
Biology and Philosophy **19:** 377–396, 2004.

Extended Phenotype – But Not *Too* Extended. A Reply to Laland, Turner and Jablonka

RICHARD DAWKINS
University Museum of Natural History, University of Oxford, UK

I am grateful to the three commentators for their thoughtful and penetrating remarks, and to the Editor for commissioning them. All three have forced me to think, re-opening neural pathways that had suffered neglect as I turned to other things in the years since *The Extended Phenotype* (henceforth *EP*) was published. Their essays raise so many interesting points, it would take another book to reply to them properly. Instead, on the basis that it is better to say a few things thoroughly than lots sketchily, I shall concentrate on what I take to be each author's central argument.

J. Scott Turner and Kevin Laland both, in their different ways, want to go further than me in extending the phenotype. Or so they see it. I am not so sure that further is the right word. Progress implies movement in a useful direction, whereas their extensions – of the organism, and into niche creation – occasionally reminded me of Stephen Leacock's knight who jumped on his horse and galloped off in all directions. I don't intend that flippantly or disrespectfully. The relevant point about the extended phenotype is that it is a *disciplined* extension. There are lots of other tempting 'extensions', which sound similar but take us off in misleading directions. I have always fought shy of misapplying the phrase to a profligate range of apparently plausible extensions.

To take a more extreme example than these commentators consider, when I am asked by lay people (as I frequently am) whether buildings count as extended phenotypes, I answer no, on the grounds that the success or failure of buildings does not affect the frequency of architects' genes in the gene pool. Extended phenotypes are worthy of the name only if they are candidate adaptations for the benefit of alleles responsible for variations in them. I might admit the theoretical possibility of generalising to other kinds of replicators such as memes (or something 'epigenetic' that Eva Jablonka might be able to explain but I wouldn't), in which case my 'no' answer might be softened. But it is enough of a problem already, getting my more hard-headed scientific colleagues to accept the extended phenotype, without arousing their active hostility by mentioning memes (which many see as simplistic) or 'epigenetic

inheritance systems' (which some might write off as obscurantist). I shall return to the important point, which I enthusiastically accept, that replicators do not have to be made of DNA in order for the logic of Darwinism to work.

Laland speaks, I suspect, for all three authors when he espouses cyclical causation. He quotes me as saying

> There are causal arrows leading from genes to body. But there is no causal arrow leading from body to genes.

Laland, who disagrees, generously wants to absolve me from responsibility for this, saying that he is quoting out of context. But I am happy to stand by it. 'Cyclical causation' leaves me cold. I must, however, make very clear that I mean causation statistically. Experimentally induced changes in bodies are never correlated with changes in genes, but changes in genes (mutations) are sometimes correlated with changes in bodies (and all evolution is the consequence). Of course most mutations occur naturally rather than experimentally, but (because corrrelation can't establish causation) I need to focus on 'experimentally induced' in order to pin down the direction of the causal arrow. It is in this statistical sense that development's arrow goes only one way. Attempts to argue for a reverse arrow recur through the history of biology, and always fail except in unimportant special-pleading senses.

Sterelny, Smith and Dickerson (1996), follow Griffiths and Gray in saying "Most acorns rot, so acorn genomes correlate better with rotting than with growth". But this is dead wrong. It misunderstands the very meaning of correlation which is, after all, a statistical technical term. Admitting that most genomes rot, the relevant question is whether *such variation as there may be* in acorn genomes correlates with *such variation as there may be* in tendency to rot. It probably does, but that isn't the point. The point is that the question of covariance is the right question to ask. Sterelny and Kitcher (1988) in their excellent paper on 'The Return of the Gene' are very clear on the matter. Think variation. Variation, variation, variation. Heritable variation; covariation between phenotype as dependent variable, and putative replicator as independent variable. This has been my *leitmotif* as I read all three commentators, and it will be my refrain throughout my reply.

Laland's main contribution to our debate is 'niche construction'. The problem I have with niche construction is that it confuses two very different impacts that organisms might have on their environments. As Sterelny (2000) put it,

> Some of these impacts are mere effects; they are byproducts of the organisms's way of life. But sometimes we should see the impact of organism on environment as the organism *engineering* its own environment: the environment is altered in ways that are adaptive for the engineering organism.

Niche construction is a suitable name only for the second of these two (and it is a special case of the extended phenotype). There is a temptation, which I regard as little short of pernicious, to invoke it for the first (byproducts) as well. Let's call the first type by the more neutral term, 'niche changing', with none of the adaptive implications of niche construction or – for that matter – of the extended phenotype.

A beaver dam, and the lake it creates, are true extended phenotypes insofar as they are adaptations for the benefit of replicators (presumably alleles but conceivably something else) that statistically have a causal influence on their construction. What crucially matters (here's the *leitmotif* again) is that *variations* in replicators have a causal link to *variations* in dams such that, over generations, replicators associated with good dams survive in the replicator pool at the expense of rival replicators associated with bad dams. Note what a stringent requirement this is. Although it is not necessary that we should already have evidence for the replicator-phenotype covariance, extended phenotype language commits us to a can only have come about through replicator-phenotype covariance. The beaver's dam is as much an adaptation as the beaver's tail. In neither case have we done the necessary research to show that it results from gene selection. In both, we have strong plausibility grounds to think it is. The same is not true – would not even be claimed by Laland and his colleagues – of most of their proposed examples of niche construction.

See how different is the 'pernicious' sense of niche construction, the byproduct that I'd prefer to sideline as 'niche changing'. Here, the dam alters the environment of the future, in some way that impinges on the life and wellbeing of beavers in general, and probably others too. Not particularly the welfare of the beavers that built the dam, not even of their children or grandchildren. The dam is good for beaverdom, and more. Beavers, frogs, fishes and marsh marigolds all benefit from a beaver-induced flooding of their niche. This is too loose and vague to count as a true extended phenotype, or as true niche construction. The deciding question is 'Who benefits?' And the reason it matters is that we have a Darwinian explanation of the dam only if dam-friendly alleles of the dam builders themselves benefit at the expense of alternative alleles.

I have no wish to downplay the importance of niche changing. It is a fair description of many important biological events, ranging from the irreversible oxygenation of Earth's early atmosphere by green bacteria and now by plants, to the greening of deserts by ecological successions of plants climaxing in dense forest communities, and including Scott Turner's *heuweltjies* (a fascinating example, of which I had been ignorant).

Most biologists would accept that the beaver dam is an evolved adaptation for the benefit of the genes of the responsible beaver. It would be a bold scientist (James Lovelock, perhaps) who would suggest that the oxygenation of the atmosphere by plants is an adaptation for the benefit of something. The oxygenation of the atmosphere is a hugely important niche change, and woe betide any creature, including any plant, that fails to adapt to it. But the presence of oxygen is nobody's adaptation (or at least, you'll have your work cut out if you want to argue that it is). It is a byproduct of plant biochemistry to which all living creatures, plants included, must adapt. Beaver dams may or may not benefit other beavers, or fishes or water beetles or pondweeds, but such diffuse and unfocused benefits cannot explain why they are there. The only benefits that can be adduced in Darwinian explanation of dams are benefits to the alleles (or other responsible replicators) of the particular beavers that build them. Otherwise, natural selection could not have shaped their evolution. Long-term consequences of niche changing are interesting and important, but they do not provide a Darwinian explanation for why animals change their niches.

Laland pays some lip service to this point when he speaks of ecological inheritance, and says that it resembles the inheritance of territory or property. Local exclusiveness is indeed a vital ingredient of true niche construction. As long as beavers have a high chance passing their lake on to their own grandchildren rather than to somebody else's grandchildren, there is at least a chance of making a workable Darwinian model of niche construction. But the rhetoric of niche construction neglects to follow the lip service, and we are left believing it to be a larger and a grander theory than it really is. Those aspects of niche construction theory that work are already included within extended phenotype theory. Those aspects that don't fit within existing extended phenotype theory don't work.

Don't work as Darwinian adaptations, that is. They can still be interesting in other ways. Earthworms are mentioned by both Laland and Turner, and Laland's splendid 'accessory kidneys' are a gift to Turner and his 'extended organism'. Earthworms radically change the environment in which they, and all other soil organisms including – significantly – rival earthworms live. Again, we certainly have niche alteration but, please, not niche construction until a lot more work has been done to establish this onerous claim.

Ecological succession is a form of niche changing – not niche construction – which follows a repeatable, regular pattern. A desert is colonised by weeds, which then change conditions sufficiently to allow the subsequent invasion by an orderly succession of plants and animals, each wave altering niches in ways that favour the next wave, culminating in a climax forest. But, important and repeatable as ecological succession is, it is not a Darwinian

adaptation on the part of prior member of the succession on behalf of later members. Rather, natural selection within the gene pools of later members of the succession favours those individuals that take advantage of the conditions inadvertently set up by earlier members. The climax forest is a consequence of colonisation by weeds decades or even centuries earlier. The forest is not an extended phenotype of the weeds' genes, nor is it helpful or illuminating to call it a niche constructed by the weeds. The same can be said of the repeatably regular pattern of development of coral reefs, in which generations of polyps build literally on the environment provided by centuries of dead predecessors, and form the foundation – literally and metaphorically – for the marine equivalent of a climax forest community.

Moving on from ecological succession to longer-term processes that look a bit like niche construction, coevolutionary arms races are the outstanding example (Dawkins and Krebs 1979). Predators impose new selection pressures on prey, which respond in evolutionary time such that future generations of prey impose changed selection pressures on future generations of predators. The coevolutionary positive feedback spirals that result are responsible for the most advanced and stunning illusions of design that the natural world has to offer. Again this is a case of animals changing future niches, and changing them in fascinating ways, but again it isn't niche construction, and no helpful purpose is served by lumping it with beaver dams or ecological succession. Understanding requires us to respect clear distinctions.

I don't denigrate niche changing as an important biological phenomenon. But it is not the same thing as true niche construction. Nothing but confusion will result from treating one as a continuation of the other. Since this seems to be a misunderstanding that is eagerly waiting to happen, niche construction is a phrase that should be abandoned forthwith.

That's all I want to say about niche construction. Now, the extended organism, which is J Scott Turner's main contribution to our debate. Turner, like Laland, is aware of the distinction between benefit to the agents responsible for a phenotype, and benefit to the world at large. But, as with Laland, his enthusiasm is in danger of misleading others into forgetting the distinction.

Turner, like Jablonka as we shall see, thinks I am too much of a genetic triumphalist. For the moment I shall leave that on one side while I focus on the wonderful examples of would-be extended organisms that Turner offers us from his own work on termites. Yes, the *Macrotermes* nest, with its underground living and brooding chambers and its overground ventilation apparatus, has many of the attributes of an organism. And yes, it is an intriguing conceit that the fungi are cultivating the termites, rather

than the other way around. Indeed, I said something pretty similar about cellulose-digesting gut microbes in *EP* (p. 208):

> Could the evolution of eusociality in the Isoptera be explained as an adaptation of the microscopic symbionts rather than of the termites themselves?

Once again, note that the extended phenotype is a *disciplined* hypothesis. Speculative as my suggestion was, it was a very specific and tightly limited speculation. Implicitly it postulated *alleles* in microorganisms (or fungi to take in Turner's hypothesis) which *vary* in their effects upon termite social behaviour (or mounds). The fact that there is no actual evidence for either speculation need not worry us at this stage. The point is to be precise about the genetic nature of the speculation. Adaptive hypotheses, however wild and speculative, must not be vaguely Panglossian but precisely limited to specified alleles (or other replicators) which *vary* and which exert a *causal* influence on *variation* in the phenotype of interest.

Let's apply these rigorous standards to the hypothesis that a termite mound is an extended organism. We shall conclude in favour, but it is important to make the case properly, in what I have called a disciplined manner. We shall take for granted the physiological, homeostatic and thermodynamic arguments put by Turner – not because they are unimportant but because he has made them so well. Instead, we concentrate on the genetics (using genes to stand for other conceivable replicators). Mound morphology is sure to be influenced by a number of genes, acting via mound embryology which, in the terms of our discussion, is another name for termite behaviour. These genes are to be found in the cells of many different organisms (using 'organism' in the conventional, non-extended sense). They include genes in the cell nuclei of numerous individual worker termites. They also might include genes in fungi, genes in gut symbionts, and genes in mitochondria or other cytoplasmic elements in the cells of termites, fungi or gut symbionts. So, we potentially have a rich pandemonium of genetic inputs to our mound phenotype, coming at it from as many as three kingdoms.

For my money, the analogy of mound with organism stands up well. The fact that we have a heterogeneously sourced genetic input to the embryology of the phenotype doesn't matter. Lots of genes affect each aspect of my bodily phenotype, including, for all I know, mitochondrial genes. My 'own' nuclear genes tug me in more or less different directions, and my phenotype is some sort of quantitative polygenic compromise. So that is not a difference that might stop the mound being an organism. What, then, is the prime characteristic of an organism? It is that, at least to a quantitatively appreciable extent, all its genes are passed on to the next generation together, in a small 'bottlenecked' propagule. The rationale for this is given in *EP*,

especially Chapter 12, 'Host phenotypes of parasite genes' and Chapter 14, 'Rediscovering the Organism', and I shall not repeat it here. Instead, let's go straight to the termite mound to see how well it holds up. Pretty well. Each new nest is founded by a single queen (or king and queen) who then, with a lot of luck, produces a colony of workers who build the mound. The founding genetic injection is, by the standards of a million-strong termite colony, an impressively small bottleneck. The same is, at least quantitatively, true of the gut symbionts with which all termites in the new nest are infected by anal licking, ultimately from the queen – the bottleneck. And the same is quantitatively true of the fungus, which is carefully transported, as a small inoculum, by the founding queen from her natal nest. All the genes that pass from a parent mound to a daughter mound do so in a small, shared package. By the bottleneck criterion, the termite mound passes muster as an extended organism, even though it is the phenotype of a teeming mass of genes sitting in many thousands of workers.

I won't miss an opportunity to emphasise (though again I shall not repeat the full argument from *EP)* that every organism (conventionally defined) is already a symbiotically cooperating union of its 'own' genes. What draws them, in a Darwinian sense, to cooperate is again 'bottlenecking': a shared statistical expectation of the future. This shared expectation follows directly from the method of reproduction, according to which all of an organism's 'own' nuclear genes, and its cytoplasmic genes for good measure, pass to the next generation in a shared propagule. To the extent that this is true of parasite genes (for example bacteria that travel inside the host's egg), to that very same extent aggressive parasitism will give way in evolutionary time to amicable and cooperative symbiosis. The parasite genes and the host genes see eye to eye on what is an optimum host phenotype. Both 'want' a host phenotype that survives and reproduces. But to the extent that parasite genes pass to their own next generation via some sideways route which is not shared with those of the host genes, to that same extent the parasite will tend to be vicious and dangerous. In such cases, the optimum phenotype from the parasite genes' point of view may well be dead – perhaps having burst in a cloud parasite spores. All our 'own' genes are mutually parasitic, but they are amicably cooperative parasites because their shared route to the future in every generation leads them to 'see eye to eye' on the optimal phenotype.

A termite mound, then, is a good extended organism. A *heuweltjie*, by my reading of Turner's description, is not. It is more like a forest or a coral reef. The genes that contribute to the putative heuweltjie phenotype don't cooperate, because they do not have a statistical expectation of sharing a propagule from the present heuweltjie to the next. Only the contingent centred around the termite genes has that shared expectation. The rest will

join the club later, from different sources, which means that, in the sense I am expounding, it is not a club. Because termite genes, with their fellow travellers, bottleneck their way from mound to mound, we can reasonably think about a form of natural selection which chooses among mounds as extended pheontypes, with adaptive consequences in an evolutionary succession of progressively improving mounds. The same will not be true of a putative natural selection of heuweltjies. Hence my statement that a heuweltjie is not a good extended organism. As in the case of Laland and his niche construction, my request to Turner is to be critical and disciplined with his notion of the extended organism. In his case, apply the bottleneck test.

At this point, I have to pick Turner up on his outrageous statement that "most would agree that the central dogma is essentially dead." It is important to do so because I suspect that many people (perhaps including present commentators who are drawn to 'cyclical causation' and similar notions) have a kind of poetic bias against Francis Crick's central dogma. This may be partly, and understandably, because of Crick's unfortunate choice of the word 'dogma', as opposed to, say, 'hypothesis' or 'theorem'. Crick's own explanation is endearing, as recounted in an interview with Horace Judson (1979). Judson asked him why he had used the word dogma and Crick replied that, because of his religious upbringing, he thought a dogma was a word for something "for which there was *no reasonable evidence.*" He had since been told by Jacques Monod that it means "something which a true believer *cannot doubt*." "You see" Crick roared with laughter as he confided in Judson, "I just didn't *know* what dogma *meant*!" Actually, the Oxford English Dictionary could be used to support either meaning.

The central dogma has been expressed in three versions, whose differences can admittedly lead to confusion: –

1. "Once information has passed into protein, *it cannot get out again*." This is Francis Crick's original wording, at the 1957 meeting of the Society for Experimental Biology and it is, as one would expect, completely clear. Note the prescience with which, long before reverse transcription was discovered, Crick in effect anticipated its irrelevance to his dogma.

> . . . the transfer of information from nucleic acid to nucleic acid, or from nucleic acid to protein may be possible, but transfer from protein to protein, or from protein to nucleic acid is impossible. Information means here the *precise* determination of sequence, either of bases in the nucleic acid or of amino acid residues in the protein (Crick 1957, quoted in Judson 1979).

In this version the central dogma has never been violated and my bet is that it never will. The genetic code, whereby nucleotide sequences are translated into amino acid sequences, is irreversible.

2. "DNA makes RNA makes protein." This sounds pithy and clever, but it is too pithy and not clever enough. Unfortunately, it is the textbook version that students learn. But it is a summary of research findings, not a theoretical principle like Crick's 'dogma'. It is technically violated by reverse transcription but, as we shall see, the fact is trivial and misses the whole point of the dogma.

3. "Embryology is irreversible." This third version is another way of saying that acquired characteristics are not inherited. It is not particularly molecular in its domain, and it owes more to Weismann than Crick, but it is interesting in being closer to 1 (theoretical principle) than to 2 (summary of known facts, now trivially violated). This version, too, has never been convincingly violated, despite many attempts.

Version 2 is disproved by reverse transcription, but this is a violation of the dogma only if we think the dogma was ever intended to apply to *both* stages of the process: transcription (DNA to RNA) as well as translation (polynucleotide to protein). But such a dogma would have been foolhardy, lacking any basis in theory, and it was explicitly excluded by Crick, with the prescience I have already praised ("the transfer of information from nucleic acid to nucleic acid"). The only ground Crick, or anybody else, ever had for confidence in his central dogma is that the information in a protein is inaccessibly buried inside the knot which the protein ties in itself – *must* tie if it is to perform its role as an enzyme. DNA is not knotted, which is why it is a lousy enzyme but very good at getting its information transcribed (into RNA, as it happens). RNA can tie itself in a kind of knot, enough to secure some sort of enzyme function (which is why some people favour it for a primitive enzyme role as well as a primitive replicator role in theories of the origin of life). But RNA doesn't always get knotted, which is why it is good at getting its information read and translated into protein. It therefore should have surprised nobody that RNA's information can sometimes be reverse transcribed back into DNA. Why should it not, given that it maps DNA information one to one, and it is necessarily accessible otherwise it could never be translated into protein? If Version 1, on the other hand, were ever disproved (which I doubt) it would only be by reverse translation of a structural protein like collagen or silk – un-knotted and therefore incapable of functioning as an enzyme.

Prions, contrary to widespread misunderstanding, do not violate Crick's careful formulation of his dogma. They are replicators after a fashion, in that their alternative conformations are infectious. But the amino acid sequence of a prion is not reverse-translated into the appropriate codon sequence of a polynucleotide (look again at Crick's prudent wording). Nor is the sequence of amino acids copied by another polypeptide chain. All that happens is that, of the alternative three dimensional conformations of a given polypeptide

sequence, one can, by its proximity, convert another existing molecule to its own shape. Nobody has ever realistically suggested that the amino acid sequence of a prion comes from any source other than DNA.

Dogma 3, the Weismannian or anti-Lamarckian pre-molecular version, is of course, the subject of old arguments, and I shall not get into all that here because it is not what Turner was talking about anyway. I'll just point out that it is a sort of whole-organism version of Crick's molecular dogma, and it is based on a similar theoretical principle. Just as amino acid sequences are inaccessibly buried in a protein, so the genetic instructions that program the development of a body are inaccessibly buried in the body itself. This is not just an empirical fact, which could be disproved at any moment by a Lamarckian finding such as a non-fraudulent case of the midwife toad. It follows from the deeper principle that embryology is not preformationistic. This is the old point about blueprints being reversible, recipes not (*EP* p. 174: 'The Poverty of Preformationism'). You can reconstruct a blueprint from a house, but not a recipe from a cake, an image that I inadvertently borrowed from my friend Patrick Bateson. Bateson's name, by the way, reminds me of my astonishment that Eva Jablonka is not the only author to sympathize with his superficially amusing but deeply misleading suggestion that a gene is a nest's way of making another nest. I shall return to this at the end.

To conclude on the central dogma, that limited part which is essentially dead (RNA cannot be reverse transcribed) should never have been born in the first place. That part of the dogma which deserved to be enunciated (and actually was enunciated by Crick) is most certainly not dead, not essentially dead, not even the tiniest bit ailing.

Let me now turn to Eva Jablonka. She, like the other two commentators, has read *EP* with flattering attention, and I am grateful for her, and their, clear disavowal of several potential misunderstandings. Genetic determinism does not follow from gene selectionism. Nor does naïve adaptationism. She is also admirably clear that "when geneticists talk about 'genes for', they are talking about genetic *differences* that make a *difference* to the phenotype." I suspect that she, like Turner, wants to have nothing to do with what he calls 'genetic triumphalism'. I agree, insofar as the 'gene' role in Darwinian models does not have to be played by DNA. If I am a triumphalist, it is a replicator triumphalist. I am happy to go along with what Sterelny (2000) has dubbed 'the extended replicator'. Indeed, I was at some pains to extend the replicator myself, in *EP*, listing several of the alternative replicators mentioned by today's three commentators: paramecium cilia, and memes, for instance. I would certainly have included prions if they had been discovered then. Jablonka is right when she says:

> Following the fortunes of heritably variable phenotypic traits in populations is common practice in evolutionary biology. We measure the genetic component of the variance in a trait in a population; models of phenotypic evolution are regularly constructed (e.g. most game theoretical models); and paleontological data, which is mostly based on morphological traits, is an accepted source of insights about evolution. Since for an entity to count as a 'fitness bearer' – a unit of adaptive evolution – it has to show (frequent) heritable variation in fitness, variant phenotypic traits are much better candidates than genes for this role.

I agree. But Jablonka should not be *surprised* that I agree. I devoted a chapter, 'Selfish Wasp or Selfish Strategy' to developing precisely the notion that a Darwinian replicator does not have to be specified as DNA, but can be a Maynard Smithian 'strategy' defined in a minimalist 'like begets like' fashion. Presumably DNA is involved in practice, but it is not a specified part of the reasoning. Jablonka's 'heritably varying phenotypic trait' is close to Williams's classic definition of the 'gene', which was the same sense in which I later called it 'selfish'.

> If there is an ultimate indivisible fragment it is, by definition, 'the gene' that is treated in the abstract definitions of population genetics (Williams 1966).

The Williams gene is only incidentally made of DNA. He later (1992) called the generalised version (what I would call a replicator) a codex, adding, "A gene is not a DNA molecule; it is the transcribable information coded by the molecule." I agree with Sterelny (and I am sure Williams would too):

> My own view is that DNA-based transmission of similarity *is* of fundamental significance. But that is not built into the structure of the theory.

Quite so. If Jablonka manages to convince the scientific community that some sort of complex feedback system of developmental cycles constitutes a true replicator, over and above its DNA content, I would be happy to embrace it. But, for the third time and at the risk of seeming pedantic, I insist on tight discipline. The criterion for recognizing a true replicator for a Darwinian model is a rigorous one. The putative replicators must vary in an open-ended way; the variants must exert phenotypic effects that influence their own survival; the variants must breed true and with high fidelity such that, when natural selection chooses one rather than its alternative, the impact persists through an indefinitely large number of generations (more precisely, survives at a high enough rate to keep pace with mutational degredation). If there is something other than DNA that meets these criteria, let us by all means include it, with enthusiasm, in our Darwinian models. But it really

must meet those criteria. Sterelny (2000) has a similar list, which he calls Hoyle Conditions because he imagines tailoring a form of life to colonise an empty world from outer space.

I am interested in the possibility that Jablonka really has a good new candidate for a true replicator, but I have to say that the use of the word 'epigenetic' makes for an unpropitious start – associated as it (no doubt unfairly) has become with obscurantism among biologists.[1] Epigenetic should be reserved for its true meaning as a historical school of embryology, hard to define except as a nebulous antonym of preformationist – which is not nebulous, is easy to define,and clearly wrong. If you want to propose an alternative replicator, extragenetic, paragenetic or quasigenetic might all be happier choices than epigenetic – not on grounds of strict etymology but because epigenetic is weighed down by inappropriate historical associations. A meme might be a quasigenetic replicator. A prion is perhaps a paragenetic replicator. Both fall down on some, but not all, of my criteria. Prions fail on the criterion of open-ended variation: the repertoire of variants for a given prion is limited to two. And memes – no, for heaven's sake don't let's get into memes now: I'll save them up to make a more worthwhile point, in a moment.

Jablonka's use of Waddington's canalization is potentially interesting (Waddington, numerous references, e.g. 1977). This isn't quite how she puts it, but canalization could play a 'self-normalizing' role. Let me explain self-normalizing, using memes in the way they are perhaps best used – by analogy. When I was a small boy at boarding school, we had to take turns in saying a goodnight prayer, kneeling up on the ends of our beds with our hands together. I can now reconstruct that the original prayer must have been that popular Evensong Collect, "Lighten our darkness, we beseech Thee O Lord, and by Thy great mercy defend us from all the perils and dangers of this night. . . ." But we only ever heard it said by each other, and none of us had a clue what most of the words meant. By the time I arrived at the school, the first line had become – and I inherited it, garbled it further, and passed it on – something like this: "Lutnar darkny sweep seech Theo Lord. . . ."

The childhood game of Chinese Whispers (American children call it Telephone) is a good model for such degradation of messages handed down over memetic 'generations'. Twenty (say) children are lined up, and a message whispered into the ear of the first. She repeats it in the ear of the second, and it passes on down the line until the twentieth child finally speaks it aloud to the assembled company – who are amused or dumbfounded at how much it has degenerated when compared with the original. As experimental memeticists we might find Chinese Whispers a useful test bed. We would compare the fidelity of various classes of message. Compare, for example, a message in a

language unknown to the children with a message they can understand. My school prayer was a sort of inadvertent running of this experiment.

When a child listens to a message and passes it on, there are two ways he can do it, one being 'normalizing' and the other not. The non-normalizing method is to imitate the sounds, phoneme by phoneme. That is approximately what the members of my dormitory were doing with 'Lighten our darkness'. The normalizing method is to treat the message, not as a set of phonemes to be imitated, but as a set of words to be looked up in a mental dictionary and then re-rendered in the child's own accents.

Cananalizing is not synonymous with digitizing but it has a similar effect. Digital codes such as DNA are protected from continuously distributed degradation, while at the same time becoming vulnerable to discrete error. Both are potential normalizing agents. Normalization is even more clearly illustrated by another meme which spread as an epidemic or craze at my father's school, and with which I re-infected the same school when I went there 26 years later. It consisted of the instructions for making an origami Chinese Junk.

> It was a remarkable feat of artificial embryology, passing through a distinctive series of intermediate stages: catamaran with two hulls, cupboard with doors, picture in a frame, and finally the junk itself, fully seaworthy or at least bathworthy, complete with deep hold, and two flat decks each surmounted by a large, square-rigged sail (Dawkins 1999).

One could imagine a version of Chinese Whispers in which what passed down the line was a hands-on demonstration of this particular skill. Unlike a drawing of a junk, which would degrade horribly down the line, the origami instructions have a good chance of making it, intact, to the twentieth child, for the reason that they are self-normalising. Here are the first five instructions for making a Chinese junk.

1. Take a square sheet of paper and fold all four corners exactly into the middle.
2. Take the reduced square so formed, and fold one side into the middle.
3. Fold the opposite side into the middle, symmetrically.
4. In the same way, take the rectangle so formed, and fold its two ends into the middle.
5. Take the small square so formed, and fold it backwards, exactly along the straight line where your last two folds met.

And so on, through 20 or 30 instructions of this kind. These instructions, though I would not wish to call them digital, are potentially of very high fidelity, just as if they were digital. This is because they all

> make reference to idealised tasks like 'fold the four corners exactly into the middle'. If the paper is not exactly square, or if a child folds ineptly so that, say, the first corner overshoots the middle and the fourth corner undershoots it, the junk that results will be inelegant. But the next child in the line will not copy the error, for she will assume that her instructor *intended* to fold all four corners into the exact centre of a perfect square. The instructions are self-normalising. The code is error correcting (Dawkins loc. cit.)

I hope the analogy to Waddingtonian canalization, and Jablonka's usage of it, is becoming clearer. A canalized embryology is resistant to change. Resistant, at least, to small, continuously distributed change, although large changes can kick Waddington's rolling ball out of the groove into a neighbouring one. Even this subtlety is well covered by the origami analogy:

> I haven't done it, but I will make the following confident prediction, assuming that we run the experiment many times on different groups of 20 children. In several of the experiments, a child somewhere along the line will forget some crucial step in the skill taught him by the previous child, and the line of phenotypes will suffer an abrupt macromutation which will presumably then be copied to the end of the line, or until another discrete mistake is made. The end result of such mutated lines will not bear any resemblance to a Chinese junk at all. But in a good number of experiments the skill will correctly pass all along the line, and the 20th junk will be no worse and no better, on average, than the first junk. If we then lay the 20 junks out in order, some will be more perfect than others, but imperfections will not be copied on down the line. If the fifth child is hamfisted and makes a clumsily asymmetrical or floppy junk, his quantitative errors will be corrected if the sixth child happens to be more dexterous (Dawkins loc. cit.).

The twenty junks will not exhibit a progressive deterioration, as they would in a game in which each child was asked to imitate a *drawing* done by the preceding child. In the light of this memetic analogy, I take it that Jablonka is proposing that canalization increases the *fidelity* of her putative replicator by resisting change, at least up to the point where the Waddingtonian 'rolling ball' is kicked into a neighbouring channel. If I am right, it is a worthwhile suggestion, which needs to be worked out more thoroughly. My hunch is that it will come to nothing, but it is interesting, nevertheless. It could have the makings of a new kind of replicator theory.

I said that I'd return to Pat Bateson and The Selfish Nest. Jablonka sympathizes with Bateson's opinion that the developmental cause-effect relationship between genes and phenotypes is circular, and that a gene can

therefore be thought of as a nest's way of making another nest. Sterelny, Smith and Dickerson (1996) go so far as to say, "Bateson was right"! No, Bateson was not right, he wasn't even close to being right, for the reasons I gave in *EP*, reasons mentioned by Jablonka, and by Sterelny et al. but, to my bafflement, not accepted by them.

> Dawkins rejected this idea on the grounds that variation is not transmitted [the *leitmotif* again, RD]. Whatever the merits of The Selfish Nest as an evolutionary hypothesis, it cannot be rejected on those grounds. First, because Dawkins here appeals to the same criterion used to exclude asexual organisms as replicators; a criterion unsatisfactory on other grounds. Second, it is not in general true. Environmentally altered patterns in cilia are inherited through fission. . . . Variation in both nesting materials and nest siting can be transmitted (Sterelny, Smith and Dickerson 1996).

My grounds for excluding asexual organisms as replicators were, in my opinion, very satisfactory. I'll reply to what Sterelny et al. went on to say:

> Dawkins appealed to fidelity to argue that asexual organisms are not replicators [*EP* p. 97]. An aphid that loses one of its legs will still give birth to six-legged offspring. . . . This criterion backfires against genetic replication. Many changes in the germline genes are not passed on. The point of the proofreading and repair mechanisms is to avoid the transmission of changes. So if genes are replicators, some changes in replicators need not be passed on; those censored by the proofreading and repair mechanisms. But then we can see the production of a six-legged aphid from its eventually five-legged forebear as a triumph of the aphid's proof-reading and correction mechanism.

Nice try. Won't do. Certainly, not all genetic changes are passed on. But no gene selectionist ever said they were. The point is that some genetic changes are passed on (otherwise there could be no evolution) but *no* environmentally acquired changes are passed on (at least not with enough high fidelity to have a chance of surviving into the indefinite future). Or, if they are passed on, they are replicators by definition and that takes care of the second part of Sterelny et al.'s objection. If environmentally altered variations in patterns of cilia are inherited (as I was happy to admit in *EP*, p. 176–177) they are replicators by definition and therefore, for present purposes, honorary genes. Aphid clones are not replicators for precisely the reason that I originally gave.

Jablonka and the school of thought dubbed 'Developmental Systems Theorists' think that the complexity of embryonic development somehow detracts from the validity of the gene's eye view of Darwinism. But we must not allow complexity to become a euphemism for muddle. Gray (1992) in 'Death of the Gene: Developmental systems strike back' says:

> . . . genetic factors do not replicate themselves nor do they physically persist across generations [*of course* they don't, that is the point of Williams's 'codex', RD]. They are replicated as part of the *reproduction* of developmental systems. Remove some part of that developmental system and genetic replication may be changed or impaired. In this sense genes are no different from any other developmental interactant.

Oh yes they are. You may be sick of hearing my *leitmotif* but we are just going to have to play it one more time as a finale. It doesn't matter how complicated the developmental support structure, nor how utterly dependent DNA may be upon it, the central question remains: which elements of the Great Batesonian Nexus of development have the property that *variations* in them are replicated, with the type of fidelity that potentially carries them through an indefinitely large number of evolutionary generations? Genes certainly meet the criterion. If anything else does, let's hear it and, if the case is well made, let's by all means elect it into membership of the replicator club. But that is a separate issue. The complexity of development itself is an obscurantist red herring. Complexity is tamed by the statistics of variation. That, for heaven's sake, is why the analysis of variance was invented, and heritability is just a special case of the analysis of variance.

This should be our response to Jablonka too, and the other commentators to the extent that they invite it. We can clearly distinguish two kinds of objection to the gene's-eye-view of selection. There is the 'genes are not the only replicators' class of objection. Let's embrace that one with open arms in principle, even though we may have to bend over backwards to accommodate some pretty specious special pleading in practice. And there is the 'Dear oh dear, development is a terribly complicated nexus, isn't it?' style of objection. Don't embrace that one. Lance the boil of obfuscatory complexity with a laser scalpel. Or mutate the metaphor, and shine a laser beam of clear statistical reasoning on what really matters, which is transgenerational covariance.

Gray repeats his error with abandon. Just one more example, in case I still have failed to get the point across.

> Lots of fun could be had with these environmentalist inversions of the gene's eye view of evolution. For example, instead of the story of the selfish gene, imagine the story of the selfish oxygen. In the evolution of the earth's atmosphere oxygen was engaged in intense competition with other atmospheric gases. With the construction of green plants oxygen developed a vehicle for its efficient replication. Chlorophyll containing organisms were thus just oxygen's way of making more oxygen (Gray, loc. cit.).

I find it disturbing that anybody could be so misled as to see this as good satire, yet I have a horrible suspicion that more than one of our three

commentators would be tempted by it. If there were alternative versions of oxygen that *varied* in their talent for exploiting plants and passed on those talents to daughter oxygens, Gray would have a point. But there aren't. Oxygen is oxygen is oxygen. There is nothing there to select.

The quality of hi-fi variation is not something cheap and easy, possessed by Bateson's nests, Gray's oxygen and just about any other unit you could think of from the world of chemistry. On the contrary, it is a precious, rare, onerous, difficult talent, possessed by genes and computer viruses and a few other things – but *genuinely* few – every one of which needs rigorous defence before biologists of critical intelligence should accept it into their Darwinian models. If it were as easy as Gray jokes, the origin of life – which means the origin of self-replicated variation – would not be the major theoretical conundrum that it is.

Hi-fi variation is not some kind of arbitrary criterion, required for scripturally dogmatic reasons stemming from the teachings of Saint George Williams. It follows from first principles, the principles that tell us why any of this matters in the first place. We are interested in evolution by natural selection. In order for anything to evolve by natural selection, there has to be variation in something that is both potentially long lasting and causally powerful, so that there emerges a difference, on the evolutionary timescale, between the state of the world if one variant survives compared with the state of the world if an alternative variant survives. If neither variant survives more than a couple of generations anyway, we are not talking evolution at all. That is why hi fi variation matters and that is why Gray's oxygen joke, Bateson's nest joke and others of their kind are not funny. There may be backwards arrows in all sorts of other senses but, in the sense that specifically matters for Darwinian evolution, the causal arrow of biological development from genotype to phenotype really is a one-way arrow.

What should I say if invited to give my own 21-year retrospective on *The Extended Phenotype?* I think Laland and Jablonka are right that the gene's-eye-view – the part of the theory that I am not responsible for inventing – really has moved to the forefront of the minds of ethologists, behavioural ecologists, sociobiologists and other evolutionary biologists in the field. This is certainly gratifying. Moreover, the study of what some people call 'ultraselfish genes' or 'selfish genetic elements' has become a major growth industry.

But the part of the theory that is wholly my own, the extended phenotype itself, unfortunately cannot yet make the same claim. It lurks somewhere near the back of some biologists' minds, but not in the lobes that plan research in the field. Twenty-one years ago, I said that nobody had done a genetic study using animal artefacts as the phenotype. I think that is still true. I would admit

to disappointment, except that it invites the obvious retort: why don't you get out there and do it yourself, then? It is a fair point. I should. Maybe I will. Idleness is a poor excuse, and preoccupation with writing books only slightly better.

Meanwhile, let me conclude with an idle pipedream. It is the beautiful Indian summer of 2010, opening day of EPI, the Extended Phenotypics Institute in one of our great university cities. After the formal unveiling by a Nobel Prizewinning scientst (Royalty wasn't considered good enough), the guests are shown wonderingly around the new building. There are three wings: the Zoological Artefact Museum (ZAM), the laboratory of Parasite Extended Genetics (PEG), and the Centre for Action at a Distance (CAD).

The artefact museum is a zoological equivalent of Oxford's Pitt Rivers, which differs from other museums of human artefacts in that its specimens are grouped functionally instead of by region of origin. Instead of sections devoted to Polynesia, Africa, Asia and pre-Columbian America, the Pitt Rivers has sections devoted to fishing nets, to wind instruments, to boats, to butchering tools, to ornamental headdresses, all gathered together with their own kind regardless of their geographic provenance. EPI's museum has all the nests together, whether made by birds, insects, mammals or spiders; all the hunting nets in another case, whether made by spiders or caddis larvae; all the sexually alluring bowers in a third, and so on. Where possible, each specimen is housed next to human equivalents, and next to functionally analogous pieces of animal anatomy: lyre bird tails next to bower bird bowers, thermoregulatory heat-exchange organs next to termite mound chimneys, and so on. A central display case shows the comparative anatomy of bird nests, each one perched on its rightful branch of a phylogenetic tree: an expanded version of the tree drawn by Winkler and Sheldon (1993) for Swallows' nests.

All around the Museum are laboratories devoted to the genetics of animal artefacts. Some would say this is, strictly speaking, the genetics of their builders, but of course the ethos of EPI acknowledges no such distinction. Artefact genetics differs from conventional genetics in that the genes whose effects bear upon any one phenotype may come from different 'organisms'. Geneticists are used to handling such summations and epistatic interactions within 'organisms' under the heading of polygenes, and our extended geneticists are well versed in the mathematical theory of polygenic inheritance (Falconer 1981). Studies in the artificial selection and genetic manipulation of silkworm cocoons enjoy a generous grant from Japan, which also supports a major project on the genetics and polymer chemistry of other silk artefacts such as spider webs and caddis larva fishing nets. The artefact museum serves as the home base for field studies of the memetics of tool making and tool use in chimpanzees, sea otters, Galapagos woodpecker finches and others.

The other two wings can be imagined by analogy with the first, and by reference to Chapters 12 and 13 of *EP*. PEG is the most prosperously endowed part of the Institute, because of the medical importance of parasite genes expressing themselves in host phenotypes. As for CAD, its generous grant from agricultural funds is prompted by the hope that artificially synthesized pheromones could revolutionise pest control. But CAD's total remit embraces nothing less than the entire field of animal communication studies and, broader yet, networks of interaction in community ecology.

In all three wings, familiar phenomena are studied from an unfamiliar perspective: different angles on a Necker cube. Everyone knows that parasites manipulate their hosts. The extended geneticists of PEG differ only in that they study variations in host behaviour and morphology as phenotypes of parasite genes. Even more than their colleagues in the artefact museum, they are never far from their well-thumbed copy of Falconer's textbook, and they are as nearly as possible indifferent to their polygenes' 'organisms' of origin. The ethologists and zoosemioticists of CAD run the risk of being mistaken for Gaian eco-mystics, as they immerse themselves in the dawn chorus and call it extended embryology. But, like their colleagues in the other two wings of EPI, they pride themselves on the disciplined rigour of their theory. The motto carved over the main door of their Institute is a one-locus mutation of St Paul: "But the greatest of these is clarity."

Note

[1] I am reminded of a satirical version of Occam's Razor, which my group of Oxford graduate students mischievously attributed to a rival establishment: "Never be satisfied with a simple explanation if a more complex one is available". And that in turn reminds me to say that Laland has missed the irony in my apparent espousal of Bateson's "Great Nexus of complex causal factors interacting in development."

References

Bateson, P.: 1978, 'Book Review: *The Selfish Gene*', *Animal Behaviour* **26**, 316–318.

Butler, Stephen, Leacock: 1971, *Nonsense Novels*, Dover Pubns, London.

Crick, F.H.C: 1958, 'On Protein Synthesis', in *Symposium of the Society for Experimental Biology XII*, Academic Press, New York, p. 153.

Dawkins, R. and Krebs, J.R.: 1979, 'Arms Race Between and Within Species', *Proceedings of the Royal. Society of London B* **205**, 489–511.

Dawkins, Richard: 1982, *The Extended Phenotype*, WH Freeman, Oxford.

Dawkins, Richard: 1999, *Foreword to The Meme Machine by Susan Blackmore*, Oxford University Press, Oxford.

Falconer, D.S.: 1981, *Introduction to Quantitative Genetics*, Longman, London.

Gray, R.: 1992, 'Death of the Gene: Developmental Systems Strike Back', in Griffiths, *Trees of Life*, Kluwer Academic, Dordrecht, pp. 165–209.

Griffiths, P. and Gray, R.: 1994, 'Developmental Systems and Evolutionary Explanation', *Journal of Philosophy* **91**(6), 277–304.

Jablonka E.: 2004, 1From Replicators to Heritably Varying Phenotypic Traits: The Extended Phenotyper Revisited', *Biology and Phylosophy* **19**, 353–375.

Judson, Horace Freeland: 1979, *The Eighth Day of Creation*, Jonathan Cape Press, London.

Laland, K.: 2004, 'Extending the Extended Phenotype', *Biology and Phylosophy* **19**, 313–325.

Maynard Smith, John: 1982, *Evolution and the Theory of Games*, Cambridge University Press, Cambridge.

Sterelny, K. and Kitcher, P.: 1988, 'The Return of the Gene', *Journal of Philosophy* **85**, 339–361.

Sterelny, K., Smith, K. and Dickerson, M.: 1996, 'The Extended Replicator', *Biology and Philosophy* **11**, 377–403.

Turner, J.S.: 2004, 'Extended Phenotypes and Extended Organisms', *Biology and Phylosophy* **19**, 327–352.

Waddington, C H.: 1977, *Tools for Thought*, Jonathan Cape, London.

Williams George, C.: 1966, *Adaptation and Natural Selection*, Princeton University Press, Princeton.

Williams George, C.: 1992, *Natural Selection: Domains, Levels and Challenges*, Oxford University Press, Oxford.

Winkler, D.W. and Sheldon, F.H.: 1993, 'Evolution of Nest Construction in Swallows (Hirundinidae): A Molecular Phylogenetic Perspective', *Proceedings of the National Academy of Sciences* **90**, 5705–5707.