

Chapter 8

Rational or associative? Imitation in Japanese quail

David Papineau and Cecilia Heyes

Abstract

Much contemporary psychology assumes a fundamental distinction between associative explanations of animal behaviour, in term of unthinking 'conditioned responses', and rational explanations, which credit animals with relevant 'knowledge' or 'understanding' or 'concepts'. This paper argues that this dichotomy is both unclear and methodologically unhelpful, serving only to distract attention from serious questions about which cognitive abilities are present in which animals. We illustrate the issues by considering recent experimental work on imitation in Japanese quail.

8.1 Introduction

According to René Descartes, a wide gulf separates animals from humans:

'... [A]fter the error of those who deny the existence of God... there is none that is more powerful in leading feeble minds astray from the straight paths of virtue than the supposition that the soul of the brutes is of the same nature with our own' (Descartes 1637)

For Descartes, the gulf between animals and humans had a metaphysical basis. The universe was composed of two fundamentally different substances, mind and matter, and animals lacked minds. Because of this, animals were incapable of rational thought and action. Where human activities were governed by the operations of the soul acting through the pineal gland, animals were mere machines, automata governed by nothing but the laws of physics.

When we ask 'Are animals rational?' the question can no longer have anything like the same significance it had for Descartes. Contemporary psychology is premised on materialism. Humans, no less than animals, are complex machines, whose operations depend on scientifically describable psychological and neurological processes. If there is still a live issue as to whether animals are rational, it cannot depend on whether or not animals are blessed with the same special mind-stuff as humans.

Perhaps nowadays we understand the question differently, in terms of two styles of cognitive explanation. This certainly seems to be the way that many psychologists understand the issue. On the one hand lie 'rational' explanations of behaviour, explanations that advert to norm-governed reasoning involving belief-like representations. On the other side lie non-rational explanations, in terms of 'behaviourist' or (more accurately) associative psychological processes.

Ideally, this contrast would give rise to fruitful experimental work. Given some potentially 'rational' behaviour, we try to get the alternative norm-governed and associative accounts to make contrasting predictions about how the behaviour would change under various experimental manipulations, and then test these predictions empirically.

However, we are doubtful that this methodological strategy is as cogent as it initially seems.¹ Of course, we are all in favour of heuristic attitudes that lead to the experimental exploration of animal behaviours and the processes that mediate them. But we doubt that a simple dichotomy between 'rational' and 'associative' is the best way to motivate such experimental work. We suspect that the rational–associative dichotomy is just Descartes dressed up in modern garb. In place of Descartes' immaterial mind we have the accolade of 'folk psychology', and in place of his brute matter we have 'associative machines'. Psychologists are asked to decide whether: (a) animals 'know about', or 'understand', or 'have a concept of' such-and-such (causality, space and time, others' minds); or whether (b) they are driven by unthinking associative mechanisms.

This way of setting things up may only obscure the real issues. We know from basic Darwinian principles that there are no unbridgeable gulfs in cognition: each cognitive innovation must proceed by adding to, or tinkering with, pre-existing cognitive systems. In line with this, we want to know precisely which cognitive mechanisms have evolved in which animals. To be told that certain animals can 'understand' such-and-such—or alternatively that they are simply displaying 'conditioned responses'—is not helpful. What we want to know is how they work, not whether they can cross the Rubicon of folk psychology.

8.2 Imitation in Japanese quail

We would like to back up the remarks above by considering a recent experiment by Akins and Zentall on imitation in Japanese quail.

¹ Editors' note: Consider how these methodological concerns may apply to empirical work described in various chapters in this volume, including those by Allen, Clayton *et al.*, Call, Shettleworth and Sutton, and Addressi and Visalberghi. Compare these concerns about the rational–associative distinction with Povinelli and Vonks' methodological concerns (this volume) about the mind reading–behaviour reading distinction.

Akins and Zentall (1998) allowed naïve observer quail to watch a trained ‘model’ or ‘demonstrator’ quail depressing a lever using either its beak or its feet. Half of the observers saw the demonstrator rewarded with food for this behaviour, while the other half saw the demonstrator responding repeatedly ‘in extinction’, that is without reward.

When the observer birds were subsequently given access to the lever themselves, those that had seen their demonstrator’s behaviour being rewarded tended to imitate the model—to peck the lever if their model had used its beak, and to step on the lever if their model had used its feet. However, when the model’s behaviour was unrewarded, there was no evidence of imitation—the proportion of pecking and stepping responses was the same for observers of beak use and foot use. Thus, the observer quail in this experiment showed sensitivity to demonstrator reward.²

8.3 The significance of sensitivity to demonstrator reward

Akins and Zentall’s result is a good test case for the heuristic utility of the ‘rationality–associative’ dichotomy for two reasons. First, it is easy, perhaps irresistible, to interpret sensitivity to demonstrator reward in folk-psychological terms. We naturally assume that the birds who imitated did so because they *wanted* food and *believed* that performing the same action as the demonstrator would enable them to get it. Second, standard associative learning theory is apparently unable to explain sensitivity to demonstrator reward. This kind of imitation would thus seem to be a good candidate for ‘rational’ behaviour, as opposed to mere associative response.

It will be worth going slowly for a moment in order to explain exactly why sensitivity to demonstrator reward is difficult to explain in standard associative terms. An initial stumbling block is known as the ‘correspondence problem’ (Heyes, in press; Nehaniv and Dautenhahn 2002). Given that a quail sees something very different when he watches another quail pecking a lever (him-pecking) and when he pecks the lever himself (me-pecking), how does the observer know that it is pecking behaviour, rather than his stepping behaviour, which is ‘the same’ as that of the pecking demonstrator? After all, neither me-stepping nor me-pecking looks or feels much like him-pecking. It appears that associative learning theory is too flimsy to deal with this conundrum, and that heavy cognitive machinery—such as ‘supramodal’ mental representation (Meltzoff and Moore 1997) or ‘perspective-taking’ (Bruner 1972; Tomasello *et al.* 1993)—is necessary to solve the correspondence problem.

However, let us suppose that the correspondence problem can somehow be overcome, by an associative mechanism (Heyes 2001) or otherwise. (When the imitating animal is a bird, rather than a person or other ape, the correspondence problem is typically evaded by calling the phenomenon ‘mimicry’, and assuming that it is the product of

² Editors’ note: see also Pepperberg’s use of model-rival training with African grey parrots, this volume.

some innate tendency.) Still, even if we assume that there is some automatic tendency for animals blindly to reproduce the behaviour of conspecifics, this does not immediately clear the way for an associative explanation of sensitivity to demonstrator reward—that is, of a tendency to reproduce behaviour *only* when it is observed to be followed by some reward.

It might seem as if it would just be a short step from blind mimicry to full sensitivity to demonstrator reward. Don't the animals just have to observe what result generally follows the demonstrator's behaviour, and draw a straightforward inference? However, if we think about the basic associative learning mechanisms by which animals might acquire information about general patterns in their environment, we can see that the inference in question is not straightforward at all.

Suppose, for the sake of the argument, that the basic associative learning mechanisms are classical and instrumental conditioning. Classical conditioning allows animals to take one stimulus, F, as a sign of another, G, whenever Fs have previously been followed by Gs in the animals' environment. Instrumental conditioning allows animals to perform some behaviour B in pursuit of reward R, whenever Bs have previously led to Rs in the animals' environment. In combination, these two kinds of mechanisms can lead to a high degree of flexibility in gearing behaviour to local environmental regularities.

Neither of these mechanisms, however, will give rise to the kind of sensitivity to demonstrator reward found in Japanese quails. It may not be immediately obvious why. But consider things from the perspective of the bird observing the demonstrator. The observer quail sees the demonstrator peck at the lever, say, and subsequently sees the demonstrator being rewarded with food. Note, however, that in this scenario the *observer* hasn't yet performed any behaviour, nor has *it* received any subsequent reward. So there is nothing here for instrumental conditioning to get a grip on, nothing to make the *observer* peck in pursuit of food.

True, if the observer quail were prone to blind mimicry, then this alone would dispose it to peck when it observes the demonstrator pecking. But this won't explain why the observer quail only pecks when it sees the demonstrator being rewarded.

What about classical conditioning? Won't this lead the observer to associate the pecking with the reward? Yes—but here the association will be between *observing* the pecking and *observing* the demonstrator receiving the reward, not between *doing* the pecking and *receiving* the reward. This classical association might lead the observer quail to respond to the sight of *another* pecking bird in the way it previously responded to the appearance of food, for instance by approaching. But this too will do nothing to explain why the observer itself learns to peck only when the demonstrator receives the food.

In short, there is nothing obviously rewarding to the observer in seeing some other bird being fed. So, even assuming some tendency to blind mimicry, there seems nothing in standard associative learning theory to explain why the observers'

pecking tendencies should get reinforced specifically when the demonstrator quail gets fed.³

8.4 Rational versus associative explanations of the Japanese quail

According to the simple dichotomy outlined earlier, the alternative to associationism is rationalism. If the behaviour of the Japanese quail cannot be explained in associative terms, then we should account for it in terms of rational cognition. Thus we should credit the observer bird with a desire for food, and a belief about how to get it, and view its behaviour as the rational product of these propositional attitudes. For instance, we might reconstruct its reasoning as follows:

The demonstrator's pecking produced food.

So my pecking will produce food.

I want food.

So I'll peck.

No doubt those who advocate some explanation along these lines will not wish to credit Japanese quail with unlimited cognitive powers. There are unquestionably some constraints on the range of subjects that the quail can think about—they can't think about electricity, for instance, or genetics. In line with this, it might be claimed that their rationality is only displayed in the operation of certain 'domain-specific modules', such as their navigation module and their social reasoning module, along with their imitation module (see and cf. Hurley, this volume, on context-bound reasons for action). Still, despite the restriction of rationality to such specific domains, the idea would remain that, within these domains, the quail are capable of forming contentful beliefs and desires, and of using reason to figure out the consequences thereof.

So, validating the rational–associative distinction, we seem to have found that sensitivity to demonstrator reward is subject to rational but not to associative explanation. But there are two twists in this tale, and here is the first: with a little imagination, old fashioned, stimulus–response (S–R) associative learning theory *can* explain sensitivity to demonstrator reward. Suppose that in the past the observer bird has fed while seeing other birds feeding (quails are highly social animals); this could well have

³ Interestingly, this is a special case of a more general limitation. As well as not being able to absorb means–end information from observation of other animals, creatures limited to standard associative mechanisms won't be able to absorb means–end information from any kind of non-participatory observation, and for the same reason. Observing that cause A produces effect B might make you anticipate B when you see A, via classical conditioning, but it won't get you *doing* A in pursuit of B, when there is nothing rewarding about merely observing B. (Cf. Papineau, in press.)

led to the sight of another bird feeding becoming a 'secondary reinforcer' for the observer, with the result that sight of another feeding itself becomes rewarding. Now put this together with a general tendency for mimicry. This incipient tendency to peck blindly, say, when observing another bird pecking, would then become paired with a rewarding event for the observer, *if* the observer sees the demonstrator feeding. This model would then explain why only the birds that saw the demonstrator being rewarded would acquire a disposition to peck—only they would have been (secondarily) rewarded by the observation of another bird feeding (see and cf. Addessi and Visalberghi, this volume, on socially facilitated feeding in capuchin monkeys).

A variety of experiments would be needed to test this associative hypothesis against the rationality alternative, but here's a simple one that would provide a good start. Consider a variant experiment in which hungry observers see demonstrators stepping and pecking, but where the two demonstrator behaviours provide access to two visually distinct but equally palatable foods. Then, before the observers are subsequently tested, they are allowed to feed to satiety on *one* of these foods. The prefed food is, therefore, devalued. If the S–R associative account is correct, one would not expect this prefeeding to have any impact on the proportion of pecking and stepping responses made by the observer quail on test. This is because the (secondary) reinforcement during observation would merely have forged a simple S–R association between, say, the sight of the lever and the urge to peck it, detached from any expectation of reward. On the other hand, if the rationality hypothesis were correct, and the animals were forming explicit beliefs about the effects of different actions, we would expect the birds to perform less of the response that produced the now devalued food during observation.

This experiment has not yet been performed, but let us suppose that, when it is done, the quail perform less of the response that had produced devalued food. Would this outcome show that quail are rational, and thereby demonstrate the heuristic value of the rational–associative dichotomy? Unfortunately not, because (and here's the second twist in the tale) the 'rational' behaviour could also be explained in associative terms. It would be incompatible with S–R associative learning, but not with associative learning of a relationship between the response and its outcome. On this account, the quail learn by demonstrator observation that, for example, pecking is associated with red food and stepping is associated with green food. However, devaluation of, say, the red food, prevents the experimental context from activating the stepping response by establishing a negative feedback loop between the animal's incentive system and its representation of the red food outcome (see Dickinson's 1994 associative cybernetic model of instrumental action for details).⁴

⁴ No similarly obvious associationist explanations offer themselves for some of the very impressive cognitive feats performed by the scrub jays in Clayton *et al.* (this volume). But this does not undermine the point, stressed in our final section below, that a bald assertion that some piece of cognition is 'rational' rather than 'associative' is no substitute for a specific hypothesis about the mechanisms responsible.

8.5 Morals

At one level, our doubts are simply about the clarity of the distinction between 'rational' and 'associative'. The standard way of drawing this distinction is in terms of *representation*: where rational thought involves the processing of genuine representational states, associative responses rest on nothing but blind mechanical response.⁵ However, it is not clear that this contrast stands up to philosophical scrutiny, especially when we reflect on the more complex kinds of associative hypotheses considered in the last section. After all, animals who have learned by conditioning must be internally modified in certain ways, and one natural way of characterizing these modifications is in terms of their content. For example an animal who has learned a response–outcome association as a result of being rewarded by D after doing B in conditions C is quite naturally described as embodying the information that B in C will lead to D. In this connection, note that it is certainly the biological purpose of such instrumental conditioning to produce such modifications just in those cases when B in C will indeed lead to D. Note also how this causal information will interact with a current perception of C and a drive for D to yield the behaviour B (in a way quite analogous to the familiar practical inference 'I know that B in C will lead to D; I can see I am in circumstances C; I want D; so, I do B').^{6,7}

Perhaps we shouldn't be trying to draw the rational–associative distinction in terms of representation *per se*, but in terms of the kind of processes to which the animals can subject their representations. It is certainly true that associative mechanisms are limited in terms of the kinds inferences that they can perform. Consider an animal who has formed a classical association between the sight of another animal doing B in C and its receiving a reward D. In a sense, this animal embodies the information that

⁵ Sometimes the distinction is also drawn in terms of normativity: only rational processes are guided by rules of correct and incorrect reasoning. But this strikes us as an obscure idea. Surely the first concern of cognitive science is to figure out how animals *do* think, not how they *ought* to think. Moreover, one of us (Papineau 1999) would argue that any normative considerations that do apply to cognitive states are *derivative* from their representational nature (you ought to believe what is true, and you ought to reason so as to preserve truth); in which case distinguishing rational from associative processes in terms of normativity collapses into distinguishing them in terms of representation.

⁶ For a more detailed discussion of representationalist readings of associative mechanisms, see Papineau, in press.

⁷ Dretske (this volume) argues that behaviour arising from associatively learned cognitive representations is minimally rational, since it is guided by the content of those representations, but that this is not true of behaviour guided by hard-wired cognitive structures. We agree that associatively learned behaviour is guided by contentful representations, and to this extent can be counted as 'rational' (though note our caveats about inferential limitations immediately below). However, we are not convinced by Dretske's contention that, if just the same cognitive structures were designed by genetic selection rather than learning, this would somehow yield less rationality.

doing B in C will be followed by D. But this information is embodied in the wrong place to influence behaviour. Since it is embodied in a classical association, rather than in an instrumentally conditioned response–outcome disposition, the information can't combine with the animal's own drive for R and its perception that it is currently in circumstances C to generate behaviour B, as would an instrumentally conditioned response–outcome disposition.⁸

So perhaps we should characterize rational cognition as cognition that is not limited in the inferences it can perform. Note how this was implicit in the rationalist explanation offered above for the Japanese quail's learning abilities: this simply presupposed that the information about a pecking–food association derived from observation could then inform the birds' own choice of action. (The demonstrator's pecking produced food/So my pecking will produce food/I want food//So I'll peck.)

However, if *this* is what rationalist explanation means, then we have substantial doubts whether it is of any serious utility for psychology. To explain behaviour by positing unlimited inferential powers is little short of offering no explanation at all. Such unlimited intellectual powers may be a kind of cognitive ideal, and maybe humans using language (or other artificial aids) come close, but cognition in non-linguistic animals must be realized by specific architectures, and such architectures will inevitably place limitations on the kinds of inferences animals can draw, and consequently on which behaviours they will display.

It runs against every principle of evolutionary thinking to suppose that some animals, in some domains, have magically jumped across the chasm dividing associative mechanisms from ideal cognition. Evolution builds adaptations by tinkering and gradual additions, not via miraculous saltations, and this applies to cognitive evolution as much as elsewhere (Heyes, in press). Many different animals can perform many varieties of flexibly intelligent behaviour. But it is not helpful to try to explain such behaviour by suggesting that these animals have somehow achieved 'rationality', if this means that they are somehow unconstrained in the inferences they can draw from their information. What we need to know is precisely which extra cognitive mechanisms evolution has endowed them with, and which extra bits of reasoning they are thereby able to perform.

Some contemporary students of animal behaviour happily characterize themselves as children of the 'cognitive revolution'. If the cognitive revolution means that other possibilities beyond strict behaviourism should be explored, then it is clearly a good thing. But if it means that psychologists should stop worrying about cognitive

⁸ Information embodied in classical associations can influence behaviour in a different way: if an animal is already disposed to perform behaviour B in condition F in pursuit of D, and then forms a classical association between G and F, this will lead it to do B in condition G too. (We can here think of it as putting together the information that B in F will lead to D, with the classical information that all Gs are Fs, to derive the conclusion that B in G will lead to D.)

mechanisms, then it is arguably a step backwards. There is little point in asking whether animals ‘know about’, or ‘understand’, or ‘have a concept of’ such-and-such, if these questions are not accompanied by an enquiry into what animals can *do* with their information.

So, to sum up, it doesn’t seem as if the rational–associative dichotomy is satisfactory, whichever way we cut it. If we try to define ‘rational’ in terms of representation, then there seems no good reason to rule out associative mechanisms as rational. And if we try to define ‘rational’ in terms of unlimited inferential abilities, then this idealizes rationality to a degree that renders it useless as a research tool.

Maybe, when we fully understand the structure of animal cognition, we will be better placed to decide what kinds of animal achievement are worth dignifying as rational. But by then we probably won’t care.

References

- Akins, C. K., and Zentall, T. R. (1998). Imitation in Japanese quail: The role of reinforcement of demonstrator responding. *Psychonomic Bulletin and Review*, **5**: 694–697.
- Bruner, J. S. (1972). Nature and uses of immaturity. *American Psychologist*, **27**: 687–708.
- Descartes, R. (1637). *A Discourse on Method*, p. 46. Translated by John Veitch (1912). London: J. M. Dent and Sons, Everyman’s Library Edition.
- Dickinson, A. (1994). Instrumental conditioning. In: N. J. Mackintosh, ed. *Animal Learning and Cognition*, pp. 45–79. San Diego: Academic Press.
- Heyes, C. M. (2001). Causes and consequences of imitation. *Trends in Cognitive Sciences*, **5**: 253–261.
- Heyes, C. M. (in press a). Four routes of cognitive evolution. *Psychological Review*.
- Heyes, C. M. (in press b). Imitation by association. In: S. Hurley and N. Chater, eds. *Perspectives on Imitation*. Cambridge, MA: MIT Press.
- Meltzoff, A. N., and Moore, M. K. (1997). Explaining facial imitation: a theoretical approach. *Early Development and Parenting*, **6**: 179–192.
- Nehaniv, C. L., and Dautenhahn, K. (2002). The correspondence problem. In: K. Dautenhahn and C. L. Nehaniv, eds. *Imitation in Animals and Artifacts*, pp. 41–61. Cambridge, MA: MIT Press.
- Papineau, D. (1999). The normativity of judgement. *Aristotelian Society Supplementary Volume*, **73**: 17–43.
- Papineau, D. (in press). Human minds. In: A. O’Hear, ed. *Minds and Persons*. Cambridge University Press.
- Tomasello, M., Kruger, A. C., and Ratner, H. H. (1993). Cultural learning. *Behavioral and Brain Sciences*, **16**: 495–552.

