an ecological niche and that obstacle complexity not necessarily pressures organisms to evolve general intelligence.

In argument 1.2 above, we proposed that accurate prediction is necessary for successful obstacle negotiation, and that increasingly complex obstacles require more accurate categorization of sensory information to choose the correct behaviour. Here, we would like to draw attention also to the number of properties of an obstacle that must be taken into account to accurately predict it. Typically, more complex behaviour requires more properties to be simulated than does less complex behaviour to achieve sufficiently accurate prediction. From a mathematical perspective each changing property corresponds to a differential equation to be solved; properties may also be mutually coupled, which implicates a set of coupled differential equations to be solved by the predicting system. This is not to imply that an organism's brain solves differential equations in the same way as does a computer, only that demands on a cognitive system should be expected to increase in correlation to the complexity it is required to model. The complexity of the cognitive system should again be expected to have a higher magnitude of intelligence. However, there are usually several avenues available for organisms to avoid directly simulating the full complexity of an adverse obstacle. Some strategies have been mentioned above with lions hunting in packs and dolphins forcing prey into shallow water. Predators may also become highly specialized such that they adapt to the specific complexities of a prey species. Hence, predator species may be able to cope with intricately behaving prey by becoming more narrowly intelligent, rather than more general.

An intriguing example of intelligent and flexible obstacle negotiation that is still specialized is *Portia*, a species of jumping spider (Cross et al. 2020). *Portia* displays a preference for preying on other spider species (Richman and Jackson 1992). This means that its prey might be categorized as an adverse obstacle, and *Portia* must cope with the risk of itself becoming a meal when hunting. The size of *Portia*'s brain is small, and can fit on the head of a pin, but these spiders display impressive behavioural flexibility. For example, *Portia* often invades the web of other spiders to capture them. According to Cross et al. (2020) this includes manipulating vibration frequencies on webs such that the resident spider calms down and refrains from attacking. An interesting aspect of this is the degree to which perturbation rhythms must be fitted not only to different prey species but even to individual spiders. *Portia* uses a kind of generate-and-test algorithm in this process and can thus flexibly adjust perturbations to an individual prey to elicit behaviour that increases likelihood of successful capture. Further, *Portia* is known to be able to perform planned detours in experimental settings (see Cross et al. 2020 for a review). One such experimental setting (also reviewed in Kabadayi et al. 2018) requires *Portia* to leave an elevated site where a prey lure can be seen, and choose between two paths at a place where the lure can not be seen. Spider subjects were tested only once and had no similar experience before testing. Still, almost all subjects successfully completed this task. Finally, *Portia* appears to maintain a working memory-like representation of the lure in the process of navigation to the correct site, and displays hesitant behaviour if the number of lures is different than what would be expected (Cross and Jackson 2014). Such hesitancy is usually interpreted as expectancy violation in the context of more complex organisms (Winters et al. 2015).

Cross et al. (2020) interpret the reviewed findings that *Portia* can be considered intelligent although there is an absence of evidence that it uses learning to any significant extent. Rather, *Portia* appears to be born with a repertoire of behaviours and cognitive abilities that allows it to flexibly cope with a broad array of hunting

situations and challenges. We concur with this interpretation; in the framework presented here, *Portia* appears to be able to accurately predict complex obstacles (its spider prey) and successfully negotiate those obstacles by means of behavioural recombination. However, to the degree that its nervous system does not support longer term memory outside working memory used for navigation planning, there is also less room for generalization and abstraction as described above. From the perspective of ecomorphic variation this is perhaps not surprising; compared to the birds and apes mentioned above, *Portia*'s ecological context is relatively stable, partly because its roaming range is limited and unlikely to span large ecological gradients. Hence the magnitude of *Portia*'s intelligence may be high compared to other spider species, but it is still apparently limited in generality due to the attenuated support for long term learning and abstraction afforded by its central nervous system. On the one hand it appears that complex obstacles do not induce sufficient evolutionary pressure to yield generalized intelligence in terms of abstraction, at least not to the same degree as do large ecological gradients. On the other hand, complex obstacles may indeed pressure organisms towards developing behavioural flexibility such as being able to recombine and adapt inborn behaviour.

5.4 Prediction and the Measurement of Intelligence

As stated by Mackintosh (2011), intelligence is difficult to measure directly in experiments, since this would require recording behaviour in individual organisms' natural environment over extended periods of time. In practice, experimenters have to make do with measurements that hopefully point at aspects of intelligence, but can be done within the confines of a laboratory setting. Several such measures have been devised, and we will discuss how some of them might relate to predictive abilities and operations in working memory.

Processing speed, in the form of choice reaction time (RT) and inspection time (IT) are aspects of cognitive processing which is strongly associated with the concept of intelligence, but which may not appear to be directly related to predictive abilities. From our point of view, it is true that these psychometric measures do not necessarily affect accuracy of predictions. Processing speed will however, affect how long it takes for an organism to learn the patterns necessary to make predictions, and it will also affect how quickly an ad-hoc model in working memory can be formed. Since processing speed indicates the efficiency and integrity of the cognitive system, it may indirectly affect prediction accuracy by means of the effort involved in forming the prediction in at least two ways. First, if the effort is too great, prediction formation may terminate prematurely, resulting in less accuracy. Second, working memory operations are dependent on inhibitory interneurons (Wang 2010) which carry a high metabolic cost to keep active (Buzsáki et al. 2007); since processing speed indicates metabolic efficiency to some degree, it also indicates whether sufficient energy can be produced to drive such expensive neuronal machinery.

5.5 Aspects Not Presently Covered by Intelligence as Prediction

Predictive ability, as stated above, is an approximation of the capacity for intelligent behaviour. This implies that there are aspects of intelligence that predictive ability, whether context specific or not, does not specifically cover. We will discuss a selection of these now.

An important contributor to being able to act intelligently is that of cognitive control. This is typically related to impulse inhibition and the ability to stay on task (Carr and Dweck 2011). It is also related to inhibiting the default mode network which produces mental activity like day dreaming (Christoff et al. 2016). Since cognitive control employs populations of inhibitory interneurons, employing it is effortful over time. Hence it can affect the process of forming predictions, as well as the degree to which a plan is adhered to, after a prediction has been generated. Cognitive control also involves attentional focus. This is used to inhibit salient distractors (Posner and Snyder 2004), as well as noise from the default mode network (Brewer et al. 2011).

Context specific prediction is typically less effortful than is the formation of predictive models in working memory, since the former is more constrained and resembles recollection more (Kane 2008). This implies that motivation is necessary both for using working memory, as well as for carrying out plans using the motor system. According to Carr and Dweck (2011), motivation for intellectual pursuits can be affected by several factors, including beliefs about intelligence is fixed or malleable, and beliefs about social belonging. These factors indicate that well-being in general is likely involved in mental behaviour in general, and activities related to prediction formation in particular, via the dynamics of motivation (Dean et al. 2016).

Creativity is related to model construction in working memory, and to prediction through the ability to imagine alternative outcomes, and different ways of negotiating obstacles. Creativity is dependent to some extent on the production of noise in the central nervous system, since some noise is necessary for the system to be able to reach all possible states (Krogh and Hertz 1992; Kirpatrick et al. 1983). This noise is some of the reason for the unpredictable associations and jumps that characterize the activity of the default mode network (Benedek and Jauk 2018). However, to be useful for prediction, these noisy creative processes must be controlled. In other words, they must be attenuated and inhibited when not useful. It is possible that abstraction in working memory can be effectively used together with creative processes by constraining the creative domain. For example, to solve a problem may require a particular property of an object like its stiffness. In the classical problem of creating a candle holder given a pack of matches and some tacs (Duncker 1945), the pack of matches must be viewed as something that can support the candle, rather than as a container of matches. This requires focusing attention on the property of stiffness and inhibition of properties like volume. In effect, this is a combination of abstraction and creativity, used to negotiate an obstacle.

An organism may be able to predict the course of events, but to influence the course of those events to its own benefit, it will need appropriate skills. Although we believe there is no real disjunction between the motor- and cognitive systems, we make a distinction here between cognitive and motor skills for emphasis. Examples of cognitive skills include those involved in playing chess, but also metacognitive skills like how to acquire knowledge, and how to think and reason. Motor system skills can be exemplified by knowing how to dance, or doing carpentry. Many, perhaps most, activities integrate both domains. An interesting aspect of skills is that they typically depend on bottom up processes, or experience to learn. That is, skills are procedural knowledge, and requires learning by doing.

5.6 Summary

We have aimed in this paper to argue that intelligence can be approximated by the ability to produce accurate predictions, and that a distinction exists whether predictions are dependent on context or not. The ability to use working memory to inhibit contextual information, we have argued, affords transfer of predictive patterns across contexts. A complement to this, is the ability to suppress impulse behaviour to maintain approach towards goals that may only be achieved over extended time. Furthermore, the ability to perform transformations in working memory affords the construction of ad-hoc models. We have argued that this ability, combined with context dependent predictive abilities like prospection, can approximate general intelligence. To form this argumentation, we have built on the evolutionary and adaptationist argumentation of Godfrey-Smith, as well as the arguments of Hohwy concerning the mind as a prediction generating entity. Our arguments integrate the two, and add refinements in the following way.

Firstly, to Godfrey-Smith's environmental complexity thesis, we add the notion of niche broadness, as well as a change in perspective on the point of cognition and control. We suggest that behaviour, physical as well as mental, come about to maintain and attain homeostatic setpoints. Such setpoints include, along with physiological ones, setpoints for information and novelty. This builds particularly on Wiener's (1948) concept of cybernetics which later matured into control theory.

Secondly, to complement Hohwy's inclusion of precision on predictions, we make explicit the significance of accuracy as a composite of trueness and precision, and the interpretation of intelligence as accurate prediction. The puzzle of what initiates action in Hohwy's framework (Vance 2015) we think can be illuminated in the context of control theory, where the nervous system acts as a distributed controller, motivating actions that keep the body in homeostasis.

Thirdly, we add to both theories the notion that general intelligence is dependent on structures that can abstract away context, and an attentional system that can selectively focus on context independent patterns. This allows transfer of predictions across contexts, but at the cost of lower precision.

6 Concluding Remarks

We have proposed a synthesis of Godfrey-Smith's environmental complexity theory and Hohwy's predictive mind theory to form the beginnings of a theory of intelligence as accurate prediction. We have also argued that the flexibility associated with general intelligence can be understood as the ability to abstract away context specific detail and transfer predictions to novel situations and environments.

There is more work to be done in considering how predictive abilities are affected by dynamical processes, in particular how they are affected by perceptions of well being, and stress.

Furthermore, our theory needs to take into account how predictions can be made on the basis of ad-hoc models formed in working memory. This in particular has the potential to elucidate the cognitive flexibility associated with general intelligence.

Our arguments for how intelligence can be approximated and understood does not immediately place it in an overarching framework of cognitive functioning. We believe that cybernetics, or control theory is the most promising candidate for providing such a framework. Understanding intelligence as the predictive abilities of a ‘planning controller,’ which can maintain and attain set points by sequential action over long periods of time, is one avenue of research in this direction.

Code Availability Not applicable.

Funding Open access funding provided by Lund University.

Data Availability Not applicable.

Declarations

Conflicts of Interest/Competing Interests No conflicts of interest or competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Asby, W.R. 1961. *An introduction to cybernetics*. Chapman & Hall Ltd..
- Atit, K., T.F. Shipley, and B. Tikoff. 2013. Twisting space: Are rigid and non-rigid mental transformations separate spatial skills? *Cognitive Processing* 14 (2): 163–173.
- Atlan, H., and I. Cohen. 1998. Immune information, self-organization, and meaning. *International Immunology* 10 (6): 711–717.
- Audet, J.N., S. Ducatez, and L. Lefebvre. 2016. The town bird and the country bird: Problem solving and immunocompetence vary with urbanization. *Behavioral Ecology* 27 (2): 637–644.
- Baluška, F., and S. Mancuso. 2018. Plant cognition and behavior: From environmental awareness to synaptic circuits navigating root apices. In *Memory and learning in plants* (pp. 51–77), ed. F. Baluška, M. Gagliano, and G. Witzany. Springer International Publishing.
- Barrett, H.C. 2015. *The shape of thought: How mental adaptations evolve*. Oxford: Oxford University Press.
- Bastos, A.M., R. Loonis, S. Kornblith, M. Lundqvist, and E.K. Miller. 2018. Laminar recordings in frontal cortex suggest distinct layers for maintenance and control of working memory. *Proceedings of the National Academy of Sciences* 115 (5): 1117–1122.
- Behrens, T.E.J., T.H. Muller, J.C.R. Whittington, S. Mark, A.B. Baram, K.L. Stachenfeld, and Z. Kurth-nelson. 2018. What is a cognitive map?: Organizing knowledge for flexible behavior. *Neuron* 100 (2): 490–509.
- Benedek, M., and E. Jauk. 2018. Spontaneous and controlled processes. In *The Oxford handbook of spontaneous thought: Mind-wandering, creativity, and dreaming*, ed. K.C.R. Fox and K. Christoff, 285–298. Oxford: Oxford University Press.
- Birmie-Gauvin, K., K.S. Peiman, A.J. Gallagher, R. De Bruijn, and S.J. Cooke. 2016. Sublethal consequences of urban life for wild vertebrates. *Environmental Reviews* 24 (4): 416–425.

- Birrell, J.M., and V.J. Brown. 2000. Medial frontal cortex mediates perceptual attentional set shifting in the rat. *Journal of Neuroscience* 20 (11): 4320–4324.
- Bobrowicz, K. (2019). *Memory for problem solving: Comparative studies in attention, working and long-term memory* (Vol. 174). Doctoral dissertation, Lund University.
- Bobrowicz, K., O'Hara, M., Carminito, C., Auersperg, A. M. I., Osvath, M. (2019). Goffin's cockatoos (*Cacatua goffiniana*) can use conflicting past experiences to solve a novel problem. In Bobrowicz, K. (2019). *Memory for problem solving: Comparative studies in attention, working and long-term memory* (Vol. 174). Doctoral dissertation, Lund University.
- Bobrowicz, K., M. Johansson, and M. Osvath. 2020. Great apes selectively retrieve relevant memories to guide action. *Scientific Reports* 10 (1): 1–13.
- Brass, M., J. Derrfuss, and D.Y. von Cramon. 2005. The inhibition of imitative and overlearned responses: A functional double dissociation. *Neuropsychologia* 43 (1): 89–98.
- Brewer, J.A., P.D. Worhunsky, J.R. Gray, Y.Y. Tang, J. Weber, and H. Kober. 2011. Meditation experience is associated with differences in default mode network activity and connectivity. *Proceedings of the National Academy of Sciences* 108 (50): 20254–20259.
- Bruineberg, J., and E. Rietveld. 2014. Self-organization, free energy minimization, and optimal grip on a field of affordances. *Frontiers in Human Neuroscience* 8: 599.
- Bruineberg, J., E. Rietveld, T. Parr, L. van Maanen, and K.J. Friston. 2018. Free-energy minimization in joint agent-environment systems: A niche construction perspective. *Journal of Theoretical Biology* 455: 161–178.
- Brüne, M., and U. Brüne-Cohrs. 2006. Theory of mind-evolution, ontogeny, brain mechanisms and psychopathology. *Neuroscience & Biobehavioral Reviews* 30 (4): 437–455.
- Buzsáki, G., K. Kaila, and M. Raichle. 2007. Inhibition and brain work. *Neuron* 56 (5): 771–783.
- Carr, P.B., and C.S. Dweck. 2011. Intelligence and motivation. In *The Cambridge handbook of intelligence*, ed. R.J. Sternberg and S.B. Kaufman, 748–770. Cambridge: Cambridge University Press.
- Castegnetti, G., A. Tzovara, S. Khemka, F. Melinščak, G.R. Barnes, R.J. Dolan, and D.R. Bach. 2020. Representation of probabilistic outcomes during risky decision-making. *Nature Communications* 11: 2419.
- Cellucci, C. (2017). *Rethinking knowledge: The heuristic view* (European studies in philosophy of science, vol. 4). Springer International Publishing.
- Chanes, L., and L.F. Barrett. 2016. Redefining the role of limbic areas in cortical processing. *Trends in Cognitive Sciences* 20 (2): 96–106.
- Christoff, K., and K. Keramiatian. 2007. Abstraction of mental representations: Theoretical considerations and neuroscientific evidence. In *Neuroscience of rule-guided behavior*, ed. S.A. Bunge and J.D. Wallis, vol. 6, 107–126. Oxford: Oxford University Press.
- Christoff, K., Z.C. Irving, K.C. Fox, R.N. Spreng, and J.R. Andrews-Hanna. 2016. Mind-wandering as spontaneous thought: A dynamic framework. *Nature Reviews Neuroscience* 17 (11): 718–731.
- Cohen, G.B., R. Ren, and D. Baltimore. 1995. Modular binding domains in signal transduction proteins. *Cell* 80 (2): 237–248.
- Corbetta, M., and G.L. Shulman. 2002. Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience* 3 (3): 201–215.
- Craik, K. 1943. *The nature of explanation*. Cambridge: Cambridge University Press.
- Cross, F.R., and R.R. Jackson. 2014. Specialised use of working memory by *Portia africana*, a spider-eating salticid. *Animal Cognition* 17: 435–444.
- Cross, F.R., G.E. Carvell, R.R. Jackson, and R.C. Grace. 2020. Arthropod intelligence? The case for *Portia*. *Frontiers in Psychology* 11 (October).
- De Bona, S., J.K. Valkonen, A. López-Sepulcre, and J. Mappes. 2015. Predator mimicry, not conspicuousness, explains the efficacy of butterfly eyespots. *Proceedings of the Royal Society B: Biological Sciences* 282 (1806): 20150202.
- Dean, E., F. Schilbach, and H. Schofield. 2016. Poverty and cognitive function. In *The economics of poverty traps*, ed. C.B. Barrett, M.R. Carter, and J.-P. Chavas, 57–119. Chicago: The University of Chicago Press.
- Dias, R., T.W. Robbins, and A.C. Roberts. 1996. Primate analogue of the Wisconsin card sorting test: Effects of excitotoxic lesions of the prefrontal cortex in the marmoset. *Behavioral Neuroscience* 110 (5): 872–886.
- Dias, R., T.W. Robbins, and A.C. Roberts. 1997. Dissociable forms of inhibitory control within prefrontal cortex with an analog of the Wisconsin card Sort test: Restriction to novel situations and independence from “on-line” processing. *Journal of Neuroscience* 17 (23): 9285–9297.
- Dretske, F. 1981. *Knowledge and the flow of information*. Cambridge: MIT Press.
- Duncker, K. 1945. On problem-solving. *Psychological Monographs*, 58(5), whole no. 270.

- Falke, J.J., R.B. Bass, S.L. Butler, S.A. Chervitz, and M.A. Danielson. 1997. The two-component signaling pathway of bacterial chemotaxis: A molecular view of signal transduction by receptors, kinases, and adaptation enzymes. *Annual Review of Cell and Developmental Biology* 13 (1): 457–512.
- Figner, B., D. Knoch, E.J. Johnson, A.R. Krosch, S.H. Lisanby, E. Fehr, and E.U. Weber. 2010. Lateral prefrontal cortex and self-control in intertemporal choice. *Nature Neuroscience* 13 (5): 538–539.
- Friston, K. 2009. The free-energy principle: A rough guide to the brain? *Trends in Cognitive Sciences* 13 (7): 293–301.
- Friston, K., F. Rigoli, D. Ognibene, C. Mathys, T. Fitzgerald, and G. Pezzulo. 2015. Active inference and epistemic value. *Cognitive Neuroscience* 6 (4): 187–214.
- Gazda, S.K., R.C. Connor, R.K. Edgar, and F. Cox. 2005. A division of labour with role specialization in group-hunting bottlenose dolphins (*Tursiops truncatus*) off cedar key, Florida. *Proceedings of the Royal Society B: Biological Sciences* 272 (1559): 135–140.
- Geary, D.C. 2005. *The origin of mind*. Washington, DC: American Psychological Association.
- Geary, D.C. 2009. The evolution of general fluid intelligence. In *Foundations in evolutionary cognitive neuroscience*, ed. S.M. Platek and T.K. Shackelford, 22–56. Cambridge: Cambridge University Press.
- Gilbert, D.T., and T.D. Wilson. 2007. Prospection: Experiencing the future. *Science* 317 (5843): 1351–1354.
- Godfrey-Smith, P. 1996. *Complexity and the function of mind in nature*. Cambridge: Cambridge University Press.
- Godfrey-Smith, P. 2002. Environmental complexity and the evolution of cognition. In *The evolution of intelligence*, ed. R.J. Sternberg and J.C. Kaufman, 233–249. New York: Psychology Press.
- Godfrey-Smith, P. 2017a. Complexity revisited. *Biology and Philosophy* 32 (3): 467–479.
- Godfrey-Smith, P. 2017b. The evolution of consciousness in phylogenetic context. In *Routledge handbook of philosophy of animal minds*, ed. K. Andrews and J. Beck, 216–226. New York: Routledge.
- Haidle, M.N. 2010. Working-memory capacity and the evolution of modern cognitive potential: Implications from animal and early human tool use. *Current Anthropology* 51 (S1): S149–S166.
- Hanley, M.E., B.B. Lamont, M.M. Fairbanks, and C.M. Rafferty. 2007. Plant structural traits and their role in anti-herbivore defence. *Perspectives in Plant Ecology, Evolution and Systematics* 8 (4): 157–178.
- Hebb, D.O. 1949. *The organization of behaviour*. New York: Wiley.
- Hohwy, J. 2013. *The predictive mind*. Oxford: Oxford University Press.
- Holding, M.L., J.E. Biardi, and H.L. Gibbs. 2016. Coevolution of venom function and venom resistance in a rattlesnake predator and its squirrel prey. *Proceedings of the Royal Society B: Biological Sciences* 283 (1829): 20152841.
- Kabadayi, C., and M. Osvath. 2017. Ravens parallel great apes in flexible planning for tool-use and bartering. *Science* 357 (6347): 202–204.
- Kabadayi, C., L.A. Taylor, A.M. von Bayern, and M. Osvath. 2016. Ravens, new Caledonian crows and jackdaws parallel great apes in motor self-regulation despite smaller brains. *Royal Society Open Science* 3 (4): 160104.
- Kabadayi, C., K. Bobrowicz, and M. Osvath. 2018. The detour paradigm in animal cognition. *Animal Cognition* 21 (1): 21–35.
- Kane, J.E. 2008. Tethered by tense: Retrospection is more constrained than prospection. In *Doctoral dissertation*. Colorado at Boulder: University of.
- Kanoski, S.E., and H.J. Grill. 2017. Hippocampus contributions to food intake control: Mnemonic, neuroanatomical, and endocrine mechanisms. *Biological Psychiatry* 81 (9): 748–756.
- Kirkpatrick, S., C. D. Gelatt, and M. P. Vecchi. 1983. Optimization by simulated annealing. *Science* 220 (4598): 671–680. <https://doi.org/10.1126/science.220.4598.671>.
- Klein, C., and A.B. Barron. 2016. Insects have the capacity for subjective experience. *Animal Sentience* 1 (9/1): 1–53.
- Klein, S.B., L. Cosmides, J. Tooby, and S. Chance. 2002. Decisions and the evolution of memory: Multiple systems, multiple functions. *Psychological Review* 109 (2): 306–329.
- Krogh, A., and J.A. Hertz. 1992. Generalization in a linear perceptron in the presence of noise. *Journal of Physics A: Mathematics and General* 25 (5): 1135–1147.
- Lara, A.H., S.W. Kennerley, and J.D. Wallis. 2009. Encoding of gustatory working memory by orbitofrontal neurons. *Journal of Neuroscience* 29 (3): 765–774.
- Legg, S., and M. Hutter. 2007a. A collection of definitions of intelligence. In *Frontiers in artificial intelligence and applications: Advances in artificial general intelligence: Concepts, architectures and algorithms*, ed. B. Goertzel and P. Wang, vol. 157, 17–24. Amsterdam: IOS Press.
- Legg, S., and M. Hutter. 2007b. Universal intelligence: A definition of machine intelligence. *Minds and Machines* 17 (4): 391–444.

- Lissek, S., B. Diekamp, and O. Güntürkün. 2002. Impaired learning of a color reversal task after NMDA receptor blockade in the pigeon (*Columbia livia*) associative forebrain (Neostriatum Caudolaterale). *Behavioral Neuroscience* 116 (4): 523–529.
- Lowry, H., A. Lill, and B.B. Wong. 2013. Behavioural responses of wildlife to urban environments. *Biological Reviews* 88 (3): 537–549.
- Luniak, M. 2004. Synurbization: Adaptation of animal wildlife to urban development. In *Proceedings of the 4th international symposium on urban wildlife conservation, may 1–5 1999*, ed. W.W. Shaw, L.K. Harris, and L. VanDruff, 50–55. Tucson, Arizona: University of Arizona.
- Mackintosh, N.J. 2011. History of theories and measurement of intelligence. In *The Cambridge handbook of intelligence*, ed. R.J. Sternberg and S.B. Kaufman, 3–19. Cambridge: Cambridge University Press.
- Manns, J.R., and H. Eichenbaum. 2006. Evolution of declarative memory. *Hippocampus* 16 (9): 795–808.
- Mason, K.A., J.B. Losos, and S.R. Singer. 2017. *Biology*. New York: McGraw-Hill.
- McMullin, E. 2002. The origins of the field concept in physics. *Physics in Perspective* 4 (1): 13–39.
- Menditto, A., M. Patriarca, and B. Magnusson. 2007. Understanding the meaning of accuracy, trueness and precision. *Accreditation and Quality Assurance* 12 (1): 45–47.
- Mink, J.W., R.J. Blumenshine, and D.B. Adams. 1981. Ratio of central nervous system to body metabolism in vertebrates: Its constancy and functional basis. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* 241 (3): R203–R212.
- Molles, M.C., and A.A. Sher. 2018. *Ecology: Concepts and applications*. Boston: McGraw-Hill Education.
- Munakata, Y., S.A. Herd, C.H. Chatham, B.E. Depue, M.T. Banich, and R.C. O'Reilly. 2011. A unified framework for inhibitory control. *Trends in Cognitive Sciences* 15 (10): 453–459.
- Olivers, C.N., F. Meijer, and J. Theeuwes. 2006. Feature-based memory-driven attentional capture: Visual working memory content affects visual attention. *Journal of Experimental Psychology: Human Perception and Performance* 32 (5): 1243–1265.
- Orban, G.A., and F. Caruana. 2014. The neural basis of human tool use. *Frontiers in Psychology* 5: 310.
- Osvath, M., and H. Osvath. 2008. Chimpanzee (*Pan troglodytes*) and orangutan (*Pongo abelii*) forethought: Self-control and pre-experience in the face of future tool use. *Animal Cognition* 11 (4): 661–674.
- Pezzulo, G., E. Carboni, F. Rigoli, L. Pio-Lopez, and K. Friston. 2016. Active inference, epistemic value, and vicarious trial and error. *Learning & Memory* 23 (7): 322–338.
- Plotkin, H.C. 1993. *Darwin machines and the nature of knowledge*. Cambridge, MA: Harvard University Press.
- Posner, M.I., and C.R.R. Snyder. 2004. Attention and cognitive control. In *Key readings in cognition: Cognitive psychology*, ed. D.A. Balota and E.J. Marsh, 205–223. New York: Psychology Press.
- Potts, C.A., S. Pastel, and D.A. Rosenbaum. 2018. How are cognitive and physical difficulty compared? *Attention, Perception, & Psychophysics* 80 (2): 500–511.
- Powell, R.P., and A.W. Still. 1979. Behaviorism and the psychology of language: An historical reassessment. *Behaviorism* 7 (1): 71–89.
- Redish, A.D. 2016. Vicarious trial and error. *Nature Reviews Neuroscience* 17 (3): 147–159.
- Reeve, J., and W. Lee. 2012. Neuroscience and human motivation. In *The Oxford handbook of human motivation*, ed. R.M. Ryan, 365–380. Oxford: Oxford University Press.
- Resnick, I., and T.F. Shipley. 2013. Breaking new ground in the mind: An initial study of mental brittle transformation and mental rigid rotation in science experts. *Cognitive Processing* 14 (2): 143–152.
- Richman, D.B., and R.R. Jackson. 1992. A review of the ethology of jumping spiders (Araneae, Salticidae). *Bulletin of the British Arachnological Society* 9 (2): 33–37.
- Schacter, D., D. Addis, and R. Buckner. 2007. Remembering the past to imagine the future: The prospective brain. *Nature Reviews Neuroscience* 8: 657–661.
- Seed, A., and R. Byrne. 2010. Animal tool-use. *Current Biology* 20 (23): R1032–R1039.
- Shannon, C.E. 1948. A mathematical theory of communication. *The Bell System Technical Journal* 27 (3): 379–423.
- Sleep, N.H., and D.K. Bird. 2007. Niches of the pre-photosynthetic biosphere and geologic preservation of Earth's earliest ecology. *Geobiology* 5 (2): 101–117.
- Smith, D.M., and S.J. Mizumori. 2006. Hippocampal place cells, context, and episodic memory. *Hippocampus* 16 (9): 716–729.
- Sohn, Y.H., K. Wiltz, and M. Hallett. 2002. Effect of volitional inhibition on cortical inhibitory mechanisms. *Journal of Neurophysiology* 88 (1): 333–338.
- Spearman, C. 1904. "General intelligence," objectively determined and measured. *The American Journal of Psychology* 15 (2): 201–292.
- Spearman, C. 1928. The abilities of man. *Science* 68 (1750): 38.

- Špinka, M., and F. Wemelsfelder. 2011. Environmental challenge and animal agency. In *Animal Welfare*, ed. M.C. Appleby, J.A. Mench, L.A.S. Olsson, and B.O. Hughes, 27–43. Cambridge, MA: Cabi.
- Stander, P.E. 1992. Cooperative hunting in lions: The role of the individual. *Behavioral Ecology and Sociobiology* 29 (6): 445–454.
- Stephens, D.W., and J.R. Krebs. 1986. *Foraging theory*. 1st ed. Monographs in Behavior and Ecology: Princeton University Press.
- Stepp, N., and M.T. Turvey. 2010. On strong anticipation. *Cognitive Systems Research* 11 (2): 148–164.
- Strauß, A., M. Wöstmann, and J. Obleser. 2014. Cortical alpha oscillations as a tool for auditory selective inhibition. *Frontiers in Human Neuroscience* 8: 350.
- van Horik, J.O., C.E. Beardsworth, P.R. Laker, E.J.G. Langley, M.A. Whiteside, and J.R. Madden. 2019. Unpredictable environments enhance inhibitory control in pheasants. *Animal Cognition* 22: 1105–1114.
- Vance, J. 2015. Review of *The predictive mind*. *Notre Dame Philosophical Reviews*.
- Wang, X.J. 2010. Neurophysiological and computational principles of cortical rhythms in cognition. *Physiological Reviews* 90 (3): 1195–1268.
- Wiener, N. 1948. *Cybernetics*. John Wiley and Sons.
- Winters, S., C. Dubuc, and J.P. Higham. 2015. Perspectives: The looking time experimental paradigm in studies of animal visual perception and cognition. *Ethology* 121: 625–640.
- Yousefian, F., H. Saberi, and T. Baniroostam. 2016. A conceptual model for ontology based intelligence. *International Journal of Computer Science and Network* 5 (1): 23–30.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.