

Philosophical Transactions of the Royal Society B, 364, 2429-2443.

doi:10.1098/rstb.2009.0061

Theme issue on Evolution, Development and Intentional Control of Imitation

Edited by: Cecilia Heyes, Ludwig Huber, György Gergely & Marcel Brass

Imitation as an Inheritance System

NICHOLAS SHEA

Abstract

What is the evolutionary significance of the various mechanisms of imitation, emulation and social learning found in humans and other animals? This paper presents an advance in the theoretical resources for addressing that question, in the light of which standard approaches from the cultural evolution literature should be refocused. The central question is whether humans have an imitation-based inheritance system – a mechanism that has the evolutionary function of transmitting behavioural phenotypes reliably down the generations. The paper goes on to apply this new perspective to the evidence about the nature of imitation mechanisms, from a variety of disciplines, featured in the theme issue.

Keywords

mechanisms of imitation; evolution of imitation; inheritance system; mirror system; cultural evolution

Nicholas Shea, Faculty of Philosophy and Somerville College, University of Oxford

Address: Philosophy Centre, 10 Merton Street, Oxford, OX1 4JJ, U.K.

Email: nicholas shea <at> philosophy ox ac uk

1. Introduction

What explains the extraordinary richness of human societies? Our culture is much richer than in other species: more artefacts, technological innovation, bodies of knowledge and complex social organisation. Doubtless, special psychological mechanisms are part of the story. For a long time, the capacity to imitate has been given centre stage. Perhaps the richness of human culture is due to an innate and uniquely human imitative ability? It is all very well to hypothesise that a more developed capacity for imitation would give rise to greater cultural complexity. But the empirical picture is by no means so clear. Recent investigations of the mechanisms and functions of imitation, many of which are reported and summarised in this theme issue, put that simple story seriously in doubt.

Although it has been known for a long time that many species are capable of forms of social learning, it is now becoming clear that some do so in surprisingly intelligent ways. One strand of developmental work in humans also brings out the rationality of infant and child imitation. The extent to which children imitate at all depends upon being in an appropriate social or 'natural pedagogical' context. Furthermore, in being selective about which actions they tend to imitate, children appear to show an appreciation of the intentional nature of a model's action and the causal structure of the problem to which it is applied. A second strand seems to pull in the opposite direction, emphasising the extent to which humans engage in 'blind' imitation, copying even irrelevant actions in a way that other species do not. Further results reported here cast doubt on the idea that there is an innate or developmentally canalized mechanism of imitation, the mirror system, endowed with a fixed library of mappings from observed actions to bodily movements. So empirical investigation has undermined the satisfyingly simple hypothesis that the ability to imitate is innate, if significant elements depend upon individual learning, and uniquely human, if other species can imitate bodily movements of conspecifics and do so in rational ways; although elements of the unfolding picture suggest that the way humans imitate does have peculiar features. The evidence could be divided into elements that support and undermine the simple hypothesis, but that would miss the extent to which the data generates new questions of its own. Why is human imitation intelligent in some ways and obtuse in others? Given that other species also have cumulative cultures which appear to be based on forms of social learning, why are their behavioural traditions so much more limited than those found in humans? That question is especially puzzling in the light of other species' abilities to imitate in rational ways, and the fact that humans seem sometimes to be less rational imitators. And what are the mechanisms involved in human imitation, if the capacity to imitate is not based on being endowed with a mirror system that develops independently of any learning?

When empirical results appear to pull in different directions and puzzles abound, a different theoretical framework is called for. One could think of social learning as being a species of individual learning, where the source of information happens to be the behaviour of conspecifics (in the same way as the behaviour of other species can carry useful information: we can learn that the song of the cuckoo is a sign of the arrival of spring). Just as features of the animate and inanimate environment can act as cues to features of the world that are relevant to an individual, so can features of conspecifics, including where and how aspects of their behaviour are performed. A quite different approach sees social learning as an inheritance system: a mechanism whose purpose is to pass phenotypic traits down the generations. From this perspective, the behaviour of conspecifics is not just another source of information on which individual learning can rely. That kind of individual learning from others is

sufficient on its own to give rise to cumulative cultures, but it is unlikely to be evolutionarily powerful. A mechanism which has the function of passing phenotypic traits down the generations will have additional features, adapted to their role in conveying information across multiple generations, features which appear paradoxical if we see social cues merely as an extra source of input for individual learning.

This paper argues that the empirical results can be usefully integrated around this question: does some type of imitation form an inheritance system? Several aspects of the unfolding empirical picture would fit together nicely if humans have an imitation-based inheritance system: the selectivity to copy only intentional actions and the disposition blindly to imitate seemingly irrelevant intentional actions; the tendency to imitate in ‘pedagogical’ contexts; and the finding that the ‘mirror system’ may not be innate or developmentally canalized. A distinctively human imitation-based inheritance system would also dissolve our central puzzle: why the cumulative cultures found in other species are so limited despite the rationality of some of the mechanisms involved. I will argue in the next section that features which are beneficial from the point of view of individual learning – like only copying those aspects of the model’s behaviour that are causally necessary to achieve some useful goal – may be maladaptive if the function is stably to transmit behavioural phenotypes down the generations, since they act as noise in the transmission system.

There are many other suggestions for explaining the uniqueness of human culture: language, powerful general intelligence, social emotions, cultural group selection, and so on. I am not offering the imitation-based inheritance system as a magic bullet that explains everything that is special about human cognition and social organisation. Instead, the hypothesis I want to assess is whether an imitation based-mechanism is one of the many things that make humans different. If we find that only humans have an inheritance-based imitation system, that would explain the comparatively limited cumulative cultures found in other species, and thereby form part of an answer to the broader question about human exceptionalism.

It is by no means clear that humans do in fact have an imitation-based inheritance system. Recent developments make that a hypothesis worthy of serious consideration, but there is evidence that counts both ways. Section 3 below offers a survey of the evidence, in particular of the results discussed in this volume. My purpose here is not to defend the claim that humans do have an imitation-based inheritance system. On that the jury is still out. Rather, I aim to show it to be a useful hypothesis for assessing the significance of several empirical results. The general idea is not wholly original, of course, but I will argue that it should be understood in a particular way. So in the next section I spell out in more detail what it would take for imitation to be evolutionarily powerful enough to qualify as an inheritance system.

2. Evolutionarily Powerful Imitation

2.1 Evolvability conditions

A Darwinian process of some kind will occur whenever certain very general conditions are satisfied. If individuals vary in their characteristics in a way that causes differential reproduction and is heritable, the population will evolve by natural selection (Lewontin 1970, Godfrey-Smith 2007). Satisfying those general conditions is a necessary but not a sufficient condition for the evolution of adaptive complexity. Different mechanisms that pass the threshold of giving rise to natural selection differ in their capacity to

produce complex adaptations and to sustain a large variety of adaptations. Some evolutionary mechanisms are more 'evolvable' than others (Sterelny 2001, 2004). Here I will focus on three sorts of conditions which are widely-accepted to make a Darwinian process more evolutionarily powerful.

Firstly, high copying fidelity: a mechanism which makes the phenotypes on which selection acts strongly heritable will be more evolvable. Copying fidelity is important if adaptations, once selected for, are to be preserved. Of course, a small rate of copying errors can play an important role in generating variation on which natural selection acts; and even with a higher error rate a broadly Darwinian process may occur, but accumulation of adaptive phenotypes depends on phenotypes, once selected for, being faithfully copied down the generations. Of course, this requires that any copy errors also be faithfully reproduced down the generations. Only then is cumulative selection possible. Otherwise, adaptations that have been selected can be lost too readily through copying errors.

Second are anti-outlaw conditions, which act to prevent phenotypic variants replicating themselves in ways that do not benefit the individuals which have them. An organism flourishes through the complex interdependence of many different phenotypes. Some of these could promote their own replication in ways that damage the organism which has them. When a somatic cell mutates and reproduces without limit, that leads to an increase in the number of cells of that new type (typically in a cancerous tumour), to the cost of the individual organism. The evolution of multi-cellular organisms depended on damping down the tendency for individual cells lines to go it alone (Maynard Smith & Szathmari 1995). Cancer cells miss out on having long-run descendants because they do not reach the next generation, since their effect is limited to the soma. So the existence of a single-celled stage in the life cycle helps to guard against such outlaws, making possible the evolution of multicellularity. If behavioural phenotypes can make use of imitation mechanisms to spread as memes, in ways that do not benefit individual organisms, then that serves to undermine the capacity of the imitation mechanism to form the basis of individual-level selection of adaptive traits. Imitation mechanisms will have greater evolutionary power if they include means of guarding against outlaws – features like the single-celled stage that is part of the life cycle of many multicellular organisms. Particularly important in this context are mechanisms that encourage vertical transmission of traits from parents to offspring. An item that can readily spread horizontally or obliquely can increase its relative frequency relatively independently of the effect it has on the individual reproductive fitness of the individuals to which it spreads. Such traits can spread like harmful viruses, according to an analogy that is familiar from the literature on memes. Mechanisms that favour vertical transmission from parents to offspring act to bring the fate of the replicating item into line with the reproductive fitness of the individual organism. So an imitation mechanism will, other things being equal, be more evolutionarily powerful if it has features encouraging vertical transmission of behavioural traits down the generations and discouraging horizontal transmission.

The third set of evolvability conditions concern the ability of an inheritance mechanism to transmit a wide range of novel variants. Novel variants may be produced by errors in the copying process, but how wide is the field of possibilities opened up thereby? How wide is the space of behavioural phenotypes that could be copied by imitation? Is variation limited to some pre-existing set of behavioural programs which the organism is capable of performing without learning; or perhaps to recombinations of elements from that set? The mechanisms of imitation will be more evolutionarily powerful if they have access to a very wide range of possible behavioural phenotypes. The broader the range of selectable variation, the easier it will be to build up adaptive complexity in a series of small,

gradual steps. Indeed, Maynard Smith & Szathmari (1995) argue that the appearance of unlimited systems of replication was a crucial transition in evolution. So it is important to know whether there are limits on the range of behavioural phenotypes that can be passed on by imitation, and whether that range varies between different species.

These evolvability conditions are representative, not exhaustive; but together they form a revealing lens through which to examine the empirical data on imitation. An imitation mechanism with these features would be more evolutionarily powerful, and hence more likely to give rise to complex cumulative culture, than one that just satisfies the bare necessary conditions on evolution by natural selection. Accordingly, if humans have an imitation mechanism with these features and other animals do not, we would have the bare bones of an explanation of why humans have more complex cumulative culture than found in other species.

2.2 Focus on individual-level selection

The focus here will be on the potential of imitation to be a driver of selection at the level of individual organisms, with patterns of behaviour being amongst the many phenotypic features of individuals that pass vertically down the generations and affect the reproductive fitness of those individuals. There may be other Darwinian processes going on as well. Boyd and Richerson have argued that social learning heuristics like conformity bias act as constraints on horizontal transmission of cultural variants so as to allow a form of cultural group selection for beneficial behavioural traits (Boyd and Richerson 1985). When taking this perspective, the entities over which the Darwinian process occurs are social groups themselves, which reproduce by giving rise to further social groups resembling the parent group in various behavioural or cultural respects. From the perspective of individual-level selection, the costs are high when useful information is shared horizontally, with non-relatives, since selection occurs on the basis of fitness differences between individuals, irrespective of changes in absolute fitness. In Boyd and Richerson's model those individual-level costs are offset by benefits to the group of sharing the useful behavioural phenotypes.

The papers in this volume that focus on culture (Whiten et al. 2009, Tennie et al. 2009) are interested in imitation as a driver of selection at the level of individual organisms. Here the fitness costs of sharing useful information are reduced by one of the evolvability conditions identified above: a restriction to vertical transmission. Although in many modern societies imitation may primarily act between unrelated individuals, in the recent evolutionary past, when people lived in much smaller kin-based groups, it is plausible that imitation operated mostly between older kin and younger kin (Mameli 2008). Such 'extended vertical cultural transmission' (Mameli) reduces the fitness cost of sharing useful information. Combined with other factors to facilitate evolvability, extended vertical cultural transmission has the capability to make behavioural traits acquired by imitation the basis of individual-level cumulative selection. In Boyd and Richerson's models of cultural group selection, conformity bias and success bias play the role of ensuring the retention of adaptive cultural variants, although they count against novel variants (which necessarily start at low frequency) being faithfully copied down the generations so as to form the basis of selection. There is not scope here to assess the importance of cultural group selection, but it is sufficient to observe that imitation is plausibly the basis of individual-level selection for behavioural phenotypes, which is at least compatible with the existence of Darwinian processes at the level of cultural groups, and may be more important in the cumulative selection of adaptive behavioural phenotypes.

Others see merit in ascribing fitness values to behavioural traits themselves – as memes – so as to analyse their spread in a population in a Darwinian way (Dawkins 1976, Blackmore 1999). As we have just seen, from the perspective of individual-level selection memes are outlaws that can undermine the accumulation of adaptive complexity. While imitation does have the potential to be an evolutionarily powerful process when it operates at the level of individual selection, it is much more controversial that the spread of memes could drive the evolution by natural selection of complex adaptations. So my focus here is not on those phenomena, but on imitation as a way of transmitting behavioural traits vertically between kin, as a potential supplement to other channels of genetic and epigenetic inheritance. If it turns out that imitation is not evolutionarily powerful at the level of individual selection, it remains open that it has an important role based on cultural group selection. Conversely, if there is good reason to think that imitation is evolutionarily powerful through vertical transmission and individual level selection, there is less need to find a group selectionist explanation for the same phenomena, although it remains possible that further group-based evolutionary processes occur in parallel.

2.3 Qualifying as an inheritance system

Scientific studies may indeed find that there is an imitation mechanism, perhaps unique to humans, that satisfies the evolvability conditions listed above (at least to some extent), and which thereby forms the basis for the cumulative selection of behavioural phenotypes. The question would then arise: why does the imitation mechanism have these neat features? One answer is that it is merely fortuitous: imitation happens to be embedded in the context of mechanisms which in fact operate to increase evolvability, but that occurred by chance. Another is that there was selection for these features, precisely because they do underpin evolvability. If these features were indeed selected, then their selective function is to give rise to heritable behavioural phenotypes. That is an additional function, over and above the particular behavioural functions of any of the phenotypes that are transmitted by this route. Compare DNA: it also has evolutionary functions of two kinds (Godfrey-Smith 1999). When a gene has been selected because of the fitness advantage conferred by the phenotype it produces by comparison to the phenotypes produced by its alleles, then that gene will thereby have an evolutionary function – to produce that phenotype. But the whole system of DNA, transcription, translation and so on has a further function of its own, a meta-function: to transmit phenotypes stably down the generations (on which natural selection will then act). Such a meta-function can only have arisen as a result of a series of selective events of the former, first-order kind. Natural selection relies on the existence of heritable variation, but selection for adaptations in this transmission mechanism can only take place over the course of a series of episodes of first-order selection.¹ It is reasonably clear that DNA does indeed have this meta-function. That can be inferred from the existence of mechanisms of DNA proof-reading and repair and their importance in germ line as well as somatic cell inheritance.

For an imitation mechanism to have this meta-function is for it, too, to have the evolutionary function reliably to transmit behavioural phenotypes, a function that we can infer from detailed knowledge of how the mechanism operates. There is still considerable disagreement about whether and

¹ At the first-order level, evolution may select for a gene for some trait T, whereas the meta-function arises by selection for the following more general capacity: if selection selects for some trait T, ensure T is reliably transmitted to future generations.

how traits that support evolvability could themselves have evolved (Maynard Smith & Szathmary 1995, Schlosser & Wagner 2004, Sterelny forthcoming). Lineage-based selection may have been important, although individual-level selection may be sufficient on its own.² It is reasonably clear that DNA has evolved by some process to have this kind of meta-function. The question here is whether there are some imitation mechanisms that have it as well. I reserve the term ‘inheritance system’ for a mechanism with the meta-function of passing phenotypic traits down the generations

So we can discern three increasingly strong claims that can be made about the evolutionary properties of an imitation mechanism. First is the claim that the imitation mechanism gives rise to a Darwinian process – by meeting the bare necessary conditions for evolution by natural selection. Second, that it founds the cumulative selection of behavioural phenotypes – by meeting the various conditions on evolvability discussed above, at least to some extent. The third and strongest claim is that there is an imitation mechanism which has the meta-function of giving rise to heritable behavioural phenotypes – which is an inheritance system. In the case of DNA, we inferred that it has a similar meta-function in order to explain the presence of various features of the mechanism that make best sense as being there for reasons of evolvability. The same logic can be applied to imitation, although here things are more complex. If empirical studies do indeed show that some form of imitation mechanism (at least in humans) operates in ways that meet the evolvability conditions, for example by increasing the fidelity with which behaviours (including novel variants) are copied or by encouraging vertical transmission between kin, then it is very plausible to infer that such features have been selected to play that evolutionary role. It follows that, just like DNA, the imitation mechanism does indeed have the meta-function of giving rise to heritable phenotypes.³

If there is a type of imitation that is an adaptation for transmitting behavioural phenotypes down the generations, then it certainly won’t be an entirely novel mechanism. It will make use of pre-existing psychological resources, many shared with other animals, and perhaps others that have evolved in humans for other reasons. Compare the evolution of the human vocal tract, for example, which

² It can seem that the capacity to evolve complex adaptations is of no benefit to individual organisms. But there is a fitness benefit for individuals, once we shift our focus from the short run fitness which is necessarily deployed in mathematical models (offspring at the first or second generation) and consider individual fitness in the very long run. Take copying fidelity as an example. When a trait arises by mutation that increases the fidelity by which selected traits are transmitted down the generations, that is likely to increase the long-run fitness of individuals bearing that trait. This is because, for each new trait for which there is first-level selection in its descendants, there will be more offspring in individuals that also have the high copying fidelity mechanism than in individuals that do not, since the advantageous trait will be driven out by copy errors in more of the individuals without the capacity for high fidelity copying (other things being equal). It follows from this pattern that the capacity for high copying fidelity will also have a selected function, the meta-function reliably to transmit behavioural phenotypes.

³ To strengthen that inference we need to rule out other explanations for the features in question which, in the case of imitation, includes the stories based on cultural group selection or trait-level memetic selection mentioned above. Since the processes which help cultural group selection get off the ground, like biases towards copying based on conformity and success, are in tension with those that are important for individual-level evolvability, the inference from individual-level evolvability to the meta-function looks plausible. But to make that case properly would require a more thorough treatment of rival group selectionist explanations than can be pursued here.

involved the re-arrangement of structural features like the tongue, epiglottis and larynx for which there were pre-existing developmental programs with their own functions, as well as the addition of novel features, so that the new complex could perform novel functions (Crelin 1987). The meta-function question asks whether, in the same way, there is a complex of features that together amount to a mechanism of imitation with the evolutionary function of giving rise to heritable behavioural phenotypes. Such a mechanism will likely co-opt psychological capacities present for other reasons but add features so as to perform the inheritance function.

2.4 Fidelity of transmission

Having motivated the interest in evolvability conditions, the next two sections show how they apply to the particular case of imitation mechanisms. The first set of evolvability conditions concern stability of transmission. Some level of heritability of phenotypes is a necessary condition for evolution by natural selection, but high fidelity transmission is more evolutionarily powerful. Selection leads trait frequencies in a population to carry useful information about the selective forces in which that population is found, and transmission errors degrade that information. So if there is an imitation-based inheritance system, we should expect it to have features designed to ensure that behavioural phenotypes are faithfully copied down the generations, including novel variants. There is clearly a balance to be struck because, although the generation of novel variants will constitute noise in the transmission system, the generation of *some* novel variants is a necessary condition on there being cumulative selection at all. In the context of an inherently noisy system like imitation learning, achieving high fidelity copying will be a more relevant constraint than ensuring that there is some noise in the copying mechanism. If there is to be cumulative selection, novel variants too must be copied – Tomasello's ratchet effect (1999) – but if individuals were regularly to vary their behaviour so that the behavioural traits they pass on to be imitated differ from those they learnt by imitation, then the copying mechanism would be too low-fidelity to allow adaptations to accumulate. Any increase in the frequency of a trait due to selection would be more than offset by copying errors that reduce the frequency of that trait in subsequent generations. So modifications to an imitated behaviour introduced by an individual, either accidentally or intentionally, even if made for good reason, will reduce the fidelity of the imitation-based inheritance system and hence, other things being equal, reduce its evolutionary power.

As well as modifications to learnt behaviour, a further source of noise is accidental bodily movements. These are around to be observed, even though the potential model is neither producing them because of having learnt them by imitation, nor because they intend them for other reasons. Imitation of accidental bodily movements would be reduced, and hence copying fidelity increased, by a bias towards copying only intentionally-produced bodily movements, mitigating the possibility of a behavioural pattern produced by accident being copied down the generations. That could be achieved without presupposing an ability to infer the mental states of others. One way is if accidental bodily movements have distinctive perceptual features that distinguish them from intentionally-produced actions, at least roughly. For example, actions may have smoother movement contours, and may slow and become more fine-grained as the limb or tool nears its objective. The fact that actions follow trajectories that correct for perturbations, whereas accidental bodily movements are deflected by external causes, may be a further perceptible clue. Goals which are emulated need not be understood as mental states, either. For these purposes, a goal is just some state of the world which a learner can observe, either an end state (e.g. getting termites out of the ground) or an intermediate state (e.g.

getting a short thick stick). Such goals can be copied without the learner having any thoughts about the mental states of the model. The same problem arises. How is the observer to distinguish changes in the model's environment that are produced intentionally (goals) from accidental effects on the environment? A rough and ready perceptually-based bias may be less readily available with respect to goals / outcomes. One alternative possibility is that the imitator only copies outcomes they have seen achieved multiple times, which would count against accidental outcomes to the extent that they occur less often than intended consequences (so there are fewer instances available to be observed);⁴ although that wouldn't stop commonly-occurring accidental side-effects being picked up and transmitted.

Another way that the fidelity of the copying process may be lowered is by individuals thinking for themselves. Individuals may be able to infer a more efficient way of bringing about the goal of some observed action. Or they may be able to discover a more efficient solution by trial and error learning. It can seem as if such modifications can only increase the potential of an imitation mechanisms to give rise to cumulative adaption, by allowing individuals to improve the adaptiveness of the behaviours they pass on. However, suppose the model has used a novel means for a good reason, but one that is transient: the model might act with their foot because their hands are occupied. Those kinds of individual innovations, although individually rational, from the point of view of an inheritance system would constitute noise in the process of transmission. Once animals have the capacity to tell for themselves what the most efficient ways of achieving a goal are, they may well be able to make distinctions amongst intentionally-produced behaviours. Where they see an inefficient means to an end which the model has good reason to adopt (e.g. because their hands are occupied), the means are unlikely to have been copied down the generations, and so only the goals of the action sequence should be emulated. However, when a model performs an action in a way which seems inefficient and is not due to some other constraint, the observer could have one of two dispositions. The rational reaction from the point of view of individual learning would be not to copy the inefficient behaviour but instead to perform an action that achieves the same goal more efficiently. However, the better disposition from the point of view of an inheritance system would be to copy the inefficient behavioural program – its very inefficiency suggests that it is only being produced because it has been copied down the generations. Of course, the model will not appreciate these reasons for having one disposition rather than another. These are the 'free floating' reasons of evolution rather than the explicitly-represented reasons of rationality (Dennett 1995). But if we find an imitation mechanism that has the latter disposition, its apparently paradoxical nature would be explained away if it is part of an inheritance system.

Accordingly, even if we find a disposition to copy irrelevant bodily movements in some contexts, we cannot conclude that the imitation is 'blind', because it may be part of a more sophisticated disposition: to copy goals and bodily movements, to ignore bodily movements where the model has a peculiar reason for achieving the goal in an inefficient way, but still copying bodily movements where there is no obvious reason for performing in an inefficient way. That is not blind imitation, but would demonstrate a sophisticated (albeit implicit) grasp of the intentional structure of the actions being copied. An analogous case is the experience of following the protocol for some complex experiment (although contrasting in being a case where the purposes are all explicitly represented and considered). Scientists generally copy all the details of a protocol. There may be no

⁴ Thanks to Cecilia Heyes for pointing out this possibility.

good reason for using 10ml rather than 20ml or 5ml of some solvent, say. That may just be the way it was first done, and since it worked no one has bothered to find out if the quantity could be varied. Indeed, some experiments are so tricky to get right that practitioners show an almost religious adherence to the letter of a known protocol. When an experiment takes hours or days to perform, there is very little motivation to put the result at risk in order to identify which steps or quantities are essential and which can be varied. So all sorts of techniques and steps are copied without any appreciation of whether or why they are necessary to achieve the goal – following an experimental protocol can feel rather like following a magic spell. But if some new piece of apparatus or technique comes out that represents a shortcut, it will likely be adopted. In such cases, the scientist can appreciate a reason why the recipe was previously performed in an inefficient way – because of a technological limitation – which reason has now changed. That is a rather sophisticated kind of selective copying, even though it involves copying some features without any appreciation of the reasons for them.

Huber et al. (2009) suggest a theoretical framework in which imitation is classified in terms of whether a movement, action or result is copied. This takes off from Tomasello's (1990) influential distinction between emulation and true imitation, based on the observation that chimpanzees will learn from observation that some new result can be achieved, and copy it, but tend not to reproduce the means that the model deployed to achieve that result. This is not the same as the distinction between goals and means to achieve them, since changes in the world can also be means to achieve a goal and, at least in humans, a performing a bodily movement can be a goal in itself. Rather, there is a continuum from the detailed motor programs deployed, assembled into a complex action sequence, through more coarse-grained descriptions of the same sequence as types of actions (e.g. reaching out to grasp an object), to the proximal and distal environmental results achieved. It is undoubtedly important to know where on this continuum a particular putatively imitated behaviour is located. But for our purposes a different way of dividing up the cases may be more important, namely how these actions are related to rewards. A seeming difference between copying bodily movements or merely copying worldly results may in fact reflect a difference in levels of reward-driven motivation. Even if an animal is capable of copying bodily movements, it will not manifest that ability if it is strongly motivated to obtain available rewards as quickly as possible by the most efficient means available. So while a difference in how much is imitated on the continuum between detailed motor programs and distal environmental results may be explained by a difference in the capacity to imitate bodily movements, it may also reflect a difference in the motivation to imitate bodily movements or in the relative strength of the motivation to obtain available rewards. A species under strong motivational pressure to obtain available rewards quickly and efficiently would be unlikely manifest the ability to imitate detailed bodily movements, even were they to possess it. So a difference that has standardly been seen as a matter of ability (lacking the ability to imitate actions or detailed bodily movements) might instead be the result of a difference in motivation. The evolutionary result is the same either way: emulation rather than strict imitation. Conversely, a species that is inclined to imitate irrespective of whether immediate rewards are achieved is more likely to imitate bodily movements as well as emulating rewarding goals.

At this point, a familiar debate in the imitation literature arises. Surely, the objection runs, imitated behaviours must be good for individuals to be selected? So the ability of an observer to weed out inefficient behaviours must be a good thing. Indeed, wouldn't a disposition to copy irrelevant behaviours undermine the utility of imitation, by causing the proliferation of all kinds of useless and costly behaviours? This debate is a great example of a theoretical tangle that can be nicely resolved by

adopting the inheritance system framework. The feature that looks good from the point of view of individual learning and outcomes in an individual lifetime – adopting the most efficient means to an end – represents damaging noise in a behavioural *inheritance* system. The optimal balance between the two dispositions depends upon the type of information individuals come across and how reliable it is.

Learning rates in individual learning offer a useful analogy. Just how heavily should a learner weigh a new piece of information? Suppose an agent is learning about probabilistic response-outcome contingencies in a stochastic environment in which those probabilities can change. In a stable environment where the probabilities linking responses with outcomes do not change much over time, an unexpected outcome, for example a response that doesn't produce an expected reward, should not weigh very heavily. It is likely to be just noise, so the learner should not adjust its view of the response-outcome contingency by very much. But in a volatile environment the same unexpected outcome carries much more information, because it may well indicate that the response-outcome contingencies have changed. So the learner should adjust its view of the response-outcome contingency much more (Behrens et al. 2007). In the stable environment, learning quickly from an unexpected outcome is maladaptive, reducing long-term payoffs. Taking the analogy across to inherited behaviours, they are well suited to carry information about long-run contingencies between behaviours and payoffs. Allowing individuals to learn quickly in response to local conditions may reduce those long-run payoffs. The imitation-based inheritance system can accumulate information about how behaviours tend to produce benefit in the environment in which they are performed down the generations, and individual learning would in some circumstances degrade the quality of that information.

When should we expect individual learning to be less adaptive, in the long run, than relying on inherited information? One case is when the outcomes are too stochastic for the sample encountered by individuals to have much reliability. Sterelny offers the example of a subsistence farmer deciding which cereal to plant (Sterelny 2006, Henrich 2002). When an individual plants a particular crop, the contingency between their choice and the eventual yield is highly statistically variable. Even when they pick the best crop for the most likely climatic conditions, there is so much variation in annual weather patterns that it will by no means always lead to a good crop; nor will badly chosen varieties always lead to poor yields. So the information available to an individual is just too impoverished to allow them on their own to build up much knowledge about likely crop yields. Sterelny follows Henrich in arguing that a tendency just to follow local customs arises to address this problem. If some folk agricultural practice has been inherited vertically down the generations, then its very persistence is a sign of its relative efficacy. An individual tendency to do what seemed to have the best outcome in individual experience would swamp this historical signal, and so prevent the accumulation of a useful body of folk agricultural practices. Food preparation taboos may be similar. Another case arises when the outcomes occur on a timescale which is of little use for individual learners. The connection between smoking and cancer is like that. The contingency is quite strong, but the cancer takes so long to arise that it is hard to detect the contingency. Surprising as it seems, it took all the power of modern medical statistics to discover the sad contingency which we now think is easy to observe in the lifetime of individual smokers. In other circumstances individual learning may be superior to a disposition to copy behavioural programs without assessing their efficacy in producing beneficial outcomes. Which disposition is normatively superior will depend on the structure of the problem space. When the information available to individual learners is highly statistically variable, or feedback takes too long to be usefully processed, then a disposition to block individual learning may be adaptive, in order to

preserve the signal of information built up over the generations. It may be that there is a greater disposition to learn by imitation in some domains rather than others, in a way that roughly tracks the variability in these trade-offs. If not, the relative cost of error assumes pressing importance. Maybe imitation would be a great way to pass on some skills that are beneficial in the long run in a way that is undetectable in the life of a single individual, but if an effect of having this imitative tendency is to block individual learning about important outcomes which the individual would be capable of learning about, then the relative fitness costs and benefits would favour individual learning over long-run imitation. The trade off depends upon the empirical details of individuals' capacities and their ecological niche.

2.5 Other evolvability conditions

So much for stability conditions and high fidelity transmission. The other two types of condition we will focus on involve less revision of theoretical perspective, and so can be dealt with more briefly. The second class of evolvability conditions are the anti-outlaw conditions, in particular a tendency for transmission vertically from parents to offspring. Vertical transmission helps guard against the proliferation of variants that replicate themselves at the expense of the long-run fitness of the individuals which bear them. As we have seen, verticality is not restricted to the parent-offspring relationship, but also includes transmission to more distant genetic relatives. So the question is whether there is a type of imitation which is constrained to encourage extended vertical transmission between kin (in the kinds of population structures in which it evolved). We can also ask whether there is any anti-outlaw mechanism equivalent to the bottleneck represented by the single-cell stage in the life cycle of multicellular organisms.

The third class involves the ability to transmit a wide variety of novel variants. The simple old picture held that imitation was an innate ability served by a special-purpose mechanism, more recently identified with the mirror system. To copy observed behaviours an imitator must somehow map perceptual inputs onto their own motor actions – the correspondence problem. An innate mirror system is postulated to 'solve' the correspondence problem by suggesting that a fixed library of mappings from perceptual input to motor output is developmentally canalized in each of us, maturing in a way that does not depend upon variable inputs from the environment (that being one sense of what it is to be innate). (Of course, that only solves the problem by pushing back the explanandum to the question of how we come to have such a developmental capacity.) However, if individuals come to imitative learning with a fixed solution to the correspondence problem – a developmentally canalized library of motor programs that they can recognise from perceptual input – then the scope for an imitation system to transmit a wide range of *novel* variants is seriously curtailed. Recombining these motor programs into arbitrary sequences may produce a range of novel patterns of behaviour to be inherited, but nowhere near as wide as if entirely new motor programs and motor-perceptual associations could be acquired. In principle, an imitation-based inheritance system could lie at many places on this continuum while still being evolutionarily powerful, but the greater the range of potential novel variants that could be passed on, the greater will be the capacity of the system to accumulate adaptive behavioural phenotypes (and hence the greater the importance to be attached to imitation as an explanation of the behavioural complexity of human culture). Although there is strong evidence that humans, at least, have a capacity to imitate various behaviours that is innate in the foregoing sense, more recent results raise the serious possibility that general purpose learning mechanisms have a role to play, either in addition to or instead of a fixed innate library of imitable motor programs.

If individuals do indeed learn perceptuo-motor mappings from individual experience, how does that square with the points about imitation-based inheritance being insulated from individual learning that we discussed a moment ago? The answer is to distinguish between solving the correspondence problem, on the one hand, and acquiring behavioural programs to be deployed in given contexts, on the other. An individual's tacit knowledge of which of their own motor programs correspond to various observed behaviours may depend on their own experience, but that does not determine which behaviours they will perform in which contexts; nor does it determine which goals they will pursue. To emulate goals and copy programs of behaviour in the pursuit of those outcomes depends on having a solution to the correspondence problem, but involves much more: chaining together action sequences and deploying them in a context-dependent way in the pursuit of relevant goals. I have tacit knowledge that would allow me to imitate an observed Nazi salute, but I don't have the disposition to perform that action in any context at all. Individual experience may be important in expanding the range of bodily movements that are available as novel variants that could be passed on by imitation. To be passed on by imitation, individuals must not only learn from experience how to execute such novel motor programs, and learn their perceptual correspondences; they must also learn by imitation in which context to perform that action. The learnt motor program must be coupled with a disposition to perform that action in some circumstances if the behaviour is to be repeated and then copied down the generations. It is the function of that complex which plausibly depends upon selection down generations of individuals. Therefore, if it turns out that the correspondence problem, or some aspect of it, is solved through individual learning, then that in no way counts against there being an imitation-based inheritance system. It may even count in favour of an inheritance function, by allowing for the transmission of novel variants. When the mirror system was discovered in macaques, and possibly in humans, the old story that viewed imitation as an innate special-purpose ability was supplemented with the claim that the mirror system is a canalized adaptation for solving the correspondence problem. Thinking carefully about evolvability causes us dramatically to change our perspective. A developmentally canalized mirror system had been thought to bolster the claim that imitation forms an inheritance system, but we now see that an imitation-based inheritance system would be more evolutionarily powerful if the mechanism for tackling the correspondence problem is not developmentally canalized, but instead learnt from individual experience, because that would dramatically increase the range of behavioural variants on which selection can act.

This section has motivated the hypothesis that humans may have an imitation mechanism which forms an inheritance system. The next section will survey recent developments in the empirical literature. Unsurprisingly, the picture that emerges is by no means clear-cut. Some lines of evidence favour the inheritance system hypothesis; others count against it. The burden of the present paper is not to establish that hypothesis, but rather to show how the inheritance system framework can provide a useful perspective on some otherwise puzzling emerging results.

3. Recent Developments in the Evidence

3.1 High-fidelity transmission

Human social learning is clearly different in many ways from social learning in other animals, just as there are many differences amongst other species. While it is reasonably clear that other species can learn by observation about new goals that are worth pursuing and new affordances of objects (emulation), it was thought that they would not imitate the actions by which such goals are achieved

(Tomasello et al. 1993). To the extent that other animals did copy bodily movements, that was an automatic effect that did not depend on the action being relevant to achieving some goal. More recent results, discussed by Huber et al., Tennie et al. and Whiten et al. in this theme issue, suggest that other species are capable of copying bodily movements, in a way that depends on the action being a means for achieving a useful outcome. For example, in a two action paradigm in which two groups of marmosets were selectively exposed to one of two different methods for achieving the very same result, marmosets tended to achieve that result by performing the same action as the one they had observed (Voelkl & Huber 2000). Voelkl & Huber (2007) went on to do a detailed movement analysis of marmoset behaviour, using a discriminant analysis classifier, which suggested that observers precisely copied the movement patterns of the models they had observed. Great apes can also be trained to copy novel actions from humans by learning the general rule ‘do as I do’ and extending it to novel actions as, it seems, can dogs (both summarised by Huber et al. 2009).

Dogs have been less thoroughly investigated than great apes. Their abilities in experimental situations be a result of domestic dogs having been selected for their ability to interact socially with humans (Hare & Tomasello 2005). Such results nevertheless offer a hint that broadly imitative abilities may spread much more widely than great apes or primates. There is evidence of goal-directed imitation in birds (Saggerson et al. 2005), dolphins (Rendell & Whitehead 2001), and even in fish. Archer fish (*Toxotes chatareus*) use a precisely directed shot of water to knock flying insect prey out of the air and into the water, where they are eaten. The best technique for hitting moving prey takes a while to learn by trial and error learning. Subordinate fish don’t themselves carry out any trials in the presence of a dominant fish, but it seems that observing the successful technique in the dominant allows them to perform it very rapidly when they first have the chance to try it themselves (Schuster et al. 2006). This is goal-directed behaviour, rather than automatic motor mimicry, in the sense that the fish only perform the action in furtherance of the goal of catching a passing flying insect.

Primates have a relatively sophisticated understanding of the causal structure of actions and the goals to which they are directed (Wood & Hauser 2008, cf. Povinelli 2000). They deploy this understanding to imitate selectively, not copying an inefficient means to a desired goal (Tennie et al. 2009, Whiten et al. 2009, Buttelmann et al. 2007). (This tendency has contributed to the view that primates are at best emulators of observed goals.) For example, all four species of great apes were able to learn by observation that the door to a box could be opened, and would then open it, but showed no tendency to do so in the same way as demonstrated (pushing vs. pulling – Tennie et al. 2006). Similarly, primates can emulate an observed goal using a novel means, other than the one they have observed (Tennie et al. 2009). Dogs, too, seem to imitate selectively to achieve the best functional fit between actions and goal (Range et al. 2007, Huber et al. 2009). These are perfectly rational tendencies, from the perspective of individual learning. But they act as serious limits on the power of imitation as an inheritance system in such species. Such individual modifications to the observed action constitute noise in the transmission of behavioural phenotypes down the generations. They block the transmission of behaviours which do not seem to observers to be efficient means of achieving goals. That may limit the range of behaviours that can be transmitted to those that individuals are able to understand the causal relevance of, which won’t extend very far beyond the range of behaviours individuals are capable of learning for themselves (the “zone of latent solutions”, Tennie et al. 2009). And if an individual does hit on an improved way of reaching a goal that is non-obvious, that information will be rapidly degraded in transmission as individuals adopt what seem to them to be more

efficient ways of reaching the goal. These tendencies, despite their undoubted advantages as aids to individual-level learning, are likely to restrict the extent to which imitation can form the basis of cumulative selection. Similarly, primates are resistant to learning a new strategy at all if they already have a strategy that works (Tennie et al. 2009). That tendency too will impair the transmission of behavioural phenotypes down the generations.

By contrast with the rational selective imitation seen in primates, human children sometimes demonstrate surprisingly unselective ‘blind over-imitation’. Meltzoff (1988) showed that 14 month old infants would imitate a seemingly inefficient bodily movement, using the forehead to switch on an illuminated lightbox one week after they had seen the same action demonstrated. It has since been confirmed that, at many ages, children are inclined to imitate a method that is transparently inefficient, in contrast to the more selective copying seen in other primates (Tomasello 1999, Horner & Whiten 2005, McGuigan et al. 2007, Lyons et al. 2007, Whiten et al. 2009). One result suggests that this ‘blind over-imitation’ may be more sensitive to rational considerations than it first appears. Gergely et al. (2002) repeated the Meltzoff paradigm – infants watching a model turn on a light with her head – but added a condition in which the model’s hands were otherwise occupied. When the model had good reason to use her forehead to turn on the lightbox, because her hands were occupied with holding a blanket, 14 month olds were much less likely to imitate the unusual head action. One interpretation is that the infants were copying the unusual head action in the ‘hands free’ condition precisely because it seemed to them unnecessary (but see Jones 2009).

Given the relative facility shown by infants in understanding and even copying goals (Carpenter et al. 2005, Csibra 2008), the contrast between apes and children demonstrated by Horner and Whiten (2005) can seem particularly puzzling. Chimpanzees will imitate a series of actions performed on an opaque box before a reward is obtained, but will cut to the chase when the same series is performed on a transparent box. When the box is transparent, it is obvious that some of the actions are unnecessary to reach the goal. By contrast, children keep on performing the unnecessary actions, even on the transparent box. Lyons et al. (2007) suggest that children know the steps are unnecessary, but imitate them anyway. There is even some evidence that children follow a developmental trajectory (Whiten et al. 2009), starting off like chimps as more selective imitators, but increasing in their tendency to imitate unnecessary actions as they grow from three to five years old.

These results become much less puzzling in the light of the hypothesis that human children have an imitation-based inheritance system whereas other animals do not. The tendency to imitate unnecessary actions will help to promote high-fidelity copying of behavioural traits down the generations. Its presence suggests an imitation-based mechanism with the meta-function of ensuring high-fidelity copying. The nascent developmental story, if vindicated in further studies, would be a nice confirmation, suggesting that the capacity for ‘over-imitation’ develops after, and is a kind of add-on to, a more ancient and widely-shared capacity to copy the actions needed to achieve a goal.

That hypothesis also throws light on why other species should primarily be emulators, even if they have the capacity to imitate the means or bodily movements used by a model to achieve a desired end. If imitation is not adapted in other species to the meta-function of faithfully copying behavioural phenotypes, then this capacity will usually be masked by the drive towards achieving the rewarding outcome directly, by the most obvious method.

Other features of human imitation are more equivocal in respect of the imitation system hypothesis. There is evidence that children will only imitate action sequences which are successful (Want & Harris 2001, Harris & Want 2005); perhaps also when they understand the intended goal of a model who tries and fails (Tomasello & Carpenter 2005). This looks like the kind of individual learning that would interfere with high-fidelity transmission, although it could be that achieving a goal, or understanding what the goal is, acts as a basic input rule, highlighting sequences of actions that are candidates for copying and distinguishing them from accidents and other bodily movements that are not. Indeed, we remarked above that a capacity to distinguish intentionally-produced behaviour from mere accidents would improve the fidelity of behavioural transmission. So it may be that, in order to trigger ‘blind over-imitation’ of details of observed actions, the child has to appreciate that the model’s action successfully achieves some result, or is at least aimed at a goal. Revealingly, although children will copy a goal that a human model attempts but fails to reach, they do not copy the goal when very similar movements are performed by inanimate devices (Meltzoff 1995, 1996, Tomasello & Carpenter 2005). Of course, children’s ability to distinguish intentionally produced actions from both inanimate movements, and mere accidents (Carpenter et al. 1998), is neutral as between the individual learning hypothesis and the inheritance system hypothesis.

We remarked above on the cost-benefit trade-off between the benefits of high-fidelity transmission down the generations and costs in terms of lost opportunities for the individual to benefit from learning in its own lifetime. Differences in ecology and selective environment may form part of an explanation of why humans have an imitation-based inheritance system (if they do). We also noted that these costs and benefits are different in different domains (e.g. growing crops vs. learning to crack a nut). The suggestion that humans deploy different social learning mechanisms to address different social problems (Gergely and Csibra 2006) may turn out to reflect some of these differences, although Gergely and Csibra themselves explain these differences in terms of features of the individual learning problem. That is quite different from an explanation in terms of the costs of adopting a relatively inflexible high-fidelity copying solution where successful strategies are selected in a process that takes place over many generations rather than in individual ontogeny.

In short, the hypothesis that there is an imitation-based inheritance system in humans (i.e. one with the meta-function of transmitting behavioural phenotypes down the generations), but not in other animals, makes some otherwise puzzling differences between the species neatly explicable. Although the evidence so far is by no means all supportive, let alone conclusive, the inheritance system hypothesis clearly merits further investigation.

3.2 Anti-outlaw conditions

If imitation is deployed in humans as an inheritance system, then we should expect to find mechanisms that operate to discourage horizontal or oblique transmission of such behavioural traits (to non-relatives, to the comparative disadvantage of the model). (Recall that enhancing the fitness of non-relatives, by non-vertical transmission, represents a cost in terms of relative fitness.) One mechanism might simply be a disposition to imitate more in childhood, while the learner is still in the proximity of its parents (Schiel and Huber 2006). Gergely and Csibra’s natural pedagogy is another obvious candidate to be a mechanism encouraging verticality (Gergely and Csibra 2006, Gergely et al. 2007, Tennie et al. 2009; although see Virányi and Range, in press, suggesting that selective imitation in

pedagogical contexts may not be limited to humans). If children tend only to imitate when a pedagogical context is signalled, that disposition may have acted in recent evolutionary history to ensure copying from parents predominates, limiting the extent to which they copy behaviours horizontally or from unrelated individuals. Why would defectors not arise that also imitate non-relatives? Perhaps because signalling a pedagogical context is a piece of co-operative signalling. It aids the learner in indicating which classes of actions ought to be imitated, but the production of the signal is under the control of the model. So the picture is of models using cues such as eye contact, contingent reactivity, and the imitator's own name to signal when actions ought to be copied; and of their doing so more often to kin than to others, at least in recent selective environments. That may be achieved by models selecting to whom they signal a pedagogical context, perhaps combined with social structures that ensure adult-child interactions are largely between relatives.

Unlike the features discussed above that suggest a high-fidelity copying mechanism, identifying natural pedagogy as an anti-outlaw mechanism is much more speculative. A line of evidence that counts against this interpretation is the data on automatic imitation effects. These will be discussed further in the next section when we ask about the library of perceptuo-motor mappings that enable imitation to take place at all. Here, it is the social role of automatic imitation that creates trouble for the inheritance system hypothesis. Research in social psychology has shown the pervasive influence of various broadly imitative effects on action, in a way that is not subject to conscious control (Bargh & Ezequiel 2008). These are effects of a different kind than the goal-directed imitation discussed above. They do not depend on the imitator understanding the model's actions to be intentional or goal-directed. Instead, features of ongoing actions can be modified or primed by the actions of a model (and by semantic cues, operating unconsciously). Social psychologists suggest that these forms of synchronous behaviour, far from being mere epiphenomena, play an important social role (Van Baaren et al. 2009). Bastiaansen et al. (2009) catalogue a variety of phenomena in which perceptual, sensory and emotional mirroring occurs between people. Bastiaansen et al. argue that such capacities have important social functions. Such mirroring can lead to emotional affiliation in dyadic interactions, and can form the basis of defining social groups. If that is right, then these kinds of broadly imitative tendencies may have a different kind of adaptationist explanation, one based on group selection (Boyd and Richerson 2005). That case is bolstered by the finding that people do deploy a variety of social learning strategies (Laland et al. 2004), which are susceptible to group-selectionist explanation. As remarked in section 2.2 above, the existence of a group selectionist explanation for imitative capacities, while not strictly incompatible with the existence of an imitation-based inheritance system based on vertical transmission, would undercut some of the motivation for thinking that a form of imitation has a meta-function based on vertical transmission. The inheritance system hypothesis might still be vindicated if a sharp separation can be made between automatic imitation effects and copying of sequences of actions deployed in goal-directed behaviour, but it is by no means clear that this fine distinction corresponds to categories that can be clearly distinguished empirically.

There may be other features of goal-directed imitation that can be seen to be anti-outlaw mechanisms. Indeed, since little investigation to date has been directly motivated by the inheritance system hypothesis, it wouldn't be surprising if there are further phenomenon associated with imitation still to be discovered. Section 3.4 suggests some places to look for anti-outlaw mechanisms.

3.3 Transmission of novel variants

To imitate the action of another, an individual has to have some way of going from a perceptual representation of an observed action / outcome to a motor representation of an action or sequence of actions suited to achieving the same. This is the ‘correspondence problem’. Mirror neurons are a candidate mechanism for achieving this correspondence. Finding appropriate mirror neurons in the brain shows that an individual has managed to solve the correspondence problem for some range of actions, but does not tell us how: whether through a mechanism selected in phylogeny and canalized in development, or through individual learning. Although having some correspondence mechanism is an essential prerequisite to the kind of goal-directed imitation that is the subject of the inheritance system hypothesis, it is by no means all that is required. So we should distinguish between correspondence mechanisms and low-level automatic mirroring, on the one hand, and broadly goal-directed emulation and imitation, on the other. Our focus above was on relatively sophisticated capacities for learning from observation, both in humans and in other species, e.g: emulating the goal of an action sequence; distinguishing actions from inanimate movements, and from accidents; imitating clearly inefficient steps directed at a goal; etc. Such abilities presuppose that the individual has a capacity that delivers correspondences between what is perceived and how they can act, but correspondence mechanisms should not be equated with the kinds of imitation that are candidates to form an inheritance system.

Rumiati et al. (2009) review the long history of neuropsychological evidence that substantiates a similar distinction between two kinds of neural mechanisms. One mechanism is the basis on which people respond to perceptual stimuli by having an idea of a corresponding movement (\approx the correspondence mechanism discussed above). Impairment of this mechanism leads to a reduced ability to imitate gestures, called ideational apraxia. On the other hand there is a set of mechanisms involved in execution or inhibition of such movements, impairment of which is characterised as ideomotor apraxia. Neuropsychological cases are consistent with the evidence from social psychology discussed above, suggesting that the default mode of operation of the correspondence mechanism leads directly to automatic mirroring actions. In normal cases, this is largely inhibited. But deficits in inhibition, often caused by lesions to the prefrontal cortex, reveal the operation of the correspondence mechanism, manifesting in patients with severe difficulties in controlling the tendency to imitate in inappropriate ways.

The sublexical or direct route to action discussed by Rumiati et al. reflects the library of perceptuo-motor mappings for which the subject has a correspondence mechanism. The fact that ideomotor apraxia is a pathological condition reinforces the point about goal-directed imitation being a complex capacity, of which a correspondence mechanism is only one component. Ferrari et al. (2009) report evidence that a similar distinction between two different routes to imitative behaviour is found in macaques. From the perspective of cognitive neuroscience, Brass et al. (2009) also find an important distinction between mirroring others’ actions and appreciating the goals, intentions and other mental states of others. Although goal-directed imitation may not require mentalizing in any rich sense, it does show a sophisticated responsiveness to the goals and reasons of the imitated model, in a way that goes far beyond mere mirroring. Interestingly, Brass et al. argue that mentalizing actually arises from systems that control low level mirror systems, in order to make a distinction between self and other that a simple correspondence mechanism will fail to keep track of. Both Rumiati et al. and Brass et al. echo aspects of Susan Hurley’s shared circuits model, in which the capacity for keeping track of the

mental states of others is argued to flow from the use of inhibition to distinguish between self and other (Hurley 2008).

In short, there are several lines of evidence substantiating the existence of two routes to broadly imitative behaviour, one involving automatic mirroring, and the other roughly at the intentional level. The correspondence mechanism underlying automatic mirroring provides the basis on which a goal-directed imitator is able to perform the actions they observe. As such, it furnishes a library of behavioural programs that are possible candidates to be passed on by an imitation-based inheritance system. Recall that the inheritance system hypothesis suggests that there should be a library of candidate behavioural programs that goes wider than those that are currently passed on by imitation – the library of variants that could be passed on by vertical transmission and on which natural selection could then act. So far, so good for the inheritance system hypothesis. However, there is another strand to the perspective offered by Rumiati et al., and by Brass et al., which is unfavourable to the inheritance system hypothesis. Both emphasise a role for high level intentional resources, like mentalizing, in the operation of the indirect or semantic route to behaviour. If this kind of imitation always involves the imitator thinking for themselves and reasoning out what to do, then it is correspondingly less likely to be conducive to the high-fidelity copying of behavioural programs down the generations. Ideomotor approaches to imitation, as championed by Wolfgang Prinz (Prinz 1990, Massen & Prinz 2009), also tend to view imitation as a much more general rational process, which would be rather unsuited to be a channel for high-fidelity copying of behavioural phenotypes. As discussed in section 3.1 above, individual thought and learning is likely to introduce too much noise into the transmission process for any signature of selection over the generations to persist. Evidence of ‘blind over-imitation’ points in the opposite direction, but it is far from clear which is the more important tendency in natural settings.

A second putative feature of the mirror system does, however, count much more clearly in favour of the inheritance system hypothesis. It is standardly assumed that the existence of an innate mirror system would support there being a domain-specific role for imitation-based learning. However, we saw in section 2.3 above that the inheritance system hypothesis makes matters look quite different. If the imitable behaviours are limited to those furnished by an innate mirror mechanism, the power of an imitation-based mechanism to pass on novel variants is significantly constrained. Recall Maynard Smith & Szathmari’s (1995) emphasis on the superior evolvability of unlimited systems of replication. The discovery that perceptuo-motor mappings can be learned in individual experience (e.g. Prather et al. 2008) therefore counts clearly, although rather unintuitively, in favour of the inheritance system hypothesis.

Ferrari et al. (2009) suggest that macaques are born with a ‘prewired’ but wide library of imitative tendencies. But they argue that, through the involvement of mesial and prefrontal cortex, behavioural sequences can be ‘parsed’ into elements that can be combined to perform novel actions. Vogt et al. (2007) interpreted corresponding activations found in an fMRI study as reflecting subjects’ parsing of observed actions into a goal and specific motor patterns. Recombination of such elements would significantly increase the range of behaviours available to be transmitted.

More radically, recent evidence suggests that even classical mirror mechanisms are acquired through individual experience. Neonatal imitation in humans (Meltzoff & Moore 1977, 1983) and other species (Ferrari 2009) are evidence that some correspondence mechanisms already exist at birth, but the experiments reported by Jones (2006, 2009) suggest that the observed behaviours, in human

neonates at least, may not depend upon a correspondence mechanism at all. Whether or not some correspondence mechanism develops independently of experience, the results reported by Catmur et al. (2007, 2009) strongly suggest that existing correspondence mechanisms can be added to or adapted through associative learning from individual experience. Rizzolatti argues that novel actions can be incorporated into the pre-existing mirror system (Rizzolatti 2005, Huber et al. 2009), whereas Catmur et al. argue that all correspondence mechanisms are acquired by experience. Either way, if it is right that the range of behaviours that can be passed on by imitation is not limited to some innate library of perceptuo-motor mappings, then the putative imitation-based inheritance system has greater evolutionary power, and its capacity to explain the exceptionalism of human culture is correspondingly increased.

3.4 New Empirical Questions

The inheritance system hypothesis produces its own agenda about which are the most important avenues for future research, which complements the research questions made salient by taking other perspectives. The hypothesis also generates some new questions of its own.

Questions about the rationality and selectivity of imitation are central to assessing the fidelity of imitation as a putative imitation system. We observed that primates, although capable of imitating an action in a two action test, are much more inclined to go for goals, especially if they already have a technique for reaching the goal. Correspondingly, how much are children willing to override a known, individually-learned action for achieving a goal when an alternative means is demonstrated? It would also be good to know more about the kinds of situations that trigger imitative behaviour. Does it depend on inferring the model's intentions, or are goal-directed actions distinguished from accidents and non-actions by more low-level features?

The suggestion that human imitation depends upon a natural pedagogical context could be explored much more. In particular, it would be useful to know how much the kind of 'blind over-imitation' which I have argued could be characteristic of high-fidelity copying, is selectively deployed in pedagogical contexts. Does 'blind over-imitation' depend upon triggering by cues like eye gaze and use of the subject's own name? Furthermore, does this tendency depend upon children seeing the model's behaviour as normative – as something they ought to do? If so, what is the source of that normativity: the social authority of the model, or just features of the learning situation? Is there anything like this normative dimension in any of the imitative behaviour of other species?

On the topic of the relative costs and benefits of individual learning vs. blind imitation, it is important to know much more about whether imitative tendencies are domain-sensitive (e.g. food preferences, tool use, social roles, etc.) and, if so, whether that sensitivity marries up with the quality of the information about the domain that would be available through a channel of vertical inheritance.

A lot of research has been done on the level at which different species imitate: goal / action / bodily movement. If the capacity to imitate bodily movements is more widespread than initially thought, then the suggestion above about differences in motivation is worth exploring. Perhaps chimps are just much more goal-focused than humans. If that is part of the explanation of the fact that primates operate largely in the domain of emulation rather than true imitation, then it should be possible to prompt greater motor-imitation by introducing contexts with less competition pressure. The

tantalising video of infant motor copying referred to in Marshall-Pescini & Whiten (2008) is a plausible illustration of how non-competitive contexts can draw out the capacity to imitate motor programs.

Another set of questions concern the mirror system, the putative library of behavioural programs available to be copied down the generations. A basic question which has been underexplored is to investigate how tight the connection is between the perceptuo-motor knowledge found in an individual's mirror system and the range of behaviours they are able to imitate. It has been largely presumed that the one enables the other, but it is not entirely clear that they even have the same range of operation. Correlatively, is the range of behaviours that can be learnt by imitation expanded or altered when an individual's mirror system is modified through associative learning (as in Catmur et al. 2007)? Once we are thinking of the mirror system as a (possibly extensible) library of selectable variants, questions about cultural variation in that library become pressing. How much do tendencies for automatic imitation vary across human cultures, and does that variation line up with the different behaviours and skills that are passed on by imitation in different societies?

In discussing features like high-fidelity copying that could underpin an imitation-based inheritance system, I have remained neutral about how those features themselves are inherited. A ready thought is that such features have arisen by genetic selection, but that is not obligatory – they too could have arisen by culture-based processes. Any processes of extended vertical transmission is a candidate. That possibility has not, as far as I am aware, yet received detailed empirical investigation, by comparison with the large body of work that investigates how particular behavioural adaptations are culturally transmitted. The picture becomes more complicated, but it does make it important to assess the extent to which tendencies like 'blind over-imitation' and natural pedagogy are themselves the result of cultural inheritance. The relevant questions there are the usual ones: developmental canalization and inter-societal variation / universality. Relatedly, we can ask just how much imitation has been a *vertical* transmission process. In modern societies, with professional teachers and a range of other media from which children learn, it seems unlikely that vertical transmission dominates. But the question remains of whether vertical transmission is plausible in the recent evolutionary past. And if modern societies differ in this respect, we should be able to find the signs of imitation-based vertical inheritance losing its adaptiveness as it is swamped by horizontal transmission.

Finally, the inheritance system hypothesis suggests other places where we could look for features that encourage evolvability, by analogy with other forms of inheritance. Are there any bottlenecks that act as anti-outlaw mechanisms? (Natural pedagogy may have a role here, too.) Is there any measure of developmental modularity in the mechanisms of imitation, so that learning a new behavioural sequence by imitation interferes only minimally with existing behavioural phenotypes? Are there any examples of biased transmission, like segregation distorter genes, or the bias towards feminising their hosts displayed by some strains of *Wolbachia* (Sterelny 2004)? Correspondingly, are there mechanisms that operate to block biased transmission? Sterelny (2004) observes that uniparental inheritance of mitochondrial DNA acts as an anti-outlaw mechanism. Gender-specific patterns of imitative behaviour might have a corresponding effect at the level of behavioural phenotypes, but gender differences in imitation have not been explored with that in mind, to test the inheritance system hypothesis. Doubtless, most of these suggestions and analogies will lead to false starts and dead ends; but at the very least the inheritance system hypothesis generates some ideas for discovering new phenomena associated with imitation, going beyond the impressively detailed hard-won knowledge that has already been gained on the topic.

4. Conclusion

The theoretical perspective developed here underlines the interest of the diverse lines of empirical enquiry reported in this theme issue; disciplines that have been brought together by the editors through their EDICI project.⁵ The payoff shows the value of pursuing such questions in the integrated interdisciplinary fashion that the EDICI programme has been so successful in promoting. When recent empirical discoveries about the mechanisms and functions of imitation are brought together, they can appear puzzling. They come into sharper focus if viewed from the perspective of the theoretical questions suggested here: is there an evolutionarily-powerful imitation-based inheritance mechanism? If so, is that a mechanism with the meta-function of transmitting heritable behavioural phenotypes down the generations – is it an *inheritance system*? We have seen that the findings so far are far from conclusive. But if the answer to one or both of these questions is yes, that would be part of a satisfying explanation of why humans have more complex cumulative cultures than those found in other animals, including in species who share many of our social learning mechanisms. It would thereby form one part of an explanation of human exceptionalism. However, as this paper has argued, to have the evolutionary power of an inheritance system, a mechanism must meet a range of demanding requirements, much stronger than the bare conditions that are enough to give rise to some form of evolution by natural selection. The explanatory payoff is substantial, but the hypothesis is exacting.

Acknowledgements

The author is grateful for comments from and discussions of this material with Dan Fessler, Peter Godfrey-Smith, Paul Griffiths, Ben Kerr, Dan Sperber and audiences at Oxford, the Australian National University and at the EDICI workshop in Vienna. Thanks are due especially to Cecilia Heyes, Kim Sterelny, Matteo Mameli and Ludwig Huber for helpful and detailed suggestions for improving the current manuscript. Work on this paper was supported by the OUP John Fell Research Fund, the James Martin 21st Century School, the Oxford Centre for Neuroethics and the Mary Somerville Junior Research Fellowship, Somerville College.

References

- Bargh, J. & Ezequiel, M. 2008. The unconscious mind. *Perspectives on Psychological Science*, 3, 73-79.
- Bastiaansen, J. A. C. J., Thioux, M. A. & Keysers, C. M. 2009. Evidence for mirror neurons in emotions. *Phil. Trans. R. Soc. B*.
- Behrens TE, Woolrich MW, Walton ME, Rushworth MFS. 2007. Learning the value of information in an uncertain world. *Nature Neuroscience* 10(9):1214-21.

⁵ A project funded by the European Community's Sixth Framework Programme under contract number NEST 012929 to Ludwig Huber, Cecilia Heyes, Marcel Brass and Gyorgy Gergely with the title 'Evolution, development and intentional control of imitation'; for further information: www.univie.ac.at/edici.

- Blackmore, S.J. 1999: *The Meme Machine*. Oxford: O.U.P.
- Boyd R. and Richerson P. 2005. *The Origin and Evolution of Cultures*. Oxford University Press, New York.
- Boyd, R. and Richerson, P.J. 1985: *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.
- Brass, M. Ruby, P. & Spengler, S. 2009. Inhibition of imitative behaviour and social cognition. *Phil. Trans. R. Soc. B*.
- Buttelmann, D. Carpenter, M., Call, J., Tomasello, M. 2007. Enculturated chimpanzees imitate rationally. *Dev. Sci.*, 10, F31–F38.
- Carpenter, M., Akhtar, N. & Tomasello, M. 1998 Fourteen- through 18-month-old infants differentially imitate intentional and accidental actions. *Inf. Behav. Dev.* 21, 315-30.
- Carpenter, M., Call, J., & Tomasello, M. 2005 Twelve- and 18-month-olds copy actions in terms of goals *Dev. Sci.* 8, F1-F8.
- Catmur, C., Walsh, V. & Heyes, C. M. 2007 Sensorimotor learning configures the human mirror system. *Curr. Biol.*, 17, 1527-1531.
- Catmur, C., Walsh, W. & Heyes, C. 2009. Associative sequence learning: The role of experience in the development of imitation and the mirror system. *Phil. Trans. R. Soc. B*.
- Crelin, E. 1987: *The Human Vocal Tract: Anatomy, Function, Development and Evolution*. New York: Vantage Press.
- Csibra, G. 2008 Goal attribution to inanimate agents by 6.5-month-old infants *Cognition* 107, 705–717
- Dawkins, R. 1976: *The Selfish Gene*. Oxford: O.U.P.
- Dennett, D.C. 1995: *Darwin's Dangerous Idea*. New York: Simon & Schuster.
- Ferrari, P. F., Bonini, L. & Fogassi, L. 2009. From monkey mirror neurons to primate behaviours: Possible ‘direct’ and ‘indirect’ pathways. *Phil. Trans. R. Soc. B*.
- Gergely, G., Bekkering, H. Kiraly, I. 2002. Rational imitation in preverbal infants. *Nature* 415, 755.
- Gergely, G., Csibra, G. 2006. Sylvia’s recipe: The role of imitation and pedagogy in the transmission of cultural knowledge In S. Levenson & N. Enfield (Eds.) *Roots of Human Sociality: Culture, Cognition, and Human Interaction*, Oxford: Berg Publishers
- Gergely, G., Egyed, K. and Király, I. 2007. On pedagogy. *Dev. Sci.*, 10, 139-46.
- Godfrey-Smith, P. 1999: Genes and codes: lessons from the philosophy of mind. In Hardcastle, V. (ed.) *Where Biology Meets Psychology: Philosophical Essays*. London / Cambridge, MA: MIT Press.
- Godfrey-Smith, P. 2007: Conditions for evolution by natural selection. *Journal of Philosophy*, 54, 489-516.

- Hare B & Tomasello M 2005 Human-like social skills in dogs? *Trends Cogn Sci* 9(9):439–444
- Harris, P. & Want, S. 2005. On learning what not to do: The emergence of selective imitation in young children's tool use. In: *Perspectives on imitation: From neuroscience to social science*, vol. 2, ed. S. Hurley & N. Chater, pp. 148–62. MIT Press.
- Henrich, J. 2002. Decision-making, cultural transmission and adaptation in economic anthropology. In J. Ensminger (ed.) *Theory in Economic Anthropology*. Berkeley, CA: AltaMira Press, 251–295.
- Horner, V. K. & Whiten, A. 2005 Causal knowledge and imitation/emulation switching in chimpanzees (*Pan troglodytes*) and children. *Anim. Cogn.* 8, 164-181.
- Huber, L., Range, F., Voekl, B., Szucsich, A., Viranyi, Z. & Miklosi, A. 2009. The evolution of imitation: What do the capacities of non-human animals tell us about the mechanisms of imitation. *Phil. Trans. R. Soc. B*.
- Hurley, S. (2008). The shared circuits model (SCM): How control, mirroring, and simulation can enable imitation, deliberation, and mindreading. *Behavioral & Brain Sciences*, 31, 1-58.
- Jones, S. S. 2006. Exploration or imitation? The effect of music on 4-week-old infants. tongue protrusions. *Infant Behavior & Development*, 29 (1), 126-130.
- Jones, S. S. 2009. The development of imitation in infancy. *Phil. Trans. R. Soc. B*.
- Laland, K. N. 2004 Social learning strategies. *Learning & Behavior*, 32, 4-14.
- Lewontin, R.C. 1970: The Units of Selection. *Annual Review of Ecology and Systematics*, 1, 1-14.
- Lyons, D. E., Young, A. G. & Keil, F. C. 2007 The hidden structure of overimitation. *Proc. Nat. Acad. Sci. USA* 104, 19751-19756.
- Mameli, M. 2008: Understanding culture: a commentary on Richerson and Boyd's Not By Genes Alone. *Biology and Philosophy*, 23, 269–81.
- Marshall-Pescini, S. & Whiten, A. 2008 Social learning of nut-cracking behaviour in East African sanctuary-living chimpanzees (*Pan troglodytes schweinfurthii*). *J. Comp. Psychol.*, 122, 186-194.
- Maynard Smith, J. and E. Szathmary 1995. *The Major Transitions in Evolution*. New York, Freeman.
- Massen, C. & Prinz, W. 2009. Movements, actions and tool-use actions: An ideomotor approach to imitation. *Phil. Trans. R. Soc. B*.
- McGuigan, N., Whiten, A., Flynn, E. & Horner, V. 2007 Imitation of causally-opaque versus causally transparent tool use by 3- and 5-year-old children. *Cognitive Development*, 22, 353-364.
- Meltzoff, A. 1988. Infant imitation after a 1-week delay: Long-term memory for novel acts and multiple stimuli. *Developmental Psychology* 24:470–76.
- Meltzoff, A. 1995. Understanding of the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology*, 31, 838–50.

- Meltzoff, A. 1996. The human infant as imitative generalist: A 20-year progress report on infant imitation with implications for comparative psychology. In *Social learning in animals: The roots of culture*, ed. C. Heyes & B. Galef, Jr., pp. 347–70. Academic Press.
- Meltzoff, A.N. & Moore, M.K. 1977 Imitation of Facial and Manual Gestures by Human Neonates. *Science*, 198, 75-78.
- Meltzoff, A.N. & Moore, M.K. 1983 Newborn infants imitate adult facial gestures. *Child Dev*, 54, 702-709.
- Povinelli, D. 2000. *Folk physics for apes: the chimpanzee's theory of how the world works*. Oxford / New York: OUP.
- Prather, J. F., Peters, S., Nowicki, S., & Mooney, R. (2008). Precise auditory-vocal mirroring in neurons for learned vocal communication. *Nature*, 451, 305-312.
- Prinz, W. (1990). A common-coding approach to perception and action. In O. Neumann & W. Prinz (Eds.), *Relationships between perception and action: Current approaches*, 167-203. Berlin: Springer-Verlag.
- Range, F., Viranyi, Z. & Huber, L. 2007. Selective imitation in domestic dogs. *Curr. Biol.* 17, 1-5.
- Rendell, L. & Whitehead, H. 2001. Culture in whales and dolphins. *Behavioral and Brain Sciences* 24, 309-382.
- Rizzolatti, G. 2005. The mirror neuron system and imitation. In *Perspectives on imitation. From neuroscience to social science*, Hurley, S. & Chater, N. (eds.), pp. 55-76. Cambridge, MA: MIT Press.
- Rumiati, R. I., Carmo, J. C. & Corradi-Dell'Acqua, C. 2009. Neuropsychological perspectives on the mechanisms of imitation. *Phil. Trans. R. Soc. B*.
- Sagerson, A.L., George, D.N. & Honey, R.C. 2005. Imitative learning of stimulus–response and response–outcome associations in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, 31, 289-300.
- Schiel, N. & Huber, L. 2006. Social influences on the development of foraging behavior in free-living common marmosets (*Callithrix jacchus*). *American Journal of Primatology* 68, 1-11.
- Schlosser, Gerhard and Wagner, Günter (eds) 2004. *Modularity in Development and Evolution*. Chicago: University of Chicago Press.
- Schuster, S., Wohl, S., Griebisch, M. & Klostermeier, I. 2006. Animal cognition: How archer fish learn to down rapidly moving targets. *Current Biology*, 16, 378-383.
- Sterelny, K. 2001. Niche Construction, Developmental Systems and the Extended Replicator. In Gray, Russell D, Griffiths, Paul E. and Oyama, Susan (eds.) *Cycles of Contingency*. Boston, MA: MIT Press pp. 333-50.
- Sterelny, K. 2004. Symbiosis, evolvability and modularity. In Schlosser, Gerhard and Wagner, Günter (eds.) *Modularity in Development and Evolution*. Chicago: University of Chicago Press.
- Sterelny, K. 2006. The Evolution and Evolvability of Culture. *Mind & Language*, 21, 137-65.

- Sterelny, K. forthcoming. Evolvability Reconsidered. In *The Major Transitions Reconsidered*, B. Calcott & K. Sterelny (eds.), to appear 2010. MIT Press. Paper delivered at *Major Transitions Revisited*, a workshop of the *Konrad Lorenz Institute for Evolution and Cognition Research, Altenberg*, 19-22 July 2007 and available online from www.victoria.ac.nz (accessed 16 October 2008).
- Tennie, C., Call, J. & Tomasello, M. 2006. Push or Pull: Imitation vs. Emulation in Great Apes and Human Children. *Ethology* 112, 1159-1169.
- Tennie, C., Call, J. & Tomasello, M. 2009. Ratcheting up the ratchet effect: On the evolution of cumulative culture. *Phil. Trans. R. Soc. B*.
- Tomasello, M. & Carpenter, M. 2005 Intention reading and imitative learning. In S. Hurley & N. Chater (eds.) *Perspectives on Imitation: From Mirror Neurons to Memes*. MIT Press.
- Tomasello, M. 1990 'Cultural transmission in the tool use and communicatory signalling of chimpanzees?' In Parker ST, Gibson K (eds) *"Language" and intelligence in monkeys and apes: Comparative developmental perspectives*. Cambridge, UK, Cambridge University Press, pp. 274-311.
- Tomasello, M. 1999. *The cultural origins of human cognition*. Harvard University Press.
- Tomasello, M., Kruger, A. & Ratner, H. 1993 Cultural learning. *Behavioral and Brain Sciences* 16:495–552.
- Van Baaren, R., Janssen, L., Chartrand, T. L. & Dijksterhuis, A. 2009. Where is the love? The social aspects of mimicry. *Phil. Trans. R. Soc. B*.
- Virányi, Z. & Range, R. (in press) Selective imitation in dogs requires demonstration in ostensive-communicative context. *Applied Animal Behaviour*
- Voelkl, B. & Huber, L. 2000 True imitation in marmosets. *Animal Behaviour* 60:195–202.
- Voelkl, B. & Huber, L. 2007. Imitation as faithful copying of a novel technique in marmoset monkeys. *PLoS ONE*, 2(7), e611.
- Vogt, S. Buccino, G. Wohlschlager, A.M. Canessa, N. Shah, N.J. Zilles, K. Eickhoff, S.B. Freund, H.J. Rizzolatti, G. & Fink, G.R. 2007 Prefrontal involvement in imitation learning of hand actions: effects of practice and expertise. *Neuroimage*, 37, 1371-1383.
- Want SC, Harris PL 2001. Learning from other peoples' mistakes: causal understanding in learning to use a tool. *Child Dev* 72:431–443
- Whiten, A., McGuigan, N., Marshall-Pescini, S. & Hopper, L. M. 2009. Emulation, imitation, overimitation and the scope of culture for child and chimpanzee. *Phil. Trans. R. Soc. B*.
- Wood, J. N. & Hauser, M. D. 2008 Action comprehension in non-human primates: motor simulation or inferential reasoning? *Trends in Cognitive Sciences* 12, 461-46