

The Evolution of Decoupled Representation

Lachlan Douglas Walmsley, BA

Thesis submitted in partial fulfilment of the requirement for the award of the degree

Master of Research

Department of Philosophy, Macquarie University

Originally submitted 10 October 2014

Submitted with revisions

Table of Contents

Synopsis.....	v
Certification of Original Content	vii
Acknowledgements	ix
Introduction.....	1
Foundations.....	4
1.1 The Environmental Complexity Thesis	4
1.1.1 Defining Complexity	5
1.1.2 Illustrating the ECT	7
1.1.3 Summary	8
1.2 The Hostility Thesis	8
1.2.1 The Problems of Hostility.....	9
1.2.2 Detection Systems and Robust Tracking Systems	10
1.2.3 Intentional Systems	15
1.2.4 Summary	18
1.3 Taking the Theses Together.....	18
1.3.1 The Relations Between the Two Accounts	19
1.3.2 Externalism and the ECT	20
1.3.3 Summary	22
1.4 Conclusion	23
Defense	24
2.1 Christensen’s Interpretation of Sterelny.....	25
2.1.1 Christensen on Decoupled Representation	25
2.1.2 Christensen on Sterelny’s Commitments.....	27
2.1.3 Summary	28

Contents

2.2 The Early Evolution of Decoupled Representation	29
2.2.1 Decoupled Representation and Early Nervous Systems	29
2.2.2 Analysis of Christensen's Evidence	31
2.2.3 Summary	33
2.3 Decoupled Representation and Inference	34
2.3.1 Rat Inference	35
2.3.2 Overcoming the Methodological Problem?	37
2.3.3 Summary	39
2.4 The Limitations of Detection	39
2.4.1 The Behaviour of Green Algae	40
2.4.2 Response to Akins and Pollon	41
2.4.3 Summary	42
2.5 Conclusion	43
Revision.....	44
3.1 Coupled Representations	45
3.1.1 The Explanation Challenge and Representation.....	46
3.1.2 Cross-channel Interaction	47
3.1.3 Coupled Representations	48
3.1.4 Summary	49
3.2 Solving the Inference Problem*	49
3.2.1 Coupled S-Representation.....	50
3.2.2 Decoupled S-representation.....	51
3.2.3 S-representation and Response Breadth	53
3.2.4 Summary	56
3.3 On-going Coordination	56
3.3.1 Solving the Tracking Problem	57

The Evolution of Decoupled Representation

3.3.2 On-going Coordination in Hostile Environments.....	59
3.3.3 Summary	61
3.4 Continuity	61
3.4.1 From Robust Tracking to Decoupled Representation.....	62
3.4.2 Precision and Accuracy	66
3.4.3 Summary	68
3.5 Conclusion	69
Conclusion	70
References	72
Appendix: List of Problems	76

Synopsis

Sterelny (2003) has developed an influential account of the origin of human-like beliefs, which complements Godfrey-Smith's (1996, 2002a) account of the evolution of cognition. In this thesis, I focus on the first part of Sterelny's account, which addresses the evolution of cognition in non-human animals.

In this thesis, I defend Sterelny's account against two criticisms. Christensen (2010) argues that Sterelny's most complex behavioural control system evolves as early as the nervous system. Akins and Pollon (in preparation) argue that Sterelny's simplest control system cannot explain a great deal of the most basic behaviour. According to both these arguments, the findings of empirical research demonstrate that Sterelny's account requires major revision or should simply be rejected. I counter these arguments, and show that Sterelny's account does not require major revision.

Although Sterelny's account does not require major revision, Christensen's and Akins and Pollon's criticisms show that there are few details of the continuity between simple and complex behavioural control in Sterelny's account. Akins and Pollon's criticism also shows that behavioural control often involves on-going coordination and that the boundaries between input and output are sometimes more indistinct than Sterelny's account indicates. In this thesis, I incorporate on-going coordination into Sterelny's account of behavioural control, and expand upon Sterelny's account of continuity.

Certification of Original Content

I certify that the content of this work (of 19,787 words) is original, except where indicated, and has not been submitted as part of any other degree. Amendments were made to this work based on the reports of three examiners.

- Lachlan Douglas Walmsley

Acknowledgements

I would like to thank my supervisor, Richard Menary. Without his advice and guidance I would not have been able to write this thesis (not the good bits, at least).

I must thank my partner, Caitlin, for her love and support throughout a difficult year.

Thanks also go to Rachael Brown, who at times was like a second supervisor, and to Robert Keys, Alex Gillett, and my sister, Donherra. All their comments were helpful, and their efforts were greatly appreciated.

I thank my mother, who always told me that what I was writing was interesting. I thank Richard, again, for telling me otherwise.

Introduction

According to folk psychology, human behaviour can be explained in terms of beliefs and desires: Brutus desired to kill Caesar, and he believed that he could kill Caesar with a knife, so Brutus stabbed Caesar. Humans also commonly use the terminology of folk psychology when explaining a range of non-human action (see Dennett 1989). For example, subjects in Heider and Simmel's (1944) experiments explain the movements of two-dimensional shapes in terms of the shapes' beliefs and desires: the big triangle 'is shut up in the house and *tries* to get out' (p. 247, my emphasis). According to Sterelny (2003) and Godfrey-Smith (2002b), these are two distinct phenomena. The first is that humans have beliefs and desires, and the second is that humans habitually interpret the behaviour of others in terms of beliefs and desires.

Sterelny develops an influential account of the origins of **cognition**, which is understood here as a system of behavioural control. In this account he aims to explain both the gradual evolution of human-like beliefs and desires, and the human habit of understanding behaviour in terms of beliefs and desires. Here, *I focus on Sterelny's sketch of the evolution of belief-like states in non-human organisms*, which has recently been criticised by Christensen (2010) and Akins and Pollon (A&P) (in preparation) on the grounds that it does not explain the findings of empirical research. Christensen argues that Sterelny's most complex control system evolves much earlier than Sterelny claims. A&P argue that Sterelny's most simple control system cannot explain a great deal of basic behaviour. Together, these criticisms indicate that Sterelny's account of the continuity between simple and complex systems of behavioural control requires major revision.

My aim is to defend Sterelny's account against these criticisms, and to show that his account does not require major revision. I counter Christensen by arguing that he misinterprets Sterelny's claims. I counter A&P by arguing that only Sterelny's simplest control system requires revision, not his entire account.

Introduction

Although I weaken Christensen's and A&P's arguments, the evidence they present cannot currently be explained by Sterelny's account. *My aim is to revise Sterelny's account to explain this evidence.* I introduce another kind of representation to explain Christensen's evidence, and a new kind of behavioural control to explain A&P's evidence. A&P's evidence indicates that on-going coordination with the environment plays an important role in basic behavioural control. I argue that it continues to play this role in more complex forms of behavioural control to ensure that representations remain accurate. Furthermore, because most responses are temporally extended, they involve simultaneous input and output. Therefore, the boundaries between input and output should be understood as indistinct in many cases.

Sterelny provides little explanation of the continuity between his behavioural control systems. *My aim is to expand upon the details of the continuity between basic and complex behavioural control.* I argue that representations must be precise in order to be used flexibly, so the evolution of complex behavioural control can be understood as the evolution of the capacity to represent the environment with precision. I also argue that selection for this capacity is limited to environments where novel problems are common and action is expensive.

Thesis Outline

In Chapter 1, I present Sterelny's account. I also present Godfrey-Smith's account of the evolution of cognition, which Sterelny uses as a framework. I present Christensen's argument that Sterelny's account is undermined by its relation to Godfrey-Smith's, and argue that Sterelny and Godfrey-Smith provide complementary accounts of the evolution of cognition.

In Chapter 2, I present Christensen's and A&P's arguments that empirical evidence indicates that Sterelny's account is either extremely limited or incorrect. I weaken these arguments, but concede that Sterelny's account requires revision to explain the evidence they present.

The Evolution of Decoupled Representation

In Chapter 3, I revise Sterelny's account to accommodate the evidence presented by Christensen and A&P. I introduce a new kind of representation to explain Christensen's evidence, and a new kind of behavioural control to explain A&P's evidence. I also show that on-going coordination plays an important role in complex forms of behavioural control. Finally, I expand upon Sterelny's account of the continuity between simple and complex systems of behavioural control.

Foundations

This chapter has four aims. *The first aim is to outline Sterelny's account of cognitive evolution* (2003, also 1995, 1999, 2001, 2006). I outline this account in section 1.2, after introducing Godfrey-Smith's account (1996, 2002a) in section 1.1, which Sterelny uses as a framework. *The second aim is to indicate that Sterelny provides little explanation of continuity between his control systems.* I indicate this in section 1.2, but do not expand upon the explanation until Chapter 3.

The third aim is to show that Sterelny and Godfrey-Smith provide complementary accounts of the evolution of cognition. Christensen claims that Sterelny's account is undermined by its relation to Godfrey-Smith's account because each provides a different reason as to why cognition evolves. In section 1.3, I argue that the relations between the two accounts are not problematic: Godfrey-Smith explains *why* cognition evolves, and Sterelny explains *how* cognition performs its function in different environments. I also show that Godfrey-Smith's commitments are no longer incompatible with Sterelny's, and argue that these new commitments are not problematic.

The fourth aim is to show that Godfrey-Smith and Sterelny are both committed to the view that organisms and environments are causally coupled and fluid. I briefly show this in section 1.3, and revisit it in Chapter 3, where I argue that these commitments should be extended from evolution and biological entities to processes of behavioural control.

1.1 The Environmental Complexity Thesis

In this section I present Godfrey-Smith's account of cognitive evolution. I present this account first because it serves as the foundation for Sterelny's account. According to Godfrey-Smith's

account (1996, 2002a), called the **Environmental Complexity Thesis** (ECT), cognition¹ deals with the problems created by environmental complexity (1996 p. 3). Godfrey-Smith claims that both variability and stability are properties of environments that select for cognition (1996, p. 216). Variability reduces the pay offs of inflexible responses but stability enables organisms to track this variation and coordinate their responses with it. The function² of cognition is to coordinate these responses. In section 1.1.1, I define Godfrey-Smith's notion of complexity. In section 1.1.2, I illustrate the ECT with an example.

1.1.1 Defining Complexity

In the context of the ECT, complexity implies neither superiority nor sophistication (Godfrey-Smith 1996, p. 24). As I explain below, it can be measured in multiple ways in the context of both environments and organisms.

Environmental Complexity

Environments can be complex in two ways: different states can obtain over space or over time within a given environment (pp. 24-25). Environments are typically a mixture of both simplicity and complexity. An organism's properties—such as its lifespan or size—determine what environmental variation is relevant (p. 153). For example, a light primate is able to move swiftly through the treetops with little consideration for the size of the branches forming its path. Any variation in branch size is irrelevant to this primate because all the branches support its weight. A heavy primate, however, must track branch size to ensure that only weight-supporting branches form its path (Sterelny 1997).

¹ Godfrey-Smith takes a broad view of cognition; the details are discussed in section 1.1.1.

² Here, function is understood in the sense of Millikan's proper functions (Millikan 1984). According to Millikan, a biological entity's function is determined by the operation it performs that has led to its reproduction. For example, the operation of pumping blood, rather than producing a beating sound, has led to the reproduction of the heart, which makes pumping blood the heart's function.

Biological Complexity

An organism's complexity can be measured according to its structure or function (Godfrey-Smith 1996, p. 26). Structural complexity can be measured in multiple ways, such as by the number of parts that constitute an organism (e.g. cells or anatomical features) or by the number of different types of parts that constitute an organism (e.g. different types of cells or different types of anatomical features). Functional complexity is measured according to the number of functional states an organism can occupy. For example, Planarians have impressive regenerative capacities, making them functionally complex in this respect (Reddien & Alvarado 2004).

Behavioural Complexity

Behaviour is a type of function (Godfrey-Smith 1996, p. 25). According to Dretske (1988), to behave is to *do* something. According to Millikan this *doing* 'is an external change or activity exhibited by an organism or external part of an organism'³ and it has a function that would normally be achieved through interacting with the environment or as a result of these interactions (1993, p. 137). Behaviours are distinct from physiological and developmental processes; behaviours are 'not most simply and usefully explained by cellular-, tissue- or organ level processes alone'; and are not 'more easily understood as developmental changes'⁴ (Levitis et al. 2009, p. 108).

³ Millikan attributes this condition to behaviours as 'a rough way to distinguish behaviours from psychological processes' (1993, p. 137).

⁴ According to Levitis et al., developmental process 'are generally much slower than the phenomena considered as behaviour, and are primarily based on ontogenetic programmes by the individual's genetic makeup' (2009, p. 108). I am unsure whether Sterelny or Godfrey-Smith would agree with this definition, or one that places less emphasis on the role of the genotype (see Stotz 2010). However, it is beyond the scope of this thesis to explore this, so a definition of developmental processes has been omitted.

Cognition

In the context of the ECT, cognition is ‘viewed as a *means to the production of behavioural complexity*’ (Godfrey-Smith 1996, p. 26, author’s emphasis), but how exactly cognition coordinates behaviour outside of the scope of the ECT (p. 127). Godfrey-Smith states that cognition ‘should be thought of as a diverse tool kit of capacities for behavioural control, including capacities for perception, internal representation of the world, memory, learning, and decision-making’ (2002a, p. 247), but he takes a broad view of cognition, so cognitive processes are not limited to, or constituted by, all these capacities. This is because different organisms encounter different problems in their environments, which are dealt with differently. One organism may not require capacities that are essential for another’s survival. For example, primates may have cognitive capacities for navigating their social environments that would be of little use to cockroaches.

1.1.2 Illustrating the ECT

The response of the planigale (*Planigale maculata*), a small frog-eating Australian marsupial, to the introduction of the cane toad (*Rhinella marina*) illustrates the main point of the ECT – that dealing with environmental variation requires behaviour flexibility enabled by cognition. Before the introduction of the poisonous cane toad, the environment was invariable with respect to frog-like organisms, which could all be eaten. Now, however, the environment is complex: sometimes frog-like organisms can be eaten and sometimes they cannot be eaten. If the planigales could not learn to respond flexibly to this environmental complexity they could employ one of two strategies, both of which deliver mediocre results. The first strategy is to eat all frog-like organisms, which leaves them vulnerable to poisoning, and the second strategy is to avoid all frog-like organisms, which prevents them from accessing food. The introduction of the cane toad has actually had little impact on the planigale population

because they are conditioned to avoid cane toads after their first unpleasant experience (Webb et al. 2008; Llewelyn et al. 2010).

The story of the planigale shows that behavioural complexity is required to deal with the problems created by environmental complexity. This behavioural complexity is, in turn, enabled by cognition. In this case, this involves the capacity to discriminate between different frog-like organisms and to learn by conditioning.

1.1.3 Summary

Above I have briefly presented the ECT. According to Godfrey-Smith, cognition evolves to coordinate behaviour in response to relevant environmental variation. Sterelny uses the ECT as a framework, so his account is better understood in the context of the ECT. I present Sterelny's account below.

1.2 The Hostility Thesis

In this section I present Sterelny's account of cognitive evolution. According to Sterelny's account (2003, also 1995, 1999, 2001, 2006), which I call the **Hostility Thesis** (HT), hostility⁵ creates environments that are more variable, less stable, and in which action is more costly. Consequently, hostile environments select for cognition that is more complex. In section 1.2.1, I describe the problems created by hostility in greater detail. In sections 1.2.2 and 1.2.3, I describe Sterelny's control systems.

⁵ Hostility in this sense is understood as competition between organisms rather than, for example, a harsh climate.

1.2.1 The Problems of Hostility

Hostility and Variability

Predictable and stereotyped behaviour can be exploited in hostile environments. If prey uses the same strategies to hide or flee, predators will learn how to locate or pursue them. Similarly, if predators use the same strategies to locate and pursue, prey will learn how to hide or flee from them. Hostility increases variability due to action being ‘agent-sensitive’ in this way (Sterelny 2006, p. 297). Agent-sensitive action can lead to ‘arms races’, where agents continually try to out-manoeuvre and out-compete each other (Dawkins 1986, ch. 7-8).

Hostility and Stability

Stable signals can be exploited in hostile environments. **Information**, which is understood here in the minimal sense of co-variation, abounds in the environment (Dretske 1981; Shannon 2001). According to the standard account, Smoke is a signal that carries information about the chances of the state Fire obtaining because these states often co-vary (Skyrms 2010). Exploitation of signals reduces the stability between co-varying states. For example, the female *Photuris* firefly preys upon male fireflies of other species by mimicking the mating signals of those species (Sterelny 2003, p. 15). This destabilises the mating signal, so that it co-varies with the states Predator and Mate instead of only Mate.

When signals co-vary with a single relevant environmental state they carry perfect information and are simple, mapping one-to-one onto the environmental state. When signals co-vary with multiple relevant environmental states they carry imperfect information and are complex, mapping one-to-many onto environmental states. Environments that contain perfect information are **transparent** and environments that contain imperfect information are **translucent**, but environments are typically a mixture of both (pp. 20-26). As the firefly example above demonstrates, hostility often increases translucency.

Hostility and the Cost of Action

Responding poorly can have disastrous consequences in hostile environments⁶. When environmental feedback is harmful, organisms cannot use it to coordinate their behaviour. One strategy for dealing with complex problems is to simplify the problem through action, but taking action becomes dangerous when the environment is hostile. Predator inspection, for example, increases the information an organism has regarding a predator but makes the inspector vulnerable. According to Sterelny other costly actions include ‘competitive interactions within a social group’, which often involve extended processes of mutual probing (2006, p. 298).

1.2.2 Detection Systems and Robust Tracking Systems

Sterelny describes three systems of behavioural control. Sterelny’s first two systems are distinguished according to the number of channels through which they track environmental variation. The notion of channels has a specific definition in information theory (see Dretske 1981), but here it is enough to understand a channel as being able to transmit a single cue (Sterelny 2001, p. 263). Sterelny invites us to take this reading: ‘When no one cue is sufficiently reliable, selection can favour the evolution of the capacity to make use of multiple channels’ (2003, p. 27). Cues also admit of several dimensions. For example, bees (*Apis* spp.) use the sun as a cue when navigating, but this cue can be transformed across multiple dimensions:

...we can define the minimum criteria for the sun as understood by bees. To a first approximation, the object should be less than 20 degrees across (up to fifty times the sun’s apparent size); it can have any amount or direction of polarization, its light must be less than 15% ultraviolet. In other words, having navigated by the sun for days, a forager will blithely

⁶ Dennett (1995) similarly claims that organisms must trial actions in a model environment once trialling actions in the environment becomes costly.

accept a 15 degree, 100% polarised, green triangle for the .4 degree, unpolarised, white circle we know as the sun. (Gould & Gould 1988, p. 128; in Sterelny 2003, p. 26)

The defensive behaviour of the hornet *Vespa mandarinia* provides us with another example of a cue. This behaviour is triggered by a ‘multi-component alarm pheromone’ with defensive behaviour, which is released as venom by other hornets, signalling to guard hornets that the nest is under threat (Ono et al. 2003, p. 637). Again, the alarm cue has multiple dimensions, which are its chemical components⁷. This suggests detecting a single cue involves some kind of integration. This will become important in Chapter 2.

Detection Systems

Sterelny’s most simple control system is the **detection system**. A system is a detection system if it: (d1) is directed toward a single behavioural response; and (d2) is sensitive to relevant environmental variation through a single channel (Sterelny 2003, p. 14). A detection agent is an organism whose behaviour is controlled entirely by detection systems. Detection systems can be hardwired or learned through association (Sterelny 2003, p. 14). The trapping response of the Venus flytrap (*Dionaea muscipula*) is an example of a very simple hardwired detection system. The Venus flytrap detects prey through a small number of mechanosensory hairs positioned on the interior sides of the lobes of its trap. Stimulation of these hairs creates a receptor potential, which, after reaching a threshold, creates an action potential that causes the trap to close (Jacobson 1965).

Sterelny, when discussing the power of detection systems, states that some behavioural sequences involve a **detection cascade** (2003, p. 18). This is a process by which one behavioural response positions the organism so that it receives another cue, which

⁷ These are the chemicals 2-pentanol, 3-methyl-1-butanol, and 1-methylbutyl 3-methylbutanoate (Ono et al. 2003, p. 637).

Foundations

triggers another response, and so on, until a complex behaviour is produced. According to Sterelny, the courtship ritual of the great crested grebe (*Podiceps cristatus*) is an example of a detection cascade.

Environmental generalisation selects for **detection systems with redundancy strategies** (pp. 23-24). The escape behaviour of the cockroach (*Periplaneta Americana*) is controlled by a hardwired detection system with redundancy. Puffs of wind, usually created by an approaching or striking toad (*Bufo marinus*), trigger the cockroach's Turn And Run escape behaviour (Camhi et al. 1978). Other predators do not create this wind displacement so must be detected through another cue. In these cases the cockroach's escape behaviour is triggered by mechanoreceptors along its antennae that respond to touch (Comer et al. 1994).

Detection systems work when cues are reliable and stable. The cockroach's predators are unable to mask themselves and so are transparent aspects of the cockroach's environment. When a single cue is unreliable or unstable organisms need to track environmental variation through multiple channels (Sterelny 2003, p. 27).

Robust Tracking

More complex than a detection system, a **robust tracking system**: (rt1) is directed toward a single behavioural response; and (rt2) is sensitive to relevant environmental variation through multiple channels (Sterelny 2003, p. 27-29). The reed warbler (*Acrocephalus scirpaceus*) provides an example of robust tracking (pp. 27-28). Reed warblers are vulnerable to brood parasitism by cuckoos (*Cuculus canorus*), so the cue Egg In The Nest does not always covary with the environmental state Reed Warbler Egg In The Nest. Relying on a single cue could have disastrous consequences if that cue was unstable, so reed warblers perform egg discrimination according to multiple cues, including relative egg size, egg colour, time at which eggs appeared in the nest, and presence of cuckoos in the local area. This example demonstrates how organisms disambiguate an unstable cue by integrating multiple cues.

Pushmi-Pullyu Representation

The detection system and robust tracking system are both orientated toward particular actions:

...robust systems, like detection systems, are behaviour-specific. Their function is to link the registration of a salient feature of the world to an appropriate response. If we think of these registrations in representational terms, as being about the states of the world they track, they are neither reports nor instructions. (Sterelny 2003, p. 29)

In Millikan's language (1989, 1995), these systems involve **pushmi-pullyu representations** (pp-representations). A representation is a pp-representation if it: (ppr1) is used to control a particular behaviour and cannot be used to control a great number of behaviours; and (ppr2) can only be involved in **on-line** processing. According to Wilson, on-line processing is 'immediate' (2002a, p. 625) and can be distinguished from **off-line** processing, which takes place 'in the absence of task relevant input and output' (p. 626). Off-line cognition, for example, involves 'stepping back, observing, assessing, planning, and only then taking action' (p. 628).

Pp-representations 'have both a descriptive and a directive function, yet they are not equivalent to the mere conjunction of a pure descriptive representation and a pure directive one but are more primitive than either' (Millikan 1995, p. 186). Sterelny provides an example of a pp-representation:

Meerkat sentries produce loud, high-pitched calls when they spot eagles, calls which sent meerkats full speed to their burrows (Avital and Jablonka 2000, p. 244). Since there is such a tight linkage between what is detected in the world – the call – and what is done, it is arbitrary to translate its registration as "Eagle Above!" rather than "Run!" or vice versa. (Sterelny 2003, p. 29)

Foundations

As Sterelny states, we cannot distinguish whether the alarm call carries the content Eagle Above (the descriptive) or Run (the directive). This is because ‘the call connects directly with action’ (Millikan 1995, p. 190). If a registration does not have this kind of tight link to action it cannot be a pp-representation. This is because we can (non-arbitrarily) attribute only descriptive content to it. If a representation does not lead directly to action then it is not a pp-representation, so pp-representations can only be involved in on-line cognition.

Although Sterelny suggests that both detection and robust tracking involve pp-representation, I argue that robust tracking systems must involve a more complex form of representation in Chapter 3.

The Explanation Challenge

Both detection and robust tracking systems function to control a single behaviour, but robust tracking systems function to do this in a way that affords the organism with ‘islands of *resilience*’ (Sterelny 2003, p. 28, author’s emphasis). This resilience is required in translucent environments in which information is imperfect. Comparing actual-sequence explanations with robust-process explanations illustrates the difference between systems with islands of resilience and those without (Sterelny 1995, 1996a).

Actual-sequence explanations state exactly how an event unfolded on a given occasion. For example, the Venus flytrap registers something fly-like through stimulation of the mechanosensory hairs on its trap, which then closes. Robust-process explanations state why an event might have unfolded to the same end if it had not unfolded in the way it did. For example, the Venus flytrap has a system which functions to catch fly-like things through stimulation of the mechanosensory hairs on its trap. However, in this case a robust-process explanation has the same predictive power as an actual-sequence explanation. This is because the Venus flytrap’s trapping response is always triggered by a *particular* proximal cue. On the other hand, when organisms are sensitive to environmental variation through multiple channels, their behaviour can be triggered by *different* cues, so the actual-sequence

explanation may differ from one time to the next. In these cases a robust-process explanation has greater predictive power than the actual-sequence explanation.

If a robust-process explanation has greater predictive power than an actual-sequence explanation, then the system being explained is not a detection system. We can think of this as a test for robust tracking. I call this the **Explanation Challenge**⁸.

1.2.3 Intentional Systems

One of Sterelny's aims is to explain the origins of belief-like states⁹. These are inner states that function only to indicate something about the state of the world, rather than to direct action (2003, p. 30). They can also be used flexibly. According to Sterelny, representations that contain information about the environment, but are not connected to specific behavioural responses, are the origins of these states. These 'minimal beliefs' are called **decoupled representations** (Sterelny 2004, p. 520). A representation is decoupled if it: (dr1) is not used to control particular behaviours and can be used to control a great number of behaviours; and (dr2) can be involved in off-line processing.

Decoupled representations are potentially relevant to many responses. Sterelny takes **response breadth** to be 'the capacity to respond in more than one way to the registration of a focal aspect of the environment' (Sterelny 2003, p. 35 footnote). Vervet monkeys (*Chlorocebus pygerythrus*), for example, may have multiple possible behavioural responses after identifying a predator. They may flee, they may hide, they may undertake further inspection, or they may continue their previous behaviour but with heightened alertness

⁸ Detection systems with redundancy strategies also pass the Explanation Challenge. This is discussed further in Chapter 3

⁹ Another of Sterelny's aims is to explain the origins of desire-like states, but incorporating this part of his account is beyond the scope of this thesis. This omission is not problematic because Christensen criticises decoupled representations and preferences separately, so I only address his criticisms of decoupled representation. Similarly, A&P's criticism is not targeted at preferences.

(Sterelny 2003, pp. 34-35). Sterelny states that decoupled representation ‘is nothing but *very* broad band response’ (p. 34, my emphasis), and that it is ‘a matter of degree’ (*ibid.*). As response breadth increases so does decoupling; however, the details of the continuity between decoupled representations and simpler representational types are lacking in the HT (the Hostility Thesis). As I show in Chapter 2, this lack of detail leaves the HT vulnerable to criticisms like Christensen’s and A&P’s. It is one of the aims of this thesis to provide some of these details, which I do in Chapter 3.

Despite its connection to flexible response, the evolution of decoupled representation is not driven alone by the need for flexible response, but by the need to respond flexibly *given* informational translucency and the cost of action (Sterelny 1997). Decoupled representation is selected for in environments where it is unclear: how information could be used; when information could be used; and how new information could modify old information (Sterelny 2003, p. 81).

The Methodological Problem

According to Sterelny, there is little empirical evidence of decoupled representation in non-humans. This is not because non-humans lack this capacity, but because of the **Methodological Problem**¹⁰. The Methodological Problem arises because: (1) response breadth is typically restricted in most experimental methodologies, making it difficult to test for (dr1); and (2) it is often unclear what information an animal uses to control its behaviour, making it difficult to test for (dr2) (Sterelny 2003, pp. 36-37). Despite the lack of evidence, Sterelny states that ‘many translucent environments select for decoupled representation, for in many such environments information becomes available in a piecemeal fashion and without its immediate significance for action being apparent’ (p. 78). Sterelny argues that one

¹⁰ It is beyond the scope of this thesis to address overcoming the Methodological Problem, but this could be addressed in future work.

The Evolution of Decoupled Representation

environment that probably selects for decoupled representation is the complex social environment of primates.

The Social Intelligence Hypothesis

Sterelny claims that complex social environments like those inhabited by primates are likely to select for decoupled representation, or, at least, for representations that are themselves the *origins* of decoupled representations (2003, p. 53). According to Sterelny, members of primate social groups must track third-party relations. For example, if group member X has been in conflict with group member Y, and group member Z is Y's kin, then X will act aggressively toward Z as well as Y (Sterelny 2003, p. 52; see also Tomasello 2000). Sterelny suggests that social maps of group members' relations are possible origins of decoupled representation because information about group structure is used when interacting with different members of the group, and different behaviours will be produced in response to different members. Furthermore they can be used in off-line processing because information about Y's position relative to Z is used when interacting with Z when Y is not present, or vice versa. These maps are needed because of the competition and cooperation that takes place within social environments. The role of competition and tactical deception that takes place in these environments is emphasised by the social intelligence hypothesis's alternative name: the Machiavellian intelligence hypothesis (Byrne & Whiten 1988).

Without going into too much detail as to why Sterelny takes these environments to likely select for representations that can be used flexibly and be involved in off-line processing, the primary reason is that relevant information about other group members may not be available at the time it is required (2003, 52). Whether or not one member trusts another will depend upon their history together. This selects for the decoupling of perception and action and for the capacity to store information off-line so that it may be used later when it does become relevant.

1.2.4 Summary

Above I have outlined the HT. According to Sterelny, hostility creates three problems: it increases the variability within environments; it decreases the stability between proximal cues and distal events; and it increases the cost of action. Detection relies on signals that are reliable and stable¹¹. Detection systems may include redundancy strategies if signals are less reliable. Robust tracking systems integrate information from multiple cues to disambiguate environmental states. Decoupled representation is required when information is potentially relevant to many actions and must be represented off-line so that it may be used at a later time when it becomes relevant. By introducing the HT in the context of the ECT I have already begun to show how the two accounts fit together. I continue to do so below.

1.3 Taking the Theses Together

In this section, I argue that the ECT and HT give complementary accounts of the evolution of cognition. The purpose of this chapter is to sketch out Sterelny's account in order to lay the foundations for improving the continuity of Sterelny's account. I have laid these foundations above, but Christensen (2010) argues that the HT itself rests on a bad foundation. Christensen argues that, although Sterelny uses the ECT as a framework, the ECT and HT give conflicting stories of what drives the evolution of cognition. Furthermore, Sterelny (1997) criticises the Godfrey-Smith's externalism. For both these reasons Sterelny appears to undermine the HT by using the ECT as a framework. In section 1.3.1, I present Christensen's criticisms, and argue that the ECT explains *why* cognition functions to coordinate behaviour and the HT explains *how* it realises this function given the problems created by hostility. In section 1.3.2, I show that the ECT and HT no longer have opposing commitments.

¹¹ Some lineages, such as arthropods, may be constrained to rely upon detection systems even in hostile environments (Sterelny 2003, p. 28)

1.3.1 The Relations Between the Two Accounts

According to Christensen, cognitive evolution is driven by the problem of flexible behaviour in the ECT and by the problem of behaviour targeting in the HT. This is his take on the relations between the two accounts:

A plausible interpretation is that Sterelny sees Godfrey-Smith's account as explaining how detection systems evolve and his account as covering, not the pressures that add further detection systems, but selection for more complex forms than detection. (2010, p. 373)

Sterelny has long taken the position that cognitive complexification is driven by more than simply the need for flexible response to single cues (1995, 1997), which supports Christensen's claim. There is further support for Christensen's interpretation elsewhere: 'Peter Godfrey-Smith explores [the idea that minds are control systems] for very simple forms of plastic response; I am interested in extending it to more cognitively complex organisms' (Sterelny 2001, p. 260). Although this is Sterelny's take on how his account relates to Godfrey-Smith's, I propose another way of understanding their relations.

Christensen shares Sterelny's interpretation of the relations between the ECT and HT. Sterelny sees the ECT as covering simple cognition and the HT as covering complex cognition. I call this the **Extension Interpretation**; however, this interpretation is inaccurate. Although the models of behavioural control used by Godfrey-Smith (1996) are only simple, the ECT does apply to cases of complex cognition. For example, Godfrey-Smith uses the ECT to explain cognitive phenomena more complex than detection such as cognitive maps and social cognition (Godfrey-Smith 2002a).

Since the Extension Interpretation is inaccurate, I propose an alternative understanding of the relations between the ECT and HT, which I call the **Why and How Interpretation**. The two accounts address two different questions that are both related to the same problem.

Foundations

The ECT addresses the question, *why* do organisms need to respond flexibly? The HT addresses the question, *how* do organisms respond flexibly? Organisms respond flexibly because environmental variation punishes inflexibility (that answers the *why*), and they respond flexibly through detection in transparent environments or where they are otherwise constrained, and through multi-channel tracking and decoupled representation in more translucent environments (that answers the *how*). The HT appears to go beyond the ECT because it is a more specific thesis than the ECT, focusing on explaining how complex human cognition evolves, but contra the Extension Interpretation, the ECT is not only concerned with the first steps of cognitive evolution.

We should accept the Why and How Interpretation because it also accounts for the differences between the ECT and HT that are explained by the Extension Interpretation. According to Christensen, the ECT cannot predict a case in which behaviour is controlled through multiple channels, such as robust tracking. According to the Extension Interpretation, this is because robust tracking is a complex cognitive system beyond the scope of the ECT. According to the Why and How Interpretation, however, this is because robust tracking is beyond the scope of the ECT because Godfrey-Smith is only concerned with explaining *why* cognition is required: ‘The version of the [ECT] developed in this book does not require any particular account of *how* cognition works’ (1996, p. 127, author’s emphasis). Godfrey-Smith’s account does not cover the evolution of specific control systems—detection systems, robust tracking, or decoupled representation—because they are the *how*.

1.3.2 Externalism and the ECT

Following Christensen and assessing the relationship between the HT and ECT, we can identify another possible source of tension between the two theses: the ECT was originally presented as an externalist explanation, but Sterelny is a critic of externalism. According to the *symmetric* externalist, properties of organic systems can be explained in terms of the environment (what is external to the organic system), and properties of the environment can

be explained in terms of organic systems. According to the *asymmetric* externalist, properties of organic systems can be explained in terms of the environment, but no property of the environment can be explained in terms of organic systems. Although both Sterelny (1997, p. 185) and Menary (2007, p. 104) attribute asymmetric externalism to Godfrey-Smith, Godfrey-Smith states that the ECT does not have to be ‘associated with an extreme form of externalism’ (1996, p. 60). Instead, he states that a ‘version of Dewey’s claim will be accepted in this book’ (p. 132); Dewey, according to Godfrey-Smith, is a symmetric externalist (*ibid.*). Furthermore, Godfrey-Smith endorses an ‘*interactionist* view of the causal traffic between mind and the rest of nature’ (p. 196, author’s emphasis). We should adopt this non-externalist understanding of the ECT since Godfrey-Smith has abandoned his externalist commitments, stating that he ‘was wrong to see the ECT as worth defending in an externalist form’ (in Marshall 2014).

Since Godfrey-Smith no longer endorses an externalist ECT I will leave the (now irrelevant) details of Sterelny’s criticism of externalism aside in the interest of keeping this chapter as simple as possible. However, it may be worth briefly outlining why an abandonment of externalism does not have problematic implications for the ECT. Given that the ECT explains internal changes by way of external changes, a non-externalist ECT may seem robbed of its explanatory power. The non-externalist is, however, still able to distinguish between internal and external, but they do not have to make the strong distinctions that the externalist does (Sterelny 1997). The non-externalist can be more flexible when it comes to drawing a line between the environment and the organism in each case. For example, we might understand some ape trait as an adaptation to the pressures of ape social groups. In another, we might understand some trait of the social group as a whole and understand some trait of the members as an adaptive response to some phenomena external to that group (see also Sterelny 1996b). In the second case we are taking, as the target of our examination, what was taken to be external to our target in the first case.

Foundations

There is more to say on the fluidity of internal and external in the context of evolutionary thinking, but it cannot be addressed here. For the purposes of demonstrating that a non-externalist ECT is possible it is enough to show that the importance role of the environment in the ECT's explanation of biological can be assured if the thesis retains its adaptationist commitments. Adaptationism answers the question of *why* with regard to the function of a biological structure:

Only in considering what function the structure might have relative to its environmental niche, and given the constraints of natural selection and genotypical variation..., can we have a reasonable answer to the question. (Menary 2007, p. 111)

By retaining its adaptationist commitments, the ECT provides an explanation of cognitive evolution that remains dependent on the role of environmental complexity.

Although Sterelny claims that the boundaries between environments and organisms are indistinct in the context of evolution (1997, p. 186), he does not claim that the boundaries between the input and output of behaviour are similarly indistinct. In Chapter 2, I show that this leaves the HT vulnerable to criticisms like A&P's. It is also one of the aims of this thesis to show how the boundaries between behavioural input and output are fluid, which I do in Chapter 3.

1.3.3 Summary

Above I have shown that the ECT and HT are compatible and that they give complementary accounts of the evolution of cognition. The ECT explains *why* cognition evolves to coordinate flexible behaviour in response to relevant environmental variation. The HT explains *how* particular behavioural control systems coordinate this behaviour in environments in which information is cheap and perfect, and those in which information is expensive and imperfect. What exactly constitutes the environment and the organism is somewhat flexible. For

example, a primate social group could be understood as an organic system (the group) or the environment of another organic system, such as a member of the group.

1.4 Conclusion

The ECT and HT provide complementary accounts of the evolution of cognition. Behavioural responses that are not sensitive to relevant environmental variation—complexity—usually have poor results. Behaviour that is coordinated with the state of the environment is typically more beneficial to the organism that produces it.

Behaviour is coordinated through multiple systems. The simplest system, the detection system, is suited to environments in which the state of the environment can be reliably discriminated through a single channel; these environments are transparent. Hostility makes it difficult to discriminate the environment through a single channel; these environments are translucent. Robust tracking systems, which control behaviours by tracking the state of the environment through multiple channels, are required in these environments. Detection and robust tracking systems are directed toward single behavioural responses, but when information about the environmental state becomes relevant to multiple responses, organisms must represent this information in an action-neutral, or decoupled, way. One environment that selects for decoupled representation is the social environment, which is both translucent and highly complex.

In the next chapter, I defend Sterelny against Christensen's and A&P's criticisms.

Defense

In Chapter 1, I introduced the ECT and HT, and argued that they provide complementary accounts of the evolution of cognition. Recently, however, Christensen and A&P have argued that the HT does not square with the findings of empirical research, and that the HT should be rejected or dramatically revised.

This chapter has two aims. *The first aim is to defend the HT against Christensen's argument that decoupled representation evolves earlier than Sterelny claims.* In section 2.1, I present Christensen's interpretation of decoupled representation and argue that it is incorrect. In section 2.2, I present Christensen's argument that Sterelny's most complex form of behavioural control—decoupled representation—controls the behaviour of any organism more complex than a sponge. I also counter Christensen in this section, arguing that his claims are based on a misinterpretation of decoupled representation. In section 2.3, I present Christensen's argument that inference requires decoupled representation and his supporting evidence. I argue that his evidence does not overcome the Methodological Problem, so it is not evidence of decoupled representation.

The second aim is to show that the HT requires moderate revision to explain the evidence that Christensen and A&P present. In section 2.3, I show that Christensen's evidence cannot be explained as either decoupled representation or robust tracking. In section 2.4, I present A&P's argument that much of even the most basic behaviour is too complex to be controlled through detection. The HT requires revision to explain Christensen's and A&P's evidence. I make these revisions in Chapter 3.

2.1 Christensen's Interpretation of Sterelny

In this section I present Christensen's interpretations of decoupled representation and Sterelny's claims, and argue that these interpretations are incorrect. Christensen presents a detailed analysis of Sterelny's framework but misinterprets Sterelny's concepts and claims. Consequently, his arguments against Sterelny are founded upon these misinterpretations. In section 2.1.1, I present Christensen's interpretation of decoupled representation. In section 2.1.2, I present his interpretation of Sterelny's claims about when, and in what environments, decoupled representation evolves.

2.1.1 Christensen on Decoupled Representation

According to Christensen, a representation is decoupled if it: (Cdr1)¹² controls more than one behavioural response (2010, p. 381); and (Cdr2) is involved in **multi-conditionalisation**. Multi-conditionalisation refers to a form of behavioural control where behaviour is controlled in response to multiple 'states or entities (for instance, respond to the red light only in the white room)' (p. 384). Below, I argue that these conditions are inaccurate.

Condition (Cdr1)

Although decoupled representations evolve as response breadth increases, decoupled representations do not simply contribute to more than one behavioural response or task. As (dr1) indicates, decoupled representations are used flexibly, so they can contribute to *a great number* of behaviours and are *not linked* to any particular behaviours. Decoupled representations are not linked to particular behaviours because they often represent information that is functionally ambiguous—that is, it is unclear how the information can be used and what behaviour it should direct. Christensen, however, argues that once a perceptual

¹² The C stands for Christensen because these are the conditions he attributes to decoupled representation.

Defense

discrimination becomes linked to more than a single behavioural response it becomes a decoupled representation. According to Christensen, there may be forms of multi-conditionalisation that involve pluri-coupling rather than decoupling: ‘multiple perceptual discriminations are coupled to a given behaviour, none of which link to any other behaviour’ (p. 385). This seems to roughly describe robust tracking, however, Christensen claims that multi-conditionalisation is distinct from robust tracking because robust tracking improves an organism’s ability to track environmental variation while multi-conditionalisation improves their ability to respond correctly to that environmental variation. I take multi-conditionlisation to be a form of robust tracking because it involves controlling a single behaviour through multiple cues, meeting conditions (rt1) and (rt2). Although Christensen acknowledges the possibility of many-to-one couplings mapping in one direction (from perception to action), he does not acknowledge the possibility of these couplings mapping in the other direction (from action to perception)¹³. Perceptual discriminations may be involved in controlling multiple behaviours without being able to be used flexibly, so (Cdr1) is not a condition of decoupled representation.

Condition (Cdr2)

Although multi-conditionalisation may be necessary for the evolution of decoupled representation, it is not an instance of decoupled representation; it better describes robust tracking. Christensen’s reasoning is that once multi-conditionalisation becomes a widespread form of control, it is likely that reusing representations to control multiple behavioural responses becomes efficient. These representations will then meet (Cdr1). Christensen’s intuition about the efficiency of information reuse and decoupled representation is correct, but information is only likely to be reused if the organism has a broad band of response to a given

¹³ These kinds of mappings will be given more attention in Chapter 3.

environmental state. Many organisms do not need to reuse information because it is only useful in a specific problem domain.

2.1.2 Christensen on Sterelny's Commitments

According to Christensen (p. 391), Sterelny claims that: (a) decoupled representation evolves only in social environments like those of primates. From this it follows that Sterelny is committed to: (b) the late emergence of decoupled representation; and (c) the restricted emergence of decoupled representation. From this it follows that Sterelny is also committed to: (d) the late emergence of multi-conditionalisation; and (e) the restricted emergence of multi-conditionalisation. Christensen claims that this is problematic because multi-conditionalisation is likely to be widespread. Therefore, decoupled representation is likely to have evolved earlier than the primate social environment and is likely to have evolved elsewhere. I call this the **Multi-conditionalisation Problem**¹⁴.

The Multi-conditionalisation Problem rests on some false assumptions. Christensen takes Sterelny to claim (a) because

Sterelny reviewed research on navigation, tool use, ecological knowledge, and social cognition and found no clear evidence for decoupled representation, although he thought it plausible that great ape social cognition would involve decoupled representation. (Christensen 2010, p. 391)

However, since Sterelny finds no clear evidence for decoupled representation because of the Methodological Problem, a lack of evidence does not imply that decoupled representations do not exist in non-humans. Sterelny is not committed to claim (a):

¹⁴ A list of these problems has been provided for convenience in the appendix.

Defense

My bet is that a broad-banded response to resource information has evolved in our lineage *and that of others*, and that it has played an important role in the evolution of decoupled representation. (Sterelny 2003, p. 50, my emphasis)

...many translucent environments select for decoupled representation, for in many such environments information becomes available in a piecemeal fashion and without its immediate significance for action being apparent. (p. 78, my emphasis)

Given that Sterelny is not committed to (a), he not is committed to (b) and (c). Consequently, he is not committed to (d) and (e). Furthermore, he could not be committed to (d) and (e) since robust tracking is multi-conditionalisation and the evolution of robust tracking is neither late nor is it restricted.

The Multi-conditionalisation Problem appears to dissolve on theoretical grounds. However, Christensen argues that multi-conditionalisation and decoupled representation emerge *very early* and are *very widespread*. If he is correct then the HT does not cover a great deal of the history of cognitive evolution, making it a very limited thesis. I call this the **Multi-conditionalisation Problem***.

2.1.3 Summary

Above, I have shown that Christensen misinterprets both the conditions of decoupled representation and Sterelny's commitments. Christensen takes representations to be decoupled if they control more than one behavioural response. However, these representations are not decoupled unless they can be used to control both a great many responses and no responses in particular. Christensen takes Sterelny to claim that decoupled representation evolves only in the complex social environments of primates, which is problematic because multi-conditionalisation, and therefore decoupled representation, is likely to be widespread. Although Sterelny is not committed to this claim, Christensen provides evidence that he takes to support the argument that multi-conditionalisation and decoupled representation evolve

very early and are very widespread, which may still challenge the HT. I present this evidence below.

2.2 The Early Evolution of Decoupled Representation

In this section, I present Christensen's evidence of the early evolution of decoupled representation, and argue that this evidence does not suggest that multi-conditionalisation and decoupled representation evolved very early and are very widespread. Christensen argues that two phenomena, which occur within all 'neurally based sensorimotor systems' (p. 390), are instances of decoupled representation. If these phenomena qualify as decoupled representation then it supports Christensen's claim that decoupled representation evolves much earlier and is much more widespread than Sterelny claims. The HT would require major revision to overcome this problem, which I have called the Multi-conditionalisation Problem*. In section 2.2.1, I present Christensen's evidence. In section 2.2.2, I argue that this evidence does not support Christensen's claim because it does not even meet Christensen's weaker conditions of decoupled representation.

2.2.1 Decoupled Representation and Early Nervous Systems

Neuronal Convergence and Divergence

Myocytes, known as independent effectors, are cells that perform both sensory and motor functions in sponges (*Poriferans*). In *Cnidarians*, which are only slightly more complex than *Poriferans*, sensory and motor functions are separated and performed by sensory neurons and effector cells respectively. This separation has several benefits: it enables specialisation of sensory neurons, which affords superior sensory functioning; and it enables neuronal divergence and convergence.

Divergence refers to the innervation of multiple effector cells by a single sensory neuron and convergence refers to the innervation of a single effector cell by multiple sensory

Defense

neurons. According to Christensen, divergence meets (Cdr1) because it causes a sensory neuron to contribute to the action of different effector cells, which Christensen takes to be equivalent to different behaviours. Christensen considers whether it is a ‘stretch’ to characterise different effector cells as different behaviours, but claims that ‘there is no obvious reason to rule this out’ (p. 387). He also takes convergence to be an instance of multi-conditionalisation, and given that divergence and convergence occur simultaneously, Christensen takes this to meet (Cdr2).

Hierarchical Feature Analysis

Hierarchical feature analysis enables downstream integration of sensory information received by multiple receptor cells.

Neurons in the lateral geniculate nucleus (LGN) are sensitive to dot-shaped stimuli. Neurons in the striate cortex (or V1) called *simple cells* take input from multiple LGN neurons and are sensitive to oriented lines. (Christensen 2010, p. 387, author’s emphasis).

A single receptor cell receives very little sensory information, so hierarchical feature analysis, like that described above, is required for registration of a single environmental aspect, such as the hornet’s alarm pheromone described in Chapter 1. According to Christensen, hierarchical feature analysis involves representations that ‘will generally count as cases of both robust tracking and decoupled representation’ (p. 388). Hierarchical feature analysis, for Christensen, is a clear case of multi-conditionalisation, and is likely to involve decoupled representations because

...many feature representations can contribute to multiple behaviours. In the case of vision, for example, many states and entities will be represented in terms of combinations of location, movement, form, and colour information. (Christensen 2010, p. 388)

Therefore, these ‘feature representations’ meet (Cdr1).

2.2.2 Analysis of Christensen’s Evidence

Effector Cells and Behaviour

According to Christensen, different effector cells are different behaviours, and different sensory neurons register different environmental aspects. Therefore, neuronal divergence and convergence between sensory neurons and effector cells creates complex mappings between cues and behaviour. It follows that divergence and convergence involves representations that meet (Cdr1). Although I have already argued that (Cdr1) is not a condition of decoupled representation, divergence and convergence does not even meet this condition because effector cells are not equivalent to behaviours.

Christensen’s claim that effector cells are equivalent to behaviours is slightly ambiguous, so there are two ways of reading this claim: (a) a given effector cell *itself* constitutes behaviour; or (b) the *activation* of a given effector cell constitutes behaviour. The first reading is less charitable because cells could only constitute behaviour according to a definition of behaviour that is completely unlike any generally accepted definition. It would also imply that an organism with more than one effector cell occupies different behavioural states simultaneously and can only expand its behavioural repertoire by growing more effector cells.

The second reading is also problematic, and should be rejected for three reasons. First, although cells might be said to behave because they ‘do’ something (Dretske 1988), this does not make cell-level behaviour equivalent to organism-level behaviour. Although there is general disagreement over the exact definition of behaviour in biology, most do not consider

Defense

the activation of single cells as instances of behaviour¹⁵ (Levitis et al. 2009). According to Millikan, behaviour has a function, which would normally be achieved either through interaction with the environment or as a result of these interactions (1993, p. 137). Taking effector cell activation as behaviour, however, reveals nothing of the function of that behaviour and nothing about how this activation interacts with the environment.

Second, Christensen's unorthodox definition of behaviour is problematic in ways that Millikan's (1993) and Levitis et al.'s (2009) is not. For example, if multiple effector cells activate simultaneously then an organism must be said to be producing multiple behavioural responses simultaneously. Furthermore, these behaviours will not be observable without special equipment. Although I consider these to be problems, Christensen may disagree.

Third, there is no clear value in this unorthodox understanding of behaviour. Not only does Christensen not articulate the value of his notion of behaviour, he undermines it by using an orthodox notion of behaviour elsewhere, which he refers to as 'distinct overt behaviours' (p. 387). I reject the claim that effector cells are equivalent to behaviours because this problematic conception of behaviour is used only in an attempt to establish similarity between divergence and convergence and decoupled representation.

Features and Representations

Hierarchical feature analysis qualifies as multi-conditionalisation only if single features carry information about states or entities—cues, in the terminology of the HT. However, as Christensen states, many states and entities are *combinations* of features—dimensions, in the terminology of the HT. Multiple features must be combined for an organism to register a state or entity because single receptors register very little information:

¹⁵ This excludes cases in which the organism has only one cell; for example, the single-cell green algae discussed below.

The Evolution of Decoupled Representation

A given receptor cell is often highly ambiguous with respect to possible external signal sources, but this ambiguity can be reduced by downstream integration... (Christensen 2010, p. 388)

Multi-conditionalisation, according to Christensen, refers to instances where behaviour is controlled according to multiple ‘states or entities’ (p. 384). Although hierarchical feature analysis may be *involved* in multi-conditionalisation, it cannot be an *instance* of multi-conditionalisation. Hierarchical feature analysis is only the process of integrating information that is received by single receptor cells to register features, and then integrating those features to register a state or entity. If the organism has only one response to the registration of that state or entity then this process is part of a detection system, not multi-conditionalisation. Christensen’s ‘feature representations’ do not meet (Cdr1) either. According to Christensen, decoupled representations can control multiple behaviours; however, these ‘feature representations’ are limited, and, by themselves, are unable to control even a single behavioural response.

2.2.3 Summary

Above, I have argued that, contra Christensen, decoupled representation does not occur within ‘neurally based sensorimotor systems’ (p. 390). If it did, decoupled representation must have evolved much earlier and be much more widespread than Sterelny claims, so the HT would cover little of the history of cognitive evolution—this was the Multi-conditionalisation Problem*.

Christensen argues that neuronal convergence and divergence involve complex mappings between perception and behaviour. However, different effector cells are not different behaviours. Therefore, neuronal convergence and divergence are not instances of complex mappings between perception and behaviour. Christensen also argues that

hierarchical feature analysis is an example of multi-conditionalisation, which demonstrates that information reuse is likely to be widespread. However, hierarchical feature analysis is not an instance of multi-conditionalisation. This is because hierarchical feature analysis involves the integration of information to register single environmental states or entities; it does not necessarily involve the combination of multiple environmental states or entities, which is necessary for multi-conditionalisation. Furthermore, the ‘feature representations’ involved in hierarchical feature analysis do not meet (Cdr1) because they are too limited to control even a single behavioural response. I have dissolved the Multi-conditionalisation Problem*, but Christensen makes another argument for the early evolution of decoupled representation. I present this below.

2.3 Decoupled Representation and Inference

In this section, I present Christensen’s evidence of decoupled representation in rats, and argue that, although the behaviour of the rats does not meet the conditions of decoupled representation, the HT cannot explain their behaviour without revision. According to Christensen, it is odd that Sterelny finds no evidence of decoupled representation given that it is so widespread. He claims that there is evidence of decoupled representation, and reports two experiments, which he takes to provide this evidence. Although I have shown that decoupled representation is not as widespread as Christensen claims, his evidence may still challenge the HT. In section 2.3.1, I present Christensen’s evidence. In section 2.3.2, I argue that this evidence does not overcome the Methodological Problem, but that the HT cannot explain this kind of behaviour without revision.

2.3.1 Rat Inference

According to Christensen, performing inference¹⁶ requires extracting and integrating information across multiple experiences. Christensen claims that inference requires decoupled representation because the information used must be represented in an action-neutral way, or it could not contribute to novel responses. Since many animals can complete inference tasks (Premack 2007), it follows that decoupled representation is more widespread than the HT indicates. I call this the **Inference Problem**. Christensen supports his claim that inference requires decoupled representation with two studies on rat cognition.

The Transitive Inference Task

The first study involves a transitive inference task (Dusek and Eichenbaum 1997). A transitive relation is one in which, if an element A is in a relation x to an element B, and B is in a relation x to an element C, then A will also be in a relation x to C (Allen 2006). For example: if A is bigger than B, and B is bigger than C, then A is bigger than C. Completion of a transitive inference task requires recognition of the shared relation between the non-adjacent elements.

Subjects are trained on pairs of adjacent relations until they reach an adequate level of performance. They are then tested on a novel pairing of non-adjacent elements. Dusek and Eichenbaum use the standard five-element test. During the training phases of a three-element test, the end element A is never conditioned against, and the end element C is always conditioned against. Therefore, when subjects are presented with the pair AC, the choice of A over C could simply be a result of this conditioning. In the five-element test, however, the elements in the novel pairing (BD) have a history of being equally conditioned for and against.

¹⁶ In this thesis I take it that performing inference does not necessarily require symbolic representations or language (Vigo & Allen 2009).

Defense

In Dusek and Eichenbaum's experiment, rats were presented with pairs of sand boxes distinguished by the odour of spices buried in each box. The reward scheme with which the rats were trained was as follows: $A > B > C > D > E$. The rats were trained with each of the 'premise' pairs (AB, BC, CD, DE), and then presented with a novel pairing (BD). Rats with disconnected hippocampuses chose the correct box at chance levels, but 88% of normal rats correctly dug in sand box B. According to Christensen, this suggests that the normal rats were able to represent the spice hierarchy.

The Morris Water Maze Task

The second study involves a Morris water maze task (Eichenbaum 2000). In this task, rats are placed in a pool of opaque water and must swim onto a platform hidden just beneath the water's surface. Because the platform is hidden the rats must navigate according to landmarks located around the room in which the pool is situated. In this experiment, rats with disconnected hippocampuses could only learn to locate the platform when they began the task from the same position. Normal rats, however, learned to locate the platform irrespective of their starting position. According to Christensen, this suggests that the normal rats represent the position of the platform in relation to the landmarks surrounding the pool.

Evidence of Decoupled Representation

According to Christensen, completing these tasks requires a hippocampus because of its association with declarative memory (memory of *what*, not *how*), which is 'constructed by extracting relational information across multiple experiences' (p. 392). Christensen claims that the normal rats can complete these tasks through the construction of **structural representations**¹⁷ (s-representations). S-representations are structurally or organisationally

¹⁷ Christensen actually uses the term 'model-based representation' (2010, p. 393), but this fits the definition of Ramsey's s-representation.

isomorphic¹⁸ to the systems they represent. This isomorphism can be exploited to make inferences about the target system (Ramsey 2007, p. 71; see also Swoyer 1991). Although I agree with Christensen that completing these tasks requires s-representations¹⁹, these s-representations are not necessarily decoupled.

2.3.2 Overcoming the Methodological Problem?

S-representations are required for the successful completion of the transitive inference and Morris water maze tasks. However, these experiments do not overcome the Methodological Problem because the subjects' responses are restricted. In the Morris water maze task, rats track the positions of multiple objects in the room and respond by swimming to the hidden platform. The rats with disconnected hippocampuses may be unable to solve this task because of a robust tracking failure—they cannot find the hidden platform because they are unable to track and integrate the information about the positions of the landmarks. The results of this experiment may suggest that the rats with disconnected hippocampuses are failing to form representations, but it does not suggest that these representations must be decoupled. All the information about the relative positions of the landmarks is used to solve a single problem—the problem of finding the hidden platform. Likewise, the completion of the transitive inference task may require a representation of the spice hierarchy, but this representation is used to solve only in a single problem, and the rats' behavioural responses are restricted to

¹⁸ Isomorphism conditions are often criticised for being indeterminate, but Ramsey provides a counterargument to these criticisms (2007, pp. 93-96).

¹⁹ According to some critics, it is unclear whether representation is required for the successful completion of these tasks (for criticism of the Morris water maze task see Shettleworth 1998; Zappettini & Allen 2013; for criticism of the transitive inference task see Premack 2007; Zentall 2001; De Lillo et al. 2001). However, the differing performance of the rats with disconnected hippocampuses and the control rats suggests that representations are required for completion of these tasks.

Dig In This Box or Dig In That Box, which are only slight variations of the behaviour Dig For Food.

Christensen's evidence does not overcome the Methodological Problem, so it is unclear whether these representations meet the conditions of decoupled representation. Furthermore, this kind of inference is not likely to require decoupled representation because the rats are trained to solve particular problems by performing particular behaviours, and are then tested on the same kinds of problems where they perform the same behaviours. Because they learn the structures required for completing the task (the spice hierarchy or the layout of the room) in such an action-specific way, it seems unlikely that they would represent them in an action-neutral way.

Despite all this, Christensen's intuitions are correct: robust tracking systems are too simple to control the rats' behaviour. Recall that robust tracking involves pp-representation, which cannot be involved in off-line processing. However, inference is a process that is, at least, **minimally off-line**²⁰. Recall that on-line processing is immediate, while off-line processing takes place in 'in the absence of task relevant input and output' (Wilson 2002, p. 626). Minimally off-line processing takes place in the absence of *some* task relevant input *or* output. So in the case of the transitive inference task, correctly choosing B over D requires experience with elements B, C, and D, but the input of element C is absent when the novel pair is presented. Similarly, when training on pairs BC and CD, the novel output of choosing B over D is absent. If robust tracking involves pp-representation then the HT struggles to account for the kind of problem-specific inference performed by the normal rats. I call this the **Inference Problem***. I address this problem in Chapter 3.

²⁰ According to the predictive brain hypothesis, perception just is a process of inference (Hohwy 2013).

Unfortunately, addressing the predictive brain hypothesis is beyond the scope of this thesis, but an examination of Sterelny's account of the evolution of decoupled representation in the context of this hypothesis could undertaken in future work.

2.3.3 Summary

Above, I have argued that Christensen's evidence does not overcome the Methodological Problem. Furthermore, it is unlikely that representations are decoupled when the structures they represent are learned in a task-specific way and are used only to solve the particular problems. Contra Christensen, decoupled representation does not evolve earlier and is not more widespread than Sterelny claims. However, Christensen is right in identifying that the HT cannot explain problem-specific²¹ inference. I make the necessary revisions in Chapter 3. Below, I present A&P's argument that robust tracking evolves earlier than Sterelny claims because detection systems are too simple to control much of even the most basic behaviour.

2.4 The Limitations of Detection

In this section, I present A&P's argument that detection systems are too simple to control a great deal of basic behaviour, and concede that, without revision, the HT cannot explain A&P's evidence. According to A&P, Sterelny addresses the question, 'How did decoupled representational states evolve' (in preparation, p. 3). However, like Christensen, A&P take Sterelny to be committed to the claim that decoupling only occurs in social environments: 'It is *the* step that drives the decoupling of perceptions and behaviour' (p. 6, authors' emphasis). As I have indicated above, however, this is not the case. A&P do not address decoupled representation; instead, they focus on basic behavioural control, and argue that Sterelny's account is not even 'roughly correct' because it focuses on behaviour rather than internal mechanisms (p. 2). Similarly to Christensen, A&P argue that examining internal mechanisms demonstrates that the HT is inaccurate.

A&P argue that Sterelny's account of basic behavioural control is incorrect. This is because a great deal of basic behavioural control requires tracking (p. 37). This 'requires a

²¹ Note that 'problem-specific' does not imply 'domain-specific'. By problem-specific I simply mean that the information is only relevant to a single problem or kind of problem.

continuously up-dated error signal', but detection systems only signal the 'presence or absence of a property' (p. 13). A&P do acknowledge that detection systems are probably well suited to some tasks such as 'bringing organisms to an abrupt halt... making an individual duck or flinch, or... vomit' (Akin & Pollon in preparation, p. 14). However, according to A&P, a great deal of basic behaviour cannot be explained as a detection system. I call this the **Tracking Problem**. In section 2.4.1, I present A&P's evidence of the limitations of detection systems. In section 2.4.2, I argue that this evidence cannot be explained as robust tracking, and concede that the HT requires revision to explain it.

2.4.1 The Behaviour of Green Algae

A&P support their claim that basic behaviour requires on-going coordination by examining the mechanism that controls the behaviour of green algae (*C. reinhardtii*), which are single cell organisms. The alga receives input through a photosensitive eye-spot on its side, and produces output in the form of swimming, which it performs with two flagella positioned on the top of the cell. This organism photosynthesises, but its DNA can be damaged by high intensity UV light. On a simple view, the alga looks like a detection agent with two detection systems. The first system controls positive phototaxis: light causes the alga to swim toward stronger light. The second system controls negative phototaxis: intense UV light causes the alga to swim toward weaker light. The alga's behaviour, however, is not this simple.

The purpose of tracking in green algae is to locate what A&P refer to as the 'sweet spot' (p. 11). This is where light is intense, but where there is not too much harmful UV light. The alga's tracking has a searching phase and a homing phase. The goal of the searching phase is to locate directional light, which is 'a sine wave with appropriately large amplitude' (p. 11). This sine wave is created as the alga rotates on its polar axis, bringing its eye-spot to face toward, and then away from, light. This creates a contrast signal between light in one place and another. A high contrast signal clearly indicates the direction from which light is coming. Once a high contrast signal has been found, the homing phase can begin. The goal of

The Evolution of Decoupled Representation

the homing phase is to locate a constant high intensity signal. The alga achieves this by repositioning itself so that it is beneath the light source. It then swims in the direction of the light based on the increasing intensity of the light signal.

A&P reach the following conclusion:

At each moment, *C. reinhardtii* uses the continuous contrast signals created by its own movement to refine/correct its heading and body orientation. This is the very nature of closed-loop tracking: On-going sensory input adjusts on-going movement and, in the other direction, on-going movement adjusts the sensory signal relative to the organism's current behaviour. Even when *C. reinhardtii* is switching between searching and homing, it merely substitutes one form of tracking for another. So if a detection system signals the presence or absence of a property, and a tracking mechanism requires a continuously up-dated error signal, it is hard to see how detection systems *could* drive genuine tracking behaviour. (in preparation, p. 13, authors' emphasis)

As far as A&P see it, if any organism should be a detection agent, it is this alga. Yet the behaviour of this organism requires on-going coordination with the environment, so its behaviour cannot be controlled by a detection system. This supports A&P's claim that the simplest form of behavioural control cannot be the detection system. It follows that robust tracking must evolve earlier than Sterelny claims to control this kind of behaviour.

2.4.2 Response to Akins and Pollon

A&P and Christensen both argue that internal mechanisms suggest that Sterelny's control systems are inaccurate. However, Sterelny's control systems *are* defined in more course-grained terms than internal mechanisms (environmental variation, cues, and behavioural responses). Therefore, internal mechanisms can only be used to assess his control systems once they have been contextualised in these terms. However, A&P's analysis *does*

demonstrate that the green alga's behaviour is controlled through more than the registration of the presence or absence of an environmental aspect.

The HT may be able to explain this behaviour if it can be controlled by a robust tracking system. However, the alga's behaviour cannot be controlled by a robust tracking system because it does not pass the Explanation Challenge. Recall that, according to the Explanation Challenge, if an actual-sequence explanation of behaviour has the same predictive power as a robust-process explanation then this behaviour is being controlled by a detection system. An actual-sequence explanation of the alga's behaviour has the same predictive power as a robust-process explanation. The processes of searching and homing unfold in the same way every time they occur; there are no islands of resilience. The alga's behaviour is always produced in response to particular proximal stimuli. The high contrast signal and high intensity signal are involved in every instance of searching and homing.

The behaviour of the alga does not pass the Explanation Challenge. Therefore, it is not controlled by a robust tracking system. The HT must be revised to solve the Tracking Problem. I make these revisions in Chapter 3. A&P, however, argue that the HT should be rejected, not revised. In Chapter 3, I also argue against this claim, showing that a new kind of detection enables the HT to explain their evidence. Furthermore, by expanding upon the details of the continuity between decoupled representation and simpler forms of behavioural control, I show that Sterelny's story is closer to A&P's than they claim.

2.4.3 Summary

Above, I have presented A&P's argument that detection systems are too simple to explain basic behavioural control, which requires on-going coordination with the environment. They use the phototaxis of green algae as evidence to support their argument. The HT cannot account for this behaviour without revision. This is the Tracking Problem and will be addressed in Chapter 3.

2.5 Conclusion

According to Christensen, the HT does not explain most of the history of the evolution of cognition. According to A&P, it does not even explain the evolution of basic cognition. It would seem that there is very little that it does explain. I have defended against Christensen's arguments, and shown that the HT is not in need of major revision. However, Christensen's and A&P's evidence demonstrates that certain aspects of the HT do indeed require alteration.

Christensen argues that decoupled representation evolves very early and is very widespread, but this argument is based on misinterpretations of the conditions of decoupled representation. Christensen also argues that completing inference tasks requires decoupled representations, but his supporting evidence does not overcome the Methodological Problem. Furthermore, it seems unlikely that decoupled representations are required for the kind of inference featured in his evidence. However, this kind of inference cannot be explained by the HT because the representations involved do not meet the conditions of either decoupled representation or pp-representation. This is the Inference Problem*.

A&P argue that detection systems are not the most basic form of behavioural control. The behaviour of green algae does not pass the Explanation Challenge, so cannot be explained as robust tracking. Therefore, the HT cannot explain this behaviour. This is the Tracking Problem. Overcoming both the Inference Problem* and the Tracking Problem is be addressed in the next chapter.

Revision

In chapter 2, I argued that Sterelny's account of cognitive evolution does not require a major overhaul because decoupled representation does not evolve early. However, Christensen presents evidence of *complex* behavioural coordination that cannot currently be explained by the HT. This is the Inference Problem*. Additionally, A&P present evidence of *basic* behavioural coordination that cannot be explained by the HT. This is the Tracking Problem.

This chapter has three aims. *The first aim is to revise the HT to explain Christensen's evidence.* Currently, the HT cannot explain the performance of rats in inference tasks because it requires processing that is, at least, minimally offline. Robust tracking, however, appears to involve pp-representation, which is exclusively on-line. In section 3.1, I argue that robust tracking, because it integrates information across multiple channels, cannot involve only pp-representation. Instead, it involves another kind of representation, which I introduce. In section 3.2, I show how the revised HT explains Christensen's evidence and overcomes the Inference Problem*.

The second aim is to revise the HT to explain A&P's evidence. Currently, the HT cannot explain much basic behaviour because it requires on-going coordination with the environment. This is the Tracking Problem. I introduce a new kind of detection to explain this behaviour. The HT presents the process of behavioural control as having distinct boundaries between input and output phases. However, input and output often occur simultaneously and are temporally extended, so the boundaries between them are sometimes indistinct. In section 3.3, I revise the HT to accommodate the on-going coordination that plays an important role in both basic and more complex forms of behavioural control. In doing so I overcome the Tracking Problem.

The third aim is to expand upon the details of continuity in the HT. In section 3.4, I argue that robust tracking is the origin of decoupled representation because it involves

representations that are minimally off-line. I also argue that Christensen's intuitions about the role of multi-conditionalisation and the efficient reuse of information are accurate, but that information reuse is rarely efficient. Information about the environmental state must be represented with precision if it is to be reused to control different behaviours, but this is likely to occur only in environments in which novel problems are common and action is expensive. The evolution of decoupled representation can be understood as the evolution of precise and accurate representational capacities. By expanding upon the details of the continuity between basic and complex behavioural control, I show, contra A&P, Sterelny's account is not inaccurate.

3.1 Coupled Representations

In this section, I show that robust tracking does not involve pp-representation but another type of representation, which I introduce. In Chapter 2, I presented Christensen's evidence of rats completing inference tasks. I showed that their performance did not qualify as decoupled representation. However, the rats' performance cannot be explained as robust tracking. This is because completing the tasks requires minimally off-line processing and robust tracking involves pp-representation, which is limited to on-line processing. This is the Inference Problem*. In section 3.1.1, I show that detection and robust tracking do not involve the same kinds of representation according to the Explanation Challenge. In section 3.1.2, I argue that robust tracking systems involve minimally off-line processing but detection systems do not, which means that robust tracking cannot involve pp-representation. In section 3.1.3, I introduce another category of representation to accommodate for robust tracking, which enables the HT to overcome the Inference Problem*.

3.1.1 The Explanation Challenge and Representation

In chapter 1, I introduced the Explanation Challenge. Recall that, according to the Explanation Challenge, if an actual-sequence explanation of behaviour has the same predictive power as a robust-process explanation of that behaviour then the behaviour being explained is being controlled by a detection system. Sterelny claims that the Explanation Challenge can also be used to determine whether a representational explanation of an organism's behaviour is appropriate (1995, p. 207-211)²². According to Sterelny, a representational explanation is appropriate when a system is directed toward environmental variation, not the channels through which that system tracks the environmental variation. Detection systems are directed toward particular stimuli and can be explained without giving a robust-process explanation. Robust tracking systems, on the other hand, are directed toward the source of the stimuli—the environmental variation—so cannot be explained without a robust-process explanation. This is because the environmental variation remains constant across different actual-sequence explanations but the channels do not. Therefore, if a system passes the Explanation Challenge, it is both robust and representational.

The Explanation Challenge indicates that detection systems are not representational; Sterelny also suggests this elsewhere (2000, pp. 210-211, 2003, pp. 17-20). However, I remain agnostic toward whether detection involves representation to avoid engaging in cumbersome debates regarding non-representational approaches to cognition (see Shapiro 2011, ch. 5 for a brief overview). To support my argument, it is enough to show that the Explanation Challenge indicates that Sterelny takes detection and robust tracking to differ with respect to representation, despite his suggestion that they both involve pp-representation.

²² I have omitted the details of why a representational explanation is appropriate in these cases because of space limitations and because my argument does not rest on the Explanation Challenge.

The Explanation Challenge is problematic because detection systems with redundancy also pass the challenge. Indeed, any multi-channel²³ system passes the Explanation Challenge. However, as I demonstrate below, not all multi-channel systems are representational equals. Instead of using the Explanation Challenge, I examine the role of cross-channel interaction in multi-channel systems to show that detection and robust tracking systems involve different kinds of representation.

3.1.2 Cross-channel Interaction

Types of Cross-channel Interaction

Detection systems with redundancy are characterised by a lack of cross-channel interaction. The behaviour controlled by these systems is triggered by any registration, so interaction is unnecessary. This can also take place in cases of robust tracking. For example, reed warblers reject eggs that appear in the nest before they lay their own. In this case, no cross-channel interaction is necessary to trigger the rejection.

Two types of cross-channel interaction can take place in multi-channel systems. The first type is *exploratory interaction* and is characteristic of robust tracking systems: ‘sometimes the signals that animals use in robustly tracking a feature of their environment are a direct result of their own exploratory behaviour’ (Sterelny 2001, p. 272). After an organism registers a cue, they may take action to gather more information about their environment. For example, if an organism sees an intruder approaching, they may stand up to better view the behaviour of the intruder.

The second type is *excitatory interaction* and is also characteristic of robust tracking systems: ‘the presence of one cue... makes other cues... more salient’ (Sterelny 2003, p. 28

²³Multi-channel systems include robust tracking systems and detection systems with redundancy. These are systems that are sensitive to environmental variation through multiple channels.

footnote). Following registration of one cue, the organism's sensitivity to other cues increases, so the organism then reacts more strongly to these other cues.

The Requirements for Cross-channel Interaction

Although detection systems with redundancy and robust tracking systems are both multi-channel systems, cross-channel interaction occurs only in robust tracking systems. This is because cross-channel interaction cannot take place if registration of a cue always leads directly to the response controlled by the system. For example, when Reject Egg (a response controlled by a robust tracking system) is produced in response to both a size and a colour cue, the first cue does not trigger the rejection response until the second cue has also been registered. For this to be the case the reed warbler must be able to represent the information without acting upon it (the output is absent), but this information is still represented in a problem-specific or behaviour-specific way. This is an instance of minimally off-line processing. Detection systems, however, are entirely on-line because registration immediately leads to the response controlled by the system. If robust tracking involves minimally off-line processing and detection involves only on-line processing, then it follows that robust tracking systems and detection systems—even detection systems with redundancy—have different representational statuses.

3.1.3 Coupled Representations

As I stated in Chapter 1, Sterelny suggests that robust tracking, like detection, can be thought of as involving pp-representation (Sterelny 2003, p. 29). However, robust tracking systems integrate information across different channels. This means that a robust tracking system's registrations do not immediately trigger the behavioural response controlled by the system. Therefore, unlike pp-representation, robust tracking is not connected directly to action. Instead, information is represented minimally off-line, so robust tracking cannot involve only pp-representation because they are exclusively on-line.

Robust tracking cannot involve only pp-representation, nor does it involve decoupled representation because it controls specific behavioural responses. Therefore, robust tracking must involve another kind of representation, which I call **coupled representation**²⁴. A representation is coupled if it: (cr1) is used to control particular behaviours and cannot be used to control a great number of behaviours; and (cr2) can be involved in minimally off-line processing.

3.1.4 Summary

Above, I have shown that there must be another kind of representation in the HT, which I call coupled representation. Coupled representation must be a part of the HT because robust tracking cannot involve pp-representation and does not involve decoupled representation. Below, I apply the notion of coupled representation to Christensen's evidence, and show that I have solved the Inference Problem*.

3.2 Solving the Inference Problem*

*In this section, I show that my notion of coupled representation explains Christensen's evidence and solves the Inference Problem**. There is nothing in the definition of s-representation about the number of behaviours they can control, so s-representation should be compatible with the conditions of both coupled and decoupled representation. This would mean that the s-representations used by the rats to solve the transitive inference and Morris water maze tasks *could* be coupled. In this section I want to support the stronger claim that the s-representations used by the rats *are* coupled. In section 3.2.1, I argue that robust tracking systems perform inferences²⁵, some of which involve s-representation. In section 3.2.2, I

²⁴ It is possible that my coupled representations are similar to Clark's (1997) action-oriented representations, but it is beyond the scope of this thesis to explore this in detail.

²⁵ These inferences are typically about the likelihood of a state obtaining given particular cues.

Revision

consider two examples of s-representation that Sterelny claims are possible origins of decoupled representation: social maps and resource maps. In section 3.2.3, I argue that the s-representations involved in Christensen's evidence are more like the coupled s-representations than the social maps and resource maps. In doing so I show that my revisions enable the HT to overcome the Inference Problem*.

3.2.1 Coupled S-Representation

The rats in Christensen's evidence perform inference without using decoupled representation. Examining two of Sterelny's examples of robust tracking shows that robust tracking systems also function to perform inference.

Case 1: Reed Warbler Egg Discrimination

Reed warblers cannot discriminate between their own eggs and cuckoo eggs through a single channel. Although the state of the egg is discriminated through multiple channels, the state of the egg is not *dependent* on the other elements of the environment to which the reed warbler is sensitive. Using two channels, such as size and colour, improves the reed warbler's chances of producing the right response because large size and odd colouring often co-vary with the state Cuckoo Egg. However, the egg is not a cuckoo egg *because* of its size and colour; it is not a cuckoo egg because of how these two aspects are *related* to each other. This system infers something about the likelihood of the state of the environment based on certain cues, but it does not infer something about the likelihood of the state of the environment based on a structure in the environment. This is unlike the threat discrimination system of the piping plover.

Case 2: Piping Plover Threat Discrimination

Piping plovers (*Charadrius melodus*) are shore-nesting birds whose nests are vulnerable to predators (Ristau 1991). They use multiple channels—the intruder's behaviour, the intruder's

position, the nest's position, and its own position—to discriminate threatening intruders from non-threatening intruders. These aspects not only identify an intruder as a threat, but the state of the intruder (Threat or Not Threat) is dependent upon the *relations* between these aspects. An intruder must occupy a certain position relative to the nest to be a threat; it is unlikely that an intruder is a threat if they stand in the Far From relation to the nest. This system infers the likelihood of something based on a structure in the environment.

The best strategy piping plovers have against these predators is distraction. They have a number of distraction displays, one of which is the broken wing display. Ristau shows that in almost all broken wing displays, the intruder is further from the nest at the end of the display than at the beginning of the display. Furthermore, in most displays, the most direct path from the intruder to the displaying piping plover never brings the intruder closer to the nest. For this to be the case the piping plover must track the relations not only between itself and the nest, and itself and the intruder, but also between the intruder and the nest. These might be thought of as akin to very simple versions of the third-party relations that social primates track.

Both of the cases presented above show that Sterelny's paradigm examples of robust tracking both involve making inference. This means that, although the systems are directed toward specific actions, they must involve at least minimal off-line processing. Of particular interest is the second case, which shows that some robust tracking systems function to track relations between environmental aspects. Following Christensen and taking tracking relations between environmental aspects to require s-representation, the second case shows that some robust tracking systems must involve s-representation given that they track these relations.

3.2.2 Decoupled S-representation

Two of Sterelny's examples of possible origins of decoupled representations are social maps and resource maps. Since both maps track relations between environmental aspects they are s-

representations²⁶. The point of these cases and the piping plover case above is to show that s-representations are part of Sterelny's account, and that some s-representations can control multiple behaviours (the cases below), while others only control single behaviours (the case above). These s-representations will then be compared with those from Christensen's evidence.

Case 1: Social Maps

Social maps are the prime example of the possible origins of decoupled representation (Sterelny 2003, pp. 52-53). Since they are used to track members and their relations, social maps are s-representations. In Chapter 1, I stated that social maps are likely to be required in complex social environments, like those of primates, because members of the group need to track third-party relations: how member X acts toward member Y may be affected by how X or Y stand in relation to member Z. Unlike the piping plover's s-representation, however, these social maps can be used to direct a number of responses. For example, an interaction with the previous alpha male (Y) may result in the current alpha male (Z) standing in the Stronger Than relation to Y. Consequently, a female group member (X) may act differently toward Y and Z compared to when Z stood in the Weaker Than relation to Y.

Case 2: Resource Maps

Resource maps²⁷ are also possible origins of decoupled representations (Sterelny 2003, p. 41). Scrub jays (*Aphelocoma californica*), a species of caching bird, navigate toward high quality food (Barrett 2011). Food quality is dependent upon the type of food (Worm or Peanut) and the amount of time since that food has been cached (<48 Hours or >48 Hours). When the time since caching is <48 Hours, the food type Worm stands in the Superior To relation to the food

²⁶ Maps are paradigm cases of s-representation (Ramsey 2007, p. 71).

²⁷ By resource maps, I mean maps of an organism's territory including important resources such as shelter, food, and water.

type Peanut, and the scrub jays respond with Pursue, navigating toward the cached worms. When the time since caching is >48 Hours, the food type Worm stands in the Inferior To relation to the food type Peanut, and the scrub jays respond with Ignore, navigating toward the cached peanuts. Resource maps can be used to track relations across space and, as this example shows, across time. Again, unlike the piping plover's s-representation, but like the primate's social map, resource maps can be used to direct multiple responses.

3.2.3 S-representation and Response Breadth

In Chapter 2, I argued that Christensen does not present evidence of decoupled representation. This is because there is no evidence that the rats' s-representations could control a great many behaviours, so Christensen's evidence does not overcome the Methodological Problem. In this section, I argue that the s-representations in Christensen's evidence are more like the coupled s-representation described above than the primate's social map or the scrub jay's resource maps. This shows that the HT can account for the rats' behaviour, and, therefore, that it can overcome the Inference Problem*.

Response Breadth and Environmental Complexity

In Chapter 1, I defined response breadth as 'the capacity to respond in more than one way to the registration of a focal aspect of the environment' (Sterelny 2003, p. 35 footnote). However, there is another way of understanding response breadth:

...we can suppose that each agent has exactly one way of behaving, given the environmental state it finds itself in. Increase in response breadth is increase in the fineness with which it partitions its environment. (*ibid.*)

Revision

Sterelny states that there is probably no ‘substantive difference’ between the two ways of thinking about response breadth (*ibid.*), but thinking about it in the second way is useful for demonstrating the connection between environmental complexity and decoupling.

The scrub jay chooses the cache to which it navigates according to both the food type and the time since caching. The aspects tracked by the scrub jay expand from Worm and Peanut, to Worm & <48 Hours, Worm & >48 Hours, Peanut & <48 Hours, and Peanut & >48 Hours. Although this is an idealisation and a simplification (it omits details about the distance of the cache from the scrub jay), it demonstrates that partitioning the environment responds to the amount of environmental variation that is relevant to the organism.

The ECT predicts that behavioural complexity increases in response to relevant environmental complexity, so an increase in response breadth follows from environmental partitioning. For example, before Worm is partitioned into Worm & <48 Hours and Worm & >48 Hours, the scrub jay has only one response to this state (excluding details about distance). After Worm is partitioned, the scrub jay no longer responds inflexibly to that state; rather, its response depends on how Worm relates to <48 Hours and >48 Hours. This illustrates that partitioning increases the number of relations between aspects of the environment.

If an environment is not partitioned finely then there are few relations between environmental aspects. Consequently, response breadth is narrow. Similarly, s-representations that contain few aspects and relations afford few responses. The piping plover threat discrimination system tracks only a few aspects and relations. The aspects are the intruder, the nest, and itself (the piping plover), and the relations are the distances between them. The primate social map, on the other hand, contains many aspects and their many relations. As a result these s-representations can control a great many behavioural responses.

Christensen’s Evidence

There is no evidence that the rats’ performance involved decoupled representation because their responses were restricted by the experiment. Furthermore, comparing the rats’ s-

representations to those described above illustrates that they do not exhibit the properties of either the social or resource maps.

Christensen shows that the transitive inference and Morris water maze tasks require tracking relations, and I agree with him that this kind of tracking requires s-representation. In the case of the transitive inference task, the location of the food depends upon the spice hierarchy. There are five aspects (A, B, C, D, E) in this representation and two relations (Superior To, Inferior To). These representations are probably simpler than the piping plover's because the piping plover tracks a structure that changes continuously. Unlike the piping plover's representation, which has to be somewhat flexible, the spice hierarchy remains the same. Social maps and resource maps also change, but all the relations between the aspects in the map of the water maze are fixed, except the position of the rat. In virtue of the different relations they track, social maps afford a great many responses, such as Mate With, Cooperate With, Steal From, Deceive, Avoid and so on. Representations of the spice hierarchy and the water maze, however, afford few responses because they contain few relations.

The representations that explain the performance of the rats in Christensen's evidence do not appear to exhibit the properties of social and resource maps: they are *rigid*; they *contain few relations and aspects*; and, consequently, they *afford few responses*.

These s-representations are relevant to particular problems and control specific responses. The representation of the spice hierarchy is used only to solve the problem of spice selection, and the representation of the water maze is used only to solve the problem of direction selection. These representations are minimally off-line because the information gathered from previous experiences is used in (somewhat) novel experiences. However, these novel experiences are not novel problems. The rats learn the structures they represent as they solve particular problems. Completing the novel test requires representing the same structures as before and nothing new. The information is used in the same way, so this information can

Revision

be represented in a problem-specific or action-specific way. Coupled representations are involved in minimally off-line processing and are action-specific, so the s-representations in Christensen's rat evidence meet the conditions of coupled representation and can be explained by Sterelny's account. This solves the Inference Problem*.

3.2.4 Summary

Above, I have shown that some robust tracking systems track the relations between environmental aspects, so they are likely to involve coupled s-representation. The rats' representations have more in common with coupled s-representations than they do with s-representations that are possibly the origins of decoupled representation. This supports my earlier claim that completion of these tasks does not require decoupled representations, but it also shows that the HT can now account for their behaviour, which solves the Inference Problem*. Next I revise the HT to solve the Tracking Problem.

3.3 On-going Coordination

In this section, I introduce a new type of detection similar to a detection cascade, and argue that the boundaries between the input and output phases of behavioural control should be understood as more fluid, which solves the Tracking Problem. A&P demonstrate that basic behavioural control often involves on-going coordination with the environment, which cannot be explained as detection because detection systems only register the presence or absence of an environmental aspect. In section 3.3.1, I introduce a new type of detection to explain A&P's evidence. I also argue that their evidence shows that the boundaries between the input and output phases of behavioural control should be understood as more fluid. This solves the Tracking Problem. In section 3.3.2, I argue that on-going coordination continues to play an important role in more complex behavioural control to ensure that information remains up-to-date.

The Evolution of Decoupled Representation
3.3.1 Solving the Tracking Problem

Detection Loops

In Chapter 1, I introduced the notion of detection cascades, which were processes that produced complex behaviours through coordination between the organism and the environment. In the case of detection cascades, each cue triggers a new behaviour, but in the case of the green alga, cues often trigger the same behaviour. I call these processes **detection loops**. I introduce the notion of detection loops to explain the kind of on-going detection characteristic of green algae's phototaxis, which does not fit the description of either detection or detection cascades.

Cricket phonotaxis is another example of a detection loop. Female crickets (*Acheta domestica*) navigate toward the location of male crickets using the strength of the males' mating songs (Webb 1994). They register this song through their right and left ears and the song activates either of two interneurons, which are connected one-to-one to these ears. If the right interneuron is activated then the song is coming from the right, so the female cricket travels in that direction. If the left interneuron is activated, she travels in that direction. Phonotaxis is an example of a feedback loop (as is phototaxis): as the cricket moves closer to her mate, she not only gets closer to the goal state of reaching her mate, but her mate's song becomes stronger and clearer, making it easier for her to locate him.

Objections

A&P may object that basic tracking requires coordination through multiple channels because error signals are constituted by the signal received and a target value. Therefore, detection could not explain this behaviour because detection systems are sensitive to environmental variation through only a single channel. This objection can be countered on the grounds that the target value is biologically determined, and is not a channel because, again, channels are environmental aspects—light, in the case of the green alga.

Revision

A similar objection is that much of this kind of on-going coordination requires sensitivity to different cues. For example, although the direction in which the cricket flies is controlled by one cue—the male’s song—other cues are required so that the cricket does not, for example, fly into an obstacle. This objection can also be countered on the grounds that multiple detection systems can operate simultaneously. Barrett (2011) argues that much of the complex behaviour of simple organisms emerges out of the on-going and simultaneous coordination of multiple simple systems with the environment. It is possible that a given behavioural response emerges from the activity of multiple detection systems. In the case of the cricket, there may be one detection system directed toward more general locomotion and another detection system directed toward the specialised task of navigating toward mating songs. I call this **parallel detection**.

A&P may also argue that the HT still cannot explain basic behavioural coordination because even basic organisms ‘*have complex sensory connections with the environment*’ (Akins & Pollon in preparation, p. 34, authors’ emphasis). For example, in the case of the green alga, the environmental aspect Light is used to control both the searching phase and homing phase of its phototaxis. However, I do not take this to be problematic for the HT because Sterelny only claims that *detection systems* have one-to-one connections between environmental aspects and behaviours; he does not claim that *basic organisms* have one-to-one connections between environmental aspects and behaviours. It seems unlikely that a given channel can be used in controlling only a single behaviour. This is especially true in the case of very basic organisms, which have the capacity to track the environment through only a single channel.

It may now seem as though I am contradicting my earlier claim that information reuse is rare, which I made in my argument against Christensen in Chapter 2. However, I do not consider multiple detection systems controlled by the same environmental aspect to be an instance of information reuse. This is because I take information reuse to imply that the information was used once to control one response, and then used again later to control

The Evolution of Decoupled Representation

another. Detection systems, however, use information immediately and then do not retain that information. Therefore, I take it that when multiple detection systems are controlled through a single channel, they are not ‘reusing’ that information.

The Fluidity of Input and Output

Sterelny (1997) argues that Godfrey-Smith should understand the boundaries between biological entities and the environment as indistinct. The boundaries between input and output in the process of behavioural control should be understood as similarly fluid in some cases. While Sterelny claims that organisms take action to gather information about their environments (2003, p. 22), A&P show that there is not always a clear information-gathering stage and acting-upon-information-gathered stage. When a detection cascade or detection loop is taken as a whole, the boundaries between input and output become fluid for two reasons. The first is that input and output are extended and so occur simultaneously. The second is that the behaviour produced by the cascade or loop requires the coordination of input and output.

In the case of cricket phonotaxis, the female cricket’s locomotion toward her male mate is temporally extended, so input and output occur simultaneously. Furthermore, output enables the reception of stronger input. In the case of the cricket, much like the case of the alga, it is unhelpful to break the process into an input and an output phase. Similarly, the green alga’s entire searching and homing processes can be understood as on-going input and output that, together, produce its behaviour. By introducing detection loops, and by arguing that the boundaries between input and output are fluid, I have solved the Tracking Problem.

3.3.2 On-going Coordination in Hostile Environments

On-going coordination plays an important role in basic behavioural control, but hostility increases the cost of action, so it might be expected that on-going coordination is rare in hostile environments. However, the behaviour of the piping plover illustrates that on-going

coordination also plays an important role in more complex systems of behavioural control, and in hostile environments, to maintain information.

The piping plover's broken wing display is a response that is extended over time, and as the piping plover produces the display it continues to receive input: piping plovers monitor the responses of intruders, and typically stop displaying if the intruder is not attending to their display (Ristau 1991). This on-going coordination with the environment demonstrates that the display should not be segmented into an input and output phase. Much like the cricket and the alga, the piping plover's tracking of the environmental state and its response to that state is nonlinear.

The process of on-going coordination begins before the display is produced. Ristau presents a hierarchy of behaviours ranging from the least aroused (the piping plover does not move at all) to the most aroused (the broken wing display). I propose that this hierarchy of arousal can be understood in terms of either exploratory behaviour or preparation. According to the hierarchy, the actions that the piping plover can perform while remaining on the nest include: changing its orientation to face the intruder; craning its neck; and standing up but remaining on the nest. All these appear to be actions that enable the piping plover to better monitor the behaviour and position of the intruder. According to the hierarchy, the actions that the piping plover can perform after getting off the nest include: staying closer to, or moving further from, the nest; staying off the nest for a short time or a long time; and performing the broken wing display. Moving from the nest and staying off the nest may similarly be examples of monitoring.

The broken wing display is the culmination of coordination that begins before the display itself. Piping plovers do not begin the display from their nest. Consequently, standing up is a necessary step to take to perform the display, as is moving off, and staying off, the nest. Many of the actions in Ristau's hierarchy are steps that bring the piping plover into a position in which it can produce the display. This is similar to a detection cascade or a detection loop. Before the display is produced, there is an extended process of preparatory

behaviour and monitoring, and the monitoring continues even as the broken wing display is being performed.

On-going coordination is not always a poor strategy in hostile worlds. Godfrey-Smith claims that correspondence between inner states and environmental states is a fuel for success (1996, ch. 6); maintaining these inner states becomes more important in hostile environments because errors are costly. The piping plover case illustrates that action is taken wherever possible to ensure that information remains up-to-date.

3.3.3 Summary

Above, I have introduced detection loops, and argued that the boundaries between input and output are fluid, which solves the Tracking Problem. A&P demonstrate the importance of on-going coordination in basic behavioural control. This on-going coordination remains important in more complex forms of behavioural control as it ensures that an organism has up-to-date information about its environment. Next, I pull the themes of my solutions to the Inference Problem* and Tracking Problem together, and provide a sketch of the continuity between decoupled representation and simpler systems of behavioural control. In doing so I show that Sterelny's account of the evolution of decoupled representation is not flawed in the way that A&P claim.

3.4 Continuity

In this section, I explore some of the details of the continuity between decoupled representation and simpler systems of behavioural control. Expanding upon the details of this continuity is necessary to argue against A&P's claim that Sterelny gets 'the story of decoupled representations *exactly* backwards' (in preparation, p. 39). Furthermore, I take my analysis of robust tracking and coupled representation to suggest that robust tracking is the origin of belief-like representations, not social maps or resource maps. In section 3.4.1, I

sketch an account of the continuity from robust tracking to decoupled representation. In section 3.4.2, I argue that the evolution of decoupled representation is the evolution of the capacity to represent the environment with greater precision and accuracy, and defend Sterelny's story of decoupled representation.

3.4.1 From Robust Tracking to Decoupled Representation

The Origins: Robust Tracking

Recall that multi-conditionalisation refers to behaviour that is controlled by multiple information sources. Christensen illustrates multi-conditionalisation with the example of a response that is only produced after registering the cues Red Light and White Room. I take robust tracking to be a form of multi-conditionalisation because it involves controlling a behavioural response through sensitivity to multiple environmental aspects. For example, the reed warbler may produce the response Reject Egg after registering the cues Large Egg and Cuckoo Nearby.

Despite my disagreements with Christensen, I take his intuitions about multi-conditionalisation and its role in the evolution of decoupled representation to be roughly correct. I take robust tracking, which is a form of multi-conditionalisation, to be the origin of decoupled representation. This is because robust tracking involves coupled representations, which can be involved in minimally off-line processing. These coupled representations can be seen as the origins of decoupled representation because they are less on-line than pp-representations: they do not immediately trigger action, and they can be used in the absence of either input or output. Although they cannot be used outside of a specific problem domain, they may control exploratory behaviours, such as probing, to gather more information about a problem. In Chapter 1, I stated that decoupled representation is selected for in environments in which it is unclear both when information could be used and how new information could modify old information. This lack of clarity partially describes the problem that robust

tracking systems solve. When information is imperfect it must be represented minimally off-line, so that it can be integrated with other imperfect information before it is acted upon.

An Intermediary Stage

Christensen claims that multi-conditionalisation is likely to lead to decoupled representation when it becomes 'efficient to reuse' representations (2010 p. 386). Clark also states that when information is relevant to 'multiple or open-ended types of activity' it is more efficient to represent this information once, but to do so in a decoupled way, so that it may be reused (1997, p. 152). Although Christensen's intuition is correct, representation reuse is unlikely to be efficient in many cases because most organisms do not have the large behavioural repertoires required to use information flexibly. In most of the cases in which they do, the scope of reuse is likely to be limited. For example, the scrub jay may use the information about the location and state of food both when storing food and retrieving food. Similarly, its Navigate To response is produced when the states Worm and <48 Hours obtain. Its Ignore response is also produced in response to the state Worm, except when the state >48 Hours obtains instead of <48 Hours. Although there is representation reuse in this case, it is only limited.

It is unclear whether Sterelny would consider the primate social maps and scrub jay resource maps I described earlier to be fully decoupled representations. Sterelny (2003, p. 41, pp. 52-53) claims that they are possible *origins* of decoupled representations. More recently, however Sterelny has stated that

In (Sterelny 2003), I suggested paradigmatically representational control states tracked relevant environmental contingencies through a number of informational channels and directed responses flexibly. Flexible tracking and response breadth are matters of degree... I suggested that an animal's cognitive map of its territory was a plausible example of this kind of multi-tracked, broad response control system. (Sterelny forthcoming, p. 4)

This could suggest either: (a) that Sterelny now takes social and resource maps to be instances of fully decoupled representations; or (b) that Sterelny still takes these cognitive maps to be decoupled to a degree, but not fully so. Either way, it is clear that some representations are not fully decoupled but can still be used flexibly. I call these **pluri-coupled representations**. A representation is pluri-coupled if it: (pcr1) can be used to control multiple behaviours; and (pcr2) can be involved in off-line processing. Pluri-coupled representations are more primitive than decoupled representations, but less primitive than coupled representations, so they are an intermediary stage of decoupling that is likely to be more common than decoupled representation.

When organisms need to solve a few closely related problems, such as storing food and retrieving food, information that is used in one task is likely to remain relevant in the other. It is also probably important that taking action to update this information is costly. For example, the scrub jay does not simply fly to each cache to test the state of the food because of the energy costs involved. Instead, it must know the location and state of the food before it takes action to retrieve it.

The Final Decoupling

Sterelny states that decoupled representation is selected for in environments where it is unclear when and how information could be used to either direct a response or modify existing information (2003, p. 81). In addition to a high cost of taking action to gather information, this lack of clarity is created by widespread novel problems and the need to maintain information.

A lack of clarity about when and how information can be used is created when novel problems are widespread. Environments that are highly variable are likely to present novel problems to the organisms that live in those environments. One strategy for solving novel problems is trial and error. However, hostility imposes ‘costs on... action, making it

impossible or expensive for an animal to gather relevant information, or position itself to take advantage of unambiguous cues' (2003, p. 26, author's emphasis). Relying on environmental feedback can be costly or dangerous in hostile environments, so novel problems must be solved by making inferences based on existing information. If this information is represented in a problem-specific way it will be of little use in these novel problems, so widespread novel problems and a high cost of action selects for decoupled representation.

A lack of clarity about when and how information can be used is also created by the need to maintain information. Up-to-date information is useful information (Godfrey-Smith 1996). As I have shown above, on-going coordination is an important part of cognition, ensuring that representations contain up-to-date information about the environment. In environments where this kind of action cannot be taken, however, representations must be kept up-to-date through inference. If information is stored in a problem-specific way then it cannot be used to update information globally or to update information related to other problems. Again, keeping information updated through inference when information is expensive selects for decoupled representation.

Decoupled representation is more than just very large response breadth; an organism can have broad-banded response to given information without being able to use that information in a truly flexible and decoupled way. Making inferences is necessary when information about new problems cannot be gathered through action, or action cannot be taken to update existing information. Christensen's rats perform inference; reed warblers seem to do the same. However, in both these cases the inference is based on information that is related only to those problems, so the information used *can* be stored in a problem-specific or action-specific way. It does not have to be decoupled. For inferences to be made about novel problems, however, the information used *cannot* be represented in a problem-specific or action-specific way. I explain why below.

Revision
3.4.2 Precision and Accuracy

Response to A&P

According to A&P, Sterelny's account of the evolution of decoupled representation is inaccurate (in preparation, p. 39). However, Sterelny's and A&P's accounts of the evolution of decoupled representation are not dissimilar. For example, A&P claim that

...we start with simple systems... evolution then moves to general systems that allow us to make fine grained and ordered distinctions over a wide variety of conditions, including novel ones. (*ibid.*)

This claim appears consistent with the connection between response breadth and environmental partitioning in Sterelny's account, as well as what I have said about the importance of novel problems in the evolution of decoupled representation. Furthermore, this claim appears consistent with what Sterelny states elsewhere:

On that standard line, thoughts do not just have representational properties: they have determinate, fine-grained representational properties; whatever the limits on a frog, I can be determinately thinking about flies, rather than approximately fly-shaped packages of food. But it is probably not true that frogs have been selectively tuned to strike at flies, rather than fly-shaped food parcels. (Sterelny forthcoming, p. 5)

A&P state that 'the evolutionary story is not from general to specific, but from the specific to the general' (in preparation, p. 39). Although A&P claim that Sterelny gets the story of the evolution of decoupled representation backwards, his story is also from the specific to the general. Detection systems are directed towards *specific* ends (because they control particular behaviours) and decoupled representations are directed toward *general* ends (because they do not control particular behaviours). However, to enable information about the environment to be used generally, organisms must evolve the capacity to represent the

environment in a determinate and fine-grained manner—they must represent the environment precisely and accurately.

Representing the Environment Precisely and Accurately

When systems are directed toward particular ends, representing the environment accurately is less important. It may even be advantageous for systems to have high error rates if the errors are cheap (Sterelny 2003, p. 32). However, as a representation ‘ceases to be tightly coupled to a specific behaviour, there ceases to be reason to protect against false positive at the expense of false negatives or vice versa’ (*ibid.*). Furthermore, when information about a given environmental state is used only in a specific task, much of the information about that state is unnecessary, and can be omitted, or represented approximately. For example, bees navigate by the sun but appear to have very imprecise information about it (see Chapter 1). Similarly, the rats that solve the Morris water maze task probably have only approximate representations of the landmarks—their general shape and position—because this is all that is required to navigate by them.

Approximate and problem-specific information does not generalise well, so is less likely to be used flexibly (Clark 1997, p. 150). In the case of the Morris water maze, for example, information about the location of one landmark relative to another cannot be used outside of the context of *this particular* water maze. When information is used across a range of tasks, however, it needs to be precise and accurate. If landmark X is represented not only as large and in the corner, but, more precisely and accurately, as something solid, weight supporting, and with draws for files—as a large filing cabinet—then there are more possible tasks in which this representation could be relevant.

Precision and accuracy is exemplified by symbol systems such as language. Decoupled representations, which are minimal beliefs, *may* only become fully belief-like—action-neutral, precise, truth evaluable, and so on—with the emergence of symbol systems

Revision

(Menary 2007; see also Clark 2006). Unfortunately, the transition from minimal beliefs to beliefs is too complex to be considered here in any detail; however, if the continuity account developed here is correct, it may be able to explain cognitive evolution from the simplest systems to the most complex. For example, if detection systems are understood as representational then my continuity account can be understood as beginning with detection systems. This is because detection involves representations that are so imprecise that they carry both descriptive and directive content.

Figure 3.1 illustrates the continuity from detection to belief. As behavioural control moves along the spectrum: (1) there is less on-going coordination with the environment to keep information up-to-date; (2) information can be used more flexibly; because (3) information is represented with greater precision and accuracy.

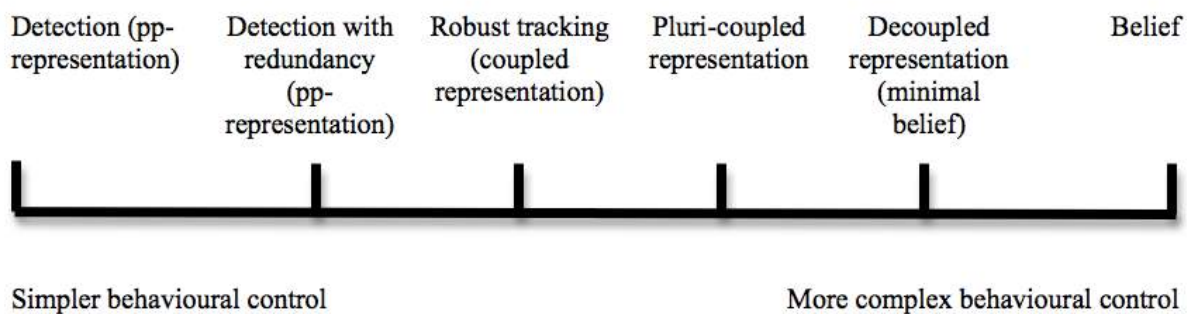


Figure 3.1: The continuity from detection to belief.

3.4.3 Summary

Above, I have sketched a story of continuity from detection to decoupled representation. I take robust tracking to be the origin of decoupled representation because it involves the integration of information across multiple experiences. When environments are highly complex they present novel problems, and in environments that are both complex and hostile, trial and error cannot be used to solve these problems. Instead, they must be solved through inference. This kind of inference requires decoupled representation because action-specific representations do not generalise well beyond their normal problem domain. This is because

they often contain partial or approximate information. Decoupled representations must be precise and accurate so they can be used flexibly. Precision is exemplified by symbol systems like language, so they probably play a crucial role in the transition from minimal beliefs to beliefs.

3.5 Conclusion

With the introduction of the notion of coupled representation, the HT can explain Christensen's evidence and overcome the Inference Problem*. With the introduction of detection loops, and by understanding the boundaries between input and output as fluid, the HT can explain A&P's evidence and overcome the Tracking Problem.

On-going coordination plays a role not only in basic behavioural control, but also in more complex forms of behavioural control, such as robust tracking. Robust tracking systems are the first steps toward decoupled representation. This is because they integrate information across experiences, which requires representing information in a less action-specific manner. When environments present many novel problems, and have a high cost of action, inference becomes a favourable strategy. Information that is represented precisely and accurately generalises well, so is better suited for use in novel problems. Therefore, the evolution of decoupled representation can be understood as the evolution of the capacity to represent the environmental state precisely and accurately.

Conclusion

Conclusion

The ECT and HT provide complementary accounts of the evolution of cognition. The ECT covers the *why*: cognition evolves to coordinate flexible behaviour in response to relevant environmental variation. The HT covers the *how*: cognition coordinates behaviour differently in different environments, such as tracking environmental variation through multiple channels when information is imperfect and using information to control multiple behaviours when information is expensive.

Christensen and A&P argue that the HT gets the *how* completely wrong, but *I have shown that it requires only moderate revision* to be able to explain the evidence that they present. *I have made these revisions* and have done so in a way that remains as faithful to Sterelny's original account as possible.

After introducing detection loops, I have shown that the boundaries between input and output in behavioural control are often fluid. Although I have avoided non-representational accounts of cognition (such as Keijzer 2001; Chemero 2009; Hutto & Myin 2013) because they are beyond the scope of this thesis, the fluidity of behavioural control supports these approaches. It may be possible to use the revised HT to reassess the scope of non-representational cognition in hostile environments. Alternatively, the ECT and HT could be examined in the context of the predictive coding approach to cognition in which interest has been growing (Hohwy 2013; Clark 2013).

After revising the role of representation in robust tracking I have also expanded upon the details about the continuity of the evolution of decoupled representation. I have sketched out the origins of decoupled representation as well as an intermediary stage and have shown that representations control multiple behaviours best when they are precise and accurate. I have argued that decoupled representations are only required in environments in which it is efficient to reuse information. However, because the accuracy of information is so critical in hostile environments, it is only efficient to reuse information in environments where the cost

The Evolution of Decoupled Representation

of action is very high and organisms are frequently presented with novel problems. The details of my continuity story could be expanded, focusing on differing precision in very simple representations and belief-like, propositional representations. It is possible that investigating this further could lead to a solution to the Methodological Problem, which remains unsolved.

References

- Akins, K, & Pollon, S in preparation, 'Three tales of pre-intentionality: green algae, water cress and water fleas, *Philosophy of Biology at Madison*, available from: <https://docs.google.com/a/students.mq.edu.au/file/d/0B3VUPVK1NQ9hc3REYkJraGZwWjQ/edit>, accessed 26 August 2014²⁸.
- Allen, C 2006, 'Transitive inference in animals: Reasoning or conditioned associations', in S Hurley and M Nudds (eds), *Rational Animals?*, Oxford University Press, Oxford, pp. 175-185.
- Avital, E, & Jablonka, E 2000, *Animal Traditions: Behavioural Inheritance in Evolution*, Cambridge University Press, Cambridge.
- Barrett, L 2011, *Beyond the Brain: How Body and Environment Shape Animal and Human Minds*, Princeton University Press, Princeton.
- Byrne, RW 1997, 'The technical intelligence hypothesis: An additional evolutionary stimulus to intelligence?', in A Whiten and RW Byrne (eds.), *Machiavellian Intelligence II: Extensions and Evaluations*, Cambridge University Press, Cambridge, pp. 289-311.
- Byrne, RW, & Whiten, A 1988, *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes and Humans*, Oxford University Press, Oxford.
- Camhi, JM, Tom, W, Volman, S 1978, 'The escape behavior of the cockroach *Periplaneta Americana*', *Journal of Comparative Physiology*, vol. 128, pp. 203-212.
- Chemero, A 2009, *Radical Embodied Cognitive Science*, MIT Press, Cambridge, MA.
- Christensen, W 2010, 'The decoupled representation theory of the evolution of cognition – a critical assessment', *The British Journal for the Philosophy of Science*, vol. 61, no. 2, pp. 361-405.
- Clark, A 1997, *Being There: Putting Brain, Body, and World Together Again*, MIT Press, Cambridge, MA.
- Clark, A 2006, 'Material Symbols', *Philosophical Psychology*, vol. 19, no. 3, pp. 291-307.
- Clark, A 2013, 'Whatever next? Predictive brains, situated agents, and the future of cognitive science', *Behavioral and Brain Sciences*, vol. 36, no. 3, pp. 181-204.
- Comer, CM, Mara, E, Murphy, KA, Getman, M, Mungy, MC 1994, 'Multisensory control of escape in the cockroach *Periplaneta americana*', *Journal of Comparative Physiology*, vol. 174, pp. 13-26.
- Dawkins, R 1986, *The Blind Watchmaker*, Longman Scientific & Technical, Harlow.

²⁸ Permission acquired from Kathleen Akins to quote from this manuscript on 28 August 2014.

The Evolution of Decoupled Representation

- De Lillo, C, Floreano, D & Antinucci, F 2001, 'Transitive choices by a simple, fully connected, backpropagation neural network: Implications for the comparative study of transitive inference', *Animal Cognition*, vol. 4, no. 1, pp. 61-68.
- Dennett, D 1989, *The Intentional Stance*, MIT Press, Cambridge, MA.
- Dennett, D 1995, *Darwin's Dangerous Idea: Evolution and the Meanings of Life*, Touchstone, New York.
- Dretske, FI 1981, *Knowledge and the Flow of Information*, Basic Blackwell, Oxford.
- Dretske, FI, 1988, *Explaining Behaviour: Reasons in a World of Causes*, MIT Press, Cambridge, MA.
- Dusek, JA, & Eichenbaum, H 1997, 'The hippocampus and memory for orderly stimulus relations', *Proceedings of the National Academy of Science of the United States of America*, vol. 94, pp. 7109-7114.
- Eichenbaum, H 2000, 'A cortical-hippocampal system for declarative memory', *Nature Reviews Neuroscience*, vol. 1, pp. 41-50.
- Godfrey-Smith, P 1996, *Complexity and the Function of Mind in Nature*, Cambridge University Press, Cambridge, MA.
- Godfrey-Smith, P 2002a, 'Environmental complexity and the evolution of cognition', in R Sternberg and J Kaufman (eds.), *The Evolution of Intelligence*, Lawrence Erlbaum, Mahwah, pp. 233-249.
- Godfrey-Smith, P 2002b, 'On the evolution of representational and interpretive capacities', *The Monist*, vol. 85, no. 1, pp. 50-69.
- Gould, JL, & Gould, CG 1988, *The Honey Bee*, WH Freeman, New York.
- Heider, F, & Simmel, M 1944, 'The experimental study of apparent behavior', *The American Journal of Psychology*, vol. 57, no. 2, pp. 243-259.
- Hohwy, J 2013, *The Predictive Mind*, Oxford University Press, Oxford.
- Hutto, D, & Myin, E 2013, *Radicalizing Enactivism*, MIT Press, Cambridge, MA.
- Jacobson, SL 1965, 'Receptor response in Venus's fly-trap', *The Journal of General Physiology*, vol. 49, pp. 117-129.
- Keijzer, FA 2001, *Representation and Behavior*, MIT Press, Cambridge MA.
- Llewellyn, J, Webb, JK, Schwarzkopf, L, Alford, R, & Shine, R 2010, 'Behavioural responses of carnivorous marsupials (*Planigale maculata*) to toxic invasive cane toads (*Bufo marinus*)', *Australian Ecology*, vol. 35, no. 5, pp. 560-567.
- Levitis, DA, Lidicker, Jr., WZ, Freund, G 2009, 'Behavioural biologists do not agree on what constitutes behaviour', *Animal Behaviour*, vol. 78, pp. 103-110.

References

- Marshall, R 2014, 'Philosophy of biology: Peter Godfrey-Smith interviewed by Richard Marshall', *3:AM Magazine*, <<http://www.3ammagazine.com/3am/philosophy-of-biology/>>, accessed 29 April 2014, .
- Menary, R 2007, *Integrated Cognition: Mind and Cognition Unbounded*, Palgrave Macmillan.
- Millikan, RG 1984, *Language, Thought, and Other Biological Categories*, MIT Press, Cambridge, MA.
- Millikan, RG 1989, 'Biosemantics', reprinted in Millikan (1993), pp. 83-102.
- Millikan, RG 1993, *White Queen Psychology and Other Essays for Alice*, MIT Press, Cambridge, MA.
- Millikan, RG 1995, 'Pushmi-Pullyu Representations', *Philosophical Perspectives*, vol. 9, no. 1, pp. 185-200.
- Ono, M, Terabe, H, Hori, H, & Sasaki, M 2003, 'Components of giant hornet alarm pheromone', *Nature*, vol. 424, pp. 637-638.
- Plummer, MR, Camhi, JM 1981, 'Discrimination of sensory signals from noise in the escape system of the cockroach: The role of wind acceleration', *Journal of Comparative Physiology*, vol. 142, pp. 347-357.
- Premack, D 2007, 'Human and animal cognition: Continuity and discontinuity', *Proceedings of the National Academy of Sciences*, vol. 104, no. 35, pp. 13861-13867.
- Ramsey, WM 2007, *Representation Reconsidered*, Cambridge University Press, Cambridge.
- Reddien, PW, & Alvarado, AS 2004, 'Fundamentals of planarian regeneration', *Annual Review of Cell and Developmental Biology*, vol. 20, pp. 725-757.
- Ristau, C 1991, 'Aspects of the cognitive ethology of an injury-feigning bird, the piping plover', in C Ristau (ed.), *Cognitive Ethology: The Minds of Other Animals*, Lawrence Erlbaum Associates, Hillsdale.
- Shannon, CE 2001 [1948], 'A mathematical theory of communication', *Mobile Computing and Communications Review*, vol. 5, no. 1, pp. 3-55.
- Shettleworth, SJ 1998, *Cognition, Evolution, and Behavior*, Oxford University Press, Oxford.
- Skyrms, B 2010, *Signals: Evolution, Learning, and Information*, Oxford University Press, Oxford.
- Sterelny, K 1995, 'Basic minds', reprinted in Sterelny (2001), pp. 198-220.
- Sterelny, K 1996a, 'Explanatory pluralism in evolutionary biology', reprinted in Sterelny (2001), pp. 129-151.
- Sterelny, K 1996b, 'The return of the group', reprinted in Sterelny (2001), pp. 81-106.
- Sterelny, K 1997, 'Where does thinking come from? A commentary on Peter Godfrey-Smith's *Complexity and the Function of Mind in Nature*', reprinted in Sterelny (2001), pp. 181-197.
- Sterelny, K 1999, 'Situated agency and the descent of desire', reprinted in Sterelny (2001), pp. 241-259.
- Sterelny, K 2000, 'Roboroach, or, the extended phenotype meets cognitive science', *Philosophy and Phenomological Research*, vol. 61, no. 1, pp. 207-215.
- Sterelny, K 2001, *The Evolution of Agency and Other Essays*, Cambridge University Press, Cambridge.
- Sterelny, K 2003, *Thought in a Hostile World*, Blackwell Publishing, Malden.

The Evolution of Decoupled Representation

- Sterelny, K 2004, 'Reply to Papineau and Stich', *Australasian Journal of Philosophy*, vol. 82, no. 3, pp. 512-522.
- Sterelny, K 2006, 'Folk logic and animal rationality', in S Hurley and M Nudds (eds), *Rational Animals?*, Oxford University Press, Oxford, pp. 293-311.
- Sterelny, K forthcoming, 'Content, control and display: the natural origins of content', *Philosophia*.
- Stotz, K 2010, 'Human nature and cognitive-developmental niche construction', *Phenomenology and the Cognitive Sciences*, vol. 9, no. 4, pp. 483-501.
- Swoyer, C 1991, 'Structural representation and surrogate reasoning', *Synthese*, vol. 87, no. 3, pp. 449-508.
- Tomasello, M 2000, 'Two hypotheses about primate cognition', in C Heyes and L Huber (eds.), *Evolution of Cognition*, MIT Press, Cambridge, MA, pp. 165-184.
- Van Duijn, M, Keijzer, F, & Franken, D 2006, 'Principles of minimal cognition: Casting cognition as sensorimotor coordination', *Adaptive Behavior*, vol. 14, no. 2, pp. 157-170.
- Vigo, R, & Allen, C 2009, 'How to reason without words: inference as categorization', *Cognitive Processing*, vol. 10, no. 1, pp. 77-88.
- Webb, B 1994, 'Robotic experiments in cricket phonotaxis'. In D Cliff, P Husbands, JA Meyer, and SW Wilson (eds.), *From Animals to Animats 3: Proceedings of the Third International Conference on Simulation of Adaptive Behavior*, pp. 45-54.
- Webb, JK, Brown, GP, Child, T, Greenlees, MJ, Phillips, BL, & Shine, R 2008, 'A native dasyurid predator (common planigale, *Planigale maculata*) rapidly learns to avoid a toxic invader', *Australian Ecology*, vol. 33, no. 7, pp. 821-827.
- Wilson, M 2002, 'Six views on embodied cognition', *Psychonomic Bulletin and Review*, vol. 9, no. 4, pp. 625-636.
- Zappettini, S & Allen, C 2013, 'Does evidence from ethology support bicoded cognitive maps?', *Behavioral and Brain Sciences*, vol. 36, no. 5, pp. 570-571.
- Zentall, TR 2001, 'The case for cognitive approach to animal learning and behaviour', *Behavioural Processes*, vol. 54, no. 1-3, pp. 65-78.

Appendix: List of Problems

Inference Problem

Inference requires decoupled representation and animals outside of primate social environments can perform inference. Therefore, decoupled representation evolves earlier and is more widespread than Sterelny claims.

Inference Problem*

The performance of non-human animals on inference tasks, such as the transitive inference and Morris water maze tasks, does not meet the conditions of decoupled representation, but neither does it meet the conditions of robust tracking systems if robust tracking involves pp-representation. Therefore, the HT cannot explain the performance of non-human animals on inference tasks.

Methodological Problem

Response breadth is typically restricted in most experimental methodologies, making it difficult to test for (dr1). It is often unclear what information an animal uses to control its behaviour, making it difficult to test for (dr2).

Multi-conditionalisation Problem

Multi-conditionalisation is likely to have evolved early and be widespread. Because multi-conditionalisation leads to decoupled representation, decoupled representation is likely to evolve earlier than primate social environments and be more widespread than Sterelny claims.

Multi-conditionisation Problem*

Multi-conditionalisation is likely to have evolved *very early* and be *very widespread*. Because multi-conditionalisation leads to decoupled representation, decoupled representation is likely to evolve *much* earlier than primate social environments and be *much* more widespread than Sterelny claims.

Tracking Problem

Basic behavioural control requires on-going coordination with the environment, and cannot be explained as detection. Therefore, the HT cannot explain a great deal of basic behavioural control.