Overextension:

The Extended Mind and Arguments from Evolutionary Biology

Armin Schulz

Department of Philosophy, Logic, and Scientific Method

London School of Economics and Political Science

Houghton St

London WC2A 2AE

UK

a.w.schulz@lse.ac.uk

(0044) 753-105-3158

Abstract

I critically assess two widely cited evolutionary biological arguments for two versions of the 'Extended Mind Thesis' (EMT): namely, an argument appealing to Dawkins's 'Extended Phenotype Thesis' (EPT) and an argument appealing to 'Developmental Systems Theory' (DST). Specifically, I argue that, firstly, appealing to the EPT is not useful for supporting the EMT (in either version), as it is structured and motivated too differently from the latter to be able to corroborate or elucidate it. Secondly, I extend and defend Rupert's argument that DST also fails to support or elucidate the EMT (in either version) by showing that the considerations in favour of the former theory have no bearing on the truth of the latter. I conclude by noting that the relevance of this discussion goes beyond the debate surrounding the EMT, as it brings out some of the difficulties of introducing evolutionary biological considerations into debates in psychology and philosophy more generally.

Overextension:

The Extended Mind and Arguments from Evolutionary Biology

I. Introduction

The 'Extended Mind Thesis' (EMT) – the claim that minds or cognitive systems can extend into the environment – has spawned much discussion and controversy (Clark 1997, 2008; Clark & Chalmers 1998; Rupert 2004, 2009a; Wilson 2004; Menary 2007; Adams & Aizawa 2008; Shapiro 2011). Given this, it is (perhaps) unsurprising that a wide range of considerations has been brought to bear on this debate. Among this range of considerations are appeals to evolutionary theory. In this paper, I assess two such appeals to evolutionary biology in more detail: firstly, an appeal to the (alleged) fact that phenotypes can extend into the environment, and secondly, an appeal to the (alleged) fact that only entire developmental systems (or at least large parts of them) can bring about phenotypes and should be seen to be the objects of analysis in evolutionary biology. While these are not the only (evolutionary) biological considerations that could be and are appealed to in the debate surrounding the EMT, these two are noteworthy, in that they are (a) fairly widely cited, but (b) not yet much discussed *in detail*.[1]

The paper is structured as follows. In section II, I make the EMT more precise. In section III, I present the 'Extended Phenotype Thesis' and critically discuss the support it is said to provide to the EMT. In section IV, I do the same for 'Developmental Systems Theory'. I conclude in section V.

[1] For an assessment of a set of more specific, natural selection-based arguments for the EMT, see Shapiro (2010). For an assessment of the similarities (or lack thereof) between cognitive systems and various specific biological systems – such as digestive ones – see Adams & Aizawa (2001). The focus of the present paper is on more general and explicitly evolutionary biological arguments for the EMT, though.

II. Two Versions of the Extended Mind Thesis

In order to make the EMT more precise, it is necessary to begin by distinguishing two quite different versions of it: the EMT_1 and the EMT_2 (Rupert 2004; Shapiro 2004, 2011). The EMT_1 is best understood with reference to standard functionalist theories of mind.[2] It claims that it is possible to point to cases in which the particular network of causes and effects that (according to functionalism) makes up a given mental state is at least partially realised by something in the organism's external environment (Clark & Chalmers 1998; Sprevak 2009). More specifically, the EMT_1 can be formulated as follows:

(EMT_1): The realisers of a mental states sometimes include – or are even fully composed of – parts of the world that are different from an organism's brain, and which might even be external to the physical boundary of the organism (e.g. its skin).

Figure 1 makes this clearer.

[Figure 1]

Note that, as implied in the above statement, it is possible to distinguish stronger and weaker forms and instances of the EMT_1. According to the strongest forms, mental states are fully and wholly located in the external world (these are what Wilson 2004 calls 'radically wide realisation cases'). According to weaker forms, mental states are *partially* realised by features external to the brain or skin of the organism (they essentially 'involve' the environment, though are not *fully*

[2] Note that it is not strictly necessary to accept functionalism in order to accept the EMT_1; however, as noted by Sprevak (2009), the two bear a close relationship, and functionalism still is the most widely accepted account of the nature of mental states. See also Clark (1997) and Shapiro (2004, 2010).

located outside of the agent's brain). Fortunately, since all the conclusions of this paper hold for both stronger and weaker forms and instances of the EMT_1, it is not necessary to consider this issue further here. In what follows, I will therefore leave it open exactly how extended the mental states in question are to be taken to be.

In contrast to the more metaphysical EMT_1, the EMT_2 is best approached through the practice of cognitive science. In this practice, it is conventionally thought that in order to study and understand cognition, it is sufficient to focus on the information processing core of a cognitive system. For example, when it comes to human cognition, it is often thought that understanding how the brain works (e.g. how it is set up to process information) is the central and most important problem that needs to be solved (Fodor 1983, 2005; Pylyshyn 1984; Damasio 1994; Clark 1997; Rupert 2009a; Shapiro 2010).

Defenders of the EMT_2 think that this picture of the appropriate method of cognitive science is mistaken, as it fails to pay sufficient attention to the embodied and extended nature of cognition (Clark 2001, 132). In particular, cognitive systems typically require many parts to be interacting closely with one another in order to be functional: a working information processor (such as a human brain) needs to be embedded in other systems that connect it to the rest of the world (such as appropriate peripheral systems), and the world needs to (be made to) in fact provide appropriately structured information. Figure 2 makes this clearer.

[Figure 2]

Given this, the EMT_2 can be formulated as follows (Clark 1997; Shapiro 2004, 2011; Rupert 2009a; Stotz 2010; Wilson 2010a):

(EMT_2): The right unit of analysis in cognitive science is the entire cognitive system, or at least large parts thereof – understanding cognition requires paying attention to the neural, organismal (i.e. extra-neural) and environmental (i.e. extra-organismal) elements of the cognitive system.[3]

Again, note that it is possible to formulate stronger and weaker forms and instances of the EMT_2. In this case, this can be done by varying the relative amount of attention given to non-neural parts of cognition, with parity being the strongest version. As before, deciding this dispute is not necessary here – all that matters for present purposes is a parallelism with Developmental Systems Theory (see below for more on this).

While there is much that could be said concerning the EMT_1 and the EMT_2, I here focus on the attempt to use certain evolutionary biological considerations to support them further. In particular, various authors (Clark 1997, 2001, 2007; Clark & Wheeler 1999; Griffiths & Stotz 2000; Shapiro 2004, 219-220; Rupert 2009a, 2009b; Wilson 2005, 2010a, 2010b; Stotz 2010; see also Sterelny 2000) have suggested that the EMT (in one or both of its versions) can be supported by appealing to (a) the 'Extended Phenotype Thesis', or (b) recent work in Developmental Systems Theory. It is the goal of the rest of this paper to provide a critical analysis of these two evolutionary biological arguments for the EMT.

[3] It is possible to divide the EMT_2 as formulated here into two parts: on the one hand, there is the claim that the study of cognition should not just focus on the brain, but also on other parts of the organism's *body* (this is sometimes called the hypothesis of 'embodied cognition' – Shapiro 2004; Clark 1997); on the other hand, there is the claim that the study of cognition should also consider features of the organism's *external environment* (this is sometimes called the hypothesis of 'situated cognition' – Wilson 2004; Griffiths & Stotz 2000). While these two theses certainly deserve a separate discussion, for present purposes, it is more useful to consider them jointly – this fits better to the argument from Developmental Systems Theory laid out below. At any rate, all that matters here is a parallelism between whatever version of the latter thesis one wants to be working with; see below for more on this.

III. The Extended Phenotype and the Extended Mind

Richard Dawkins claims that the phenotype of an organism can extend into the environment of that organism (Dawkins 1982, 2004; see also Sterelny et al. 2001). Interestingly, various scholars have claimed that the existence of extended phenotypes also speaks in favour of the EMT (in at least one of its versions). Here, for example, is Clark:

> 'The form of reasoning [in support of the EMT_1] is thus similar to that which leads Dawkins to describe the web as the 'extended phenotype' of the spider [...] and that leads J. Scott Turner to treat the sound amplifying ('singing') burrows of the mole cricket as external physiological organs. In each case we start out with a working sense of some baseline concept (phenotype, organ). We then notice that stuff we do not ordinarily treat in those terms plays the play the right kind of role to be considered as belonging to that class.' (Clark 2007, note 17)

Shapiro expresses the point as follows:

> 'Recently, some philosophers [...] have promoted the idea that the use of various props external to the body extends the mind into the environment [...]. These philosophers are, in essence, applying Richard Dawkins's idea of an extended phenotype to the individuation of minds'. (Shapiro 2004, 219).

To understand this sort of argument better, it is best to begin by laying out the structure of Dawkins's defence of the extended phenotype in more detail.

1. The Extended Phenotype Thesis

The starting point of the Extended Phenotype Thesis (EPT) is the fact that a gene (i.e., in this context, a sequence of nucleotide bases on an organism's DNA) generates a chain of effects that stretches from the microphysical, molecular level of the gene itself to the macrophysical level of the organism's environment.[4] For present purposes, three elements of this chain are usefully

[4] For a useful introductory treatment of the way in which genes cause phenotypes, see Griffiths et al. (2000). A good philosophical introduction is in Sterelny & Griffiths (1999).

made explicit. Firstly, (coding) genes, with the help of the environment and the biochemical machinery inside an organism's cells, generate the so-called gene products (such as proteins). Secondly, these gene products, again modulated by a complex network of intervening causes and feedback loops, have various effects on the morphology and nervous apparatus of the organism. Thirdly, the structure of the organism's morphology and its nervous apparatus (sometimes) have various further effects in the external world – like spider webs and beaver dams. Figure 3 makes this clearer.

[Figure 3]

Given this, the key insight behind the EPT is that there is no privileged stage of this chain of effects to which the label 'phenotype' must be attached (Nijhout 2001; Lewontin 1992). The term 'phenotype' really just refers to whatever part of the above causal chain a particular researcher is interested in (see e.g. Dawkins 1982, 230; Sterelny et al. 2001).

This matters, as it implies that considerations of theoretical utility might, at least at times, strongly speak in favour of seeing the later, external parts of the chain as the phenotype of a gene complex: this will make it easier to see why this phenotype evolved, how it relates to other features of the organism, and what further effects it has (Dawkins 1982, 196-202; see also Sterelny et al. 2001, 71-73, 77-79). In particular, when investigating changes in gene frequency (i.e. on a common definition, biological evolution), it will often be more convenient to consider the distal effects of a gene directly, instead of having them merely be the 'after effects' of the phenotype proper (Dawkins 1982, 206-207).[5] While it is true that focusing on the latter is a

[5] In the background here is also Dawkins's view that the relevant unit of selection is always the gene. While this is questionable, nothing hangs on this here.

principled possibility, it is often theoretically quite cumbersome, and we are normally better off to focus on the former (Dawkins 1982, 232; Dawkins 2004, 382).[6] In short:

(EPT) It is sometimes useful to see phenotypes as (at least partly) physically disconnected from an organism and as comprising features of the external world.

2. From the Extended Phenotype to the Extended Mind

The EPT matters here, as one might see it as an evolutionary biological analogue to the EMT_1: what the EPT says about the evolutionary biology realm would seem to bear structural similarities to what the EMT_1 says about the cognitive scientific realm. In particular, one might think that, just as the EPT shows that phenotypes can be (wholly or partly) located in the external world, so the EMT_1 shows that mental states can be (wholly or partly) located in the external world. Indeed, as noted above, several scholars seem to think precisely this (Clark 2001, 132; Clark 2007, note 17; see also Sterelny 2000; Shapiro 2004, 219-220). This would be important, as this analogy between the two theses could then be used to support the latter thesis further: to the extent that we are unclear about what the EMT_1 says, for example, we could elucidate this thesis further by appeal to the EPT. Also, given this analogy, we might realise that the types of considerations speaking in favour of the EPT also speak in favour of the EMT_1.[7] (I will consider below whether the same could be said concerning the EMT_2.)

[6] So, for example, Dawkins (1982, 232) writes 'The language of extended genetics is not demonstrably more correct. […] I suggest that the way the extended geneticist tells the story of the Bruce effect [the fact that the semen of a second male mouse can prevent the semen of a first male from impregnating a given female] is more elegant and parsimonious that the way the conventional geneticist would have told it.' Similarly, Dawkins (2004, 382) says: '[Termite] [m]ound 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.' (italics added).

[7] For a classic discussion of reasoning by analogy, see Hesse (1966).

Alas, there are some good reasons to doubt that there is in fact an interesting analogy between the EPT and the EMT_1 that could be appealed to here. In particular, when considered in more detail, the two theses turn out to be structurally very different from each other. This is due to the fact that the realisers of mental states – such as brain states – do not relate to mental states in the way that genes relate to phenotypes. While phenotypes are the (more or less distant) direct effects of genes (among other things), mental states are not the (more or less distant) direct effects of their realisers – after all, the latter *realise* the former, and do not (directly) cause them. This point also comes out clearly when comparing figures 1, and 3 above: while the extendedness of a phenotype is a matter of the location of the direct effects of a gene, the extendedness of a mental state is a matter of the location of its realisers (note that phenotypes are not *realised* by the effects of a gene – they *are* the effects of a gene – see also Nijhout 2001). For this reason, the fact that phenotypes might be external does not imply anything about whether mental states might be external: the (supposed) fact that the external effects of a gene sometimes should be seen to be its phenotype does not make it more plausible that mental states sometimes are constituted by external structures. These two claims are based on very different kinds of considerations.

A different way of putting this point is that the EPT and the EMT_1 are very different *kinds* of theses: the EPT is an inherently causal and methodological thesis, whereas the EMT_1 is an inherently constitutional and ontological thesis. Indeed, the reasons why a mental state and a phenotype are said to be extended are very different in the case of the EMT_1 and the EPT: in the former case, this is a mind and language-independent fact about the world; in the latter case, this is a mind and language-dependent fact about what we find useful in theorising about the world.

Importantly, much the same holds for the relationship between the EPT and the EMT_2. This is so, as it is not even clear how one would begin to make the case for EMT_2 using the EPT: the latter is a thesis about which of the many effects of a gene evolutionary biologists should consider its phenotype, the former is a thesis about what kind of entity cognitive scientists should study. The fact that evolutionary biologists might find it useful to see the extended effects of a gene as its phenotype seems, as such, to say nothing whatsoever about whether cognitive scientists should find it useful to focus on entire cognitive systems (or at least large subsets thereof) as the relevant subjects of their studies: one is claim about the theoretically most perspicuous way of understanding how gene frequencies change in a population, and the other is claim about the right unit of analysis in cognitive science. For this reason, we might agree that phenotypes are (often) best seen as features of the organism's environment, and still deny that the right unit of analysis in cognitive science is the cognitive system as a whole (or at least in large part) – we might be happy to talk about the evolution of different kinds of spider webs, but we might claim that the way webs are built by spiders is best analysed by focusing on their brains. So, while it is true that both the EPT and the EMT_2 make claims about which methodological approaches it is useful to pursue in order to push the relevant sciences forward, the kind of methodological approaches in questions are too different to make any meaningful connections between them possible. In short: the truth of the EPT has no clear bearings on that of either the EMT_1 or the EMT_2.

IV. Developmental Systems Theory and the Extended Mind

A number of biologists and philosophers have recently called for a new and quite radial approach towards evolutionary biological theorising – one that emphasises the interconnected nature of

many biological phenomena (Oyama 1985; Griffiths & Gray 1994, 2004, 2005; Neumann-Held 1999; Sterelny & Griffiths 1999; Robert 2004; Wilson 2005; Stotz 2006; Gilbert & Epel 2009; see also Rosenberg 2000). This new approach has become known as 'Developmental Systems Theory' (DST). Interestingly also, it might be thought that this approach can be used to support either version of the EMT. So, for example, Wilson says:

> '[T]here is a fairly direct parallel between the DST and traditional views of inheritance and development, on the one hand, and locational externalism and traditional views of cognition, on the other.' (Wilson 2010a, 178)

Similarly, Rupert says:

> 'Developmental systems theory thus appears to support the extended view in a fairly straightforward way. [...] The extended theorist simply takes the systems instantiating cognitive properties to be, or to at least be similar in scope to, the [extended] systems of fundamental importance in respect of the biological processes that give rise to cognitive phenomena.' (Rupert 2009a, 113)

In what follows, I first lay out DST, and then consider the extent to which it can be seen to support the EMT_1 or EMT_2.

1. Developmental Systems Theory

There are two key ideas underlying DST. First, DST takes a cue from the EPT in noting that there are many elements to the gene-phenotype mapping, and that the relationship between these elements is very complex. However, unlike the EPT, DST conceives of the relationship between genes and phenotypes less linearly, and more enmeshed in a network of causes (Griffiths & Gray 1994, 288-290; Stotz 2006, 2010; see also Waters 2007). In particular, this theory emphasises that genes interact with other components of the cellular and organismal environment, and it is really only this entire interactive collection that is able to bring a phenotype into existence

(Neumann-Held 1999; Oyama et al. 2001; Griffiths & Gray 2004; Robert 2004; Jablonka & Lamb 2005; Gilbert & Epel 2009; Purugganan 2010; Hallgrimsson & Hall 2011; Gissis & Jablonka 2011). For example, the way DNA sequences are transcribed into amino acids depends on the presence or absence of various promoters and regulator molecules, as well as the existence of certain kinds and quantities of mRNA and tRNA molecules; also, how amino acids are folded and combined so as to (partially) make up proteins is equally dependent on various other molecules (e.g. 'chaperons' that guide them through the cell environment); finally, how proteins are expressed as morphological or behavioural effects is a matter of various further complex mechanisms (Oyama et al. 2001; Griffiths & Gray 2004; Jablonka & Lamb 2005, 2010). Importantly moreover, these different elements can influence each other: regulator DNA sequences can influence the transcription of other DNA sequences, the way mRNA is spliced, etc. – and vice versa. In short: phenotypes are caused by an interactive combination of genetic, cytological, and environmental factors. Figure 4 makes this clearer.

[Figure 4]

The second key idea underlying DST is that all the elements of the network of causes and effects that brings about an organism's phenotype – whether genetic, cellular, or environmental – should be taken seriously in evolutionary and developmental biology (Oyama 1985; Griffiths & Gray 1994, 2004; Neumann-Held 1999; Oyama et al. 2001; Robert 2004; Stotz 2006; Gilbert & Epel 2009, chap. 10 and 404-408).[8] Given this, stronger and weaker form and instances of DST

[8] As noted in the context of the EMT_2 (see note 3), there are several further distinctions that could be made here: in particular, one could argue that it is in particular the *cellular and wider organism-internal* environment that should be paid (more) attention to, or that both the *internal and external* environment should be considered important for evolutionary and developmental investigations. As also noted earlier, though, making these finer distinctions is not

can be distinguished by focusing on the degree of attention that should be given to non-genetic factors – with parity being the most extreme form. Fortunately, since this is exactly parallel to the case of the EMT_2, it is once again not necessary to discuss these complications further – I here assume that all that is attempted by the defenders of the DST-EMT_2 link is to argue for a form of EMT_2 that is as strong as the most plausible version of DST (whatever form that is taken to be).

Two main arguments in favour of DST have been put forward. On the one hand, it has been argued that, in general, there are no good reasons to prioritise genes as 'the' causes of the phenotype. At least in many realistic (i.e. not artificially simplified) cases, cytological features (e.g. the types of RNA polymerase available in the cellular environment) and environmental features (e.g. the kinds of nutrition the organism has access to) are just as much responsible for making a phenotype what it is as genetic features are (Griffiths & Gray 1994, 2004; Sterelny & Griffiths, 1999; Oyama et al. 2001; Robert 2004; Jablonka & Lamb 2005; Maestripieri & Mateo 2009).[9] For example, it is often as easily possible to change which proteins are synthesised (and when) by changing an organism's DNA as by changing the way its mRNA is spliced or which foodstuffs it is consuming.

On the other hand, it can be argued that there are good reasons to consider developmental systems in their entirety when considering various evolutionary and developmental biological

necessary here: I simply assume that defenders of some specific version of the EMT_2 want to argue for this version of the thesis by appealing to a *parallel* version of the DST.

[9] Note also that various theorists have suggested that it is not even *in principle* possible to mark genes as especially important causes of a phenotype (see e.g. Sterelny & Griffiths 1999 for an overview of some of the arguments here). However, this more philosophical issue is quite hotly contested (see e.g. Waters 2007 and Stotz 2006 for a recent exchange on this matter). At any rate, the point emphasised here is just that defenders of DST can and do point to the *methodological* challenges of distinguishing the influence of genes from those of other cellular and environmental causes of a phenotype in many real life cases. This point is less controversial than the more principled issue of apportioning causal or informational responsibility, and fits better to the discussion of the EMT_1 and EMT_2 at stake here.

problems.[10] In particular, by taking entire developmental systems as the basic unit of analysis, hitherto disparate phenomena can be seen to be unified. For example, recent work in evolutionary ecology on the phenomenon of 'niche construction' has shown that it can be equally evolutionarily important when an organism changes its environment (e.g. through building dams or subterranean dwellings) as when it changes genetically (Odling-Smee et al. 2003; Jablonka & Lamb 2005; Gilbert & Epel 2009; Gissis & Jablonka 2011). DST makes it easy to handle these sorts of cases, as it does away with a strict organism / environment distinction in the first place (Griffiths & Gray 1994, 2004; Oyama et al. 2001; Sterelny & Griffiths 1999; Stotz 2006, 2010; Gilbert & Epel 2009).

2. From Developmental Systems Theory to the Extended Mind

In the discussion surrounding the EMT_2, DST may be thought to matter, as there seems to be a kind of meta-level parity between the two kinds of parity theses: both of these theses emphasise how a complex interplay among the components of a system is necessary for the system to function properly (as is illustrated by the many seeming similarities between figures 2 and 4). Also, in both cases, it is said that only by considering the system *in toto* (or at least large parts of it) can progress in the relevant science be made. Hence, one might conclude that accepting DST goes hand in hand with accepting the EMT_2 – to the extent that one is convinced by DST (for which, as noted earlier, there are some good reasons), one ought also to be convinced by the EMT_2. Indeed, as noted above, several scholars seem to think exactly this (Wilson 2010a, 2010b,

[10] There is some variation, though, in terms of the exact practical implications of DST for biological research. Some scholars – such as Oyama (2000) and Sterelny & Griffiths (1999) – suggest that it is possible to analyse developmental systems 'piecemeal' , while others – such as Roberts (2004) – think only a properly holistic perspective is appropriate. However, there can be no question that, for DST to be an interesting alternative to conventional work in evolutionary biology, only the entire developmental system (or at least large parts thereof) can, at least ultimately, be an appropriate unit of evolutionary biological analysis.

2005; Stotz 2010; Clark 2001, 132; Griffiths & Stotz 2000; see also Stotz & Allen 2012; Rupert 2009a, 2009b; Clark & Wheeler 1999). (We will consider below whether something similar could also be said to be true for the EMT_1.)

However, when considered in more detail, this biological route towards the EMT_2 cannot be seen to be very plausible either.[11] The overarching reason for this is that it is not clear how the (assumed) truth of DST is meant to carry over to that of the EMT_2 – that is, how much 'parity' there really is between the two theses. To see this, note that the argument for DST is based on the claim that, in order to do evolutionary and developmental biology well, there are no reasons to privilege genes over other elements of the developmental system, and some good reasons to treat them equally (or nearly so). However, as such, this says nothing about whether it is also true that, in order to do cognitive science well, there are no reasons to privilege brains (and other information processors) over other elements of the cognitive system, and some good reasons to treat them equally (or nearly so).

A similar worry has also been expressed by Rupert – though, as will become clearer momentarily, he arrives at it in a different way from how I do here:

> 'In the end, then, two systems-based concerns threaten the marriage of Developmental Systems Theory and the extended view. First, a plausible Developmental Systems Theory validates only a narrow range of genuinely extended biological individuals; the shared-fate criterion severely limits the number of such extended systems. Second, no matter how things turn out in respect of biology, we cannot ignore the potential for mismatch between the extended individuals established by Developmental Systems Theory and those systems claimed by the extended view to be extended cognitive systems.' (Rupert 2009a, 117)

[11] Note that questioning the DST-EMT_2 argument by denying the plausibility of DST in and of itself is not very compelling, as the key tenets of this thesis are now sufficiently well corroborated that even mainstream biologists and philosophers have come to accept them – at least in part (Oyama et al. 2001; Robert 2004; Gilbert & Epel 2009; Maestripieri & Mateo 2009; Stotz 2010). Of course, this is not to say that DST has now been fully confirmed and that there is no more debate surrounding it (see e.g. the exchange mentioned in note 9 above and the complications noted in note 8). However, the latter's evidential backing is sufficiently strong to at least warrant taking a possible derivation of the EMT_2 from DST as a point in the former's favour. This is all that is needed here.

The reason Rupert gives to support the first concern is that, to avoid the threat of triviality, defenders of DST have to adopt some criterion of where the boundaries of developmental systems are: where does one end and the next one begin (Rupert 2009, 114-116)? Furthermore, according to Rupert, the most plausible such criterion is based on when the components of a system have a shared fate (form a 'trait group') – i.e. when their evolutionary and developmental trajectories are correlated (see also Sober & Wilson 1998). The reason Rupert gives in support of the second concern is that 'the narrowness of the range of extended biological systems dims the prospects for cross-disciplinary fit. It may well be true that natural forces sometimes select for extended systems, but such systems might not be the ones of interest to psychology' (Rupert 2009, 117).

However, Rupert's argument has been attacked by defenders of a connection between DST and the EMT_2. In particular, Stotz (2010) claims that this argument is confused: she suggests that Rupert is wrong to claim that the most plausible criterion for individuating developmental systems is based on the existence of 'trait groups'. In fact (she goes on to say), this last criterion is part of an account of when groups of organisms can be a *unit of selection*, not when some entities form a developmental system (Stotz 2010, 486-487). Furthermore, she thinks that there are other (and more plausible) criteria that would lead extended developmental systems to be a much more populated class than what is claimed by Rupert (Stotz 2010, 487). Hence, she thinks that Rupert has failed to show that it is not possible to argue from DST to the EMT_2.

I think that both authors are partly right here. Stotz (2010) is correct in noting that the existence of trait groups is not the best way of picking out developmental systems: not every trait group needs to form a developmental system (e.g. various organisms – insects, fungi, etc. – that live on the same leaf might form a trait group, but *need not* be a developmental system in the

DST sense), and not every developmental system needs to be easily definable in terms of trait groups (indeed, this will be hard for organisms that live in relative isolation – such as plants like the Saguaro cactus – but which might still be well analysed as developmental systems with many non-organic parts). Furthermore, I agree with her that the class of extended developmental systems might well be quite large, if a different (and more plausible) criterion for individuating them is employed.

However, I also think that, ultimately, this does not invalidate the worries expressed above concerning the DST-based case for the EMT_2. In particular, I agree with Rupert that there is a potential mismatch between DST and the EMT_2 – however, unlike him, I do not think that the 'width' of the class of extended developmental systems is the reason for this. In fact, I think that the real worry for a DST-EMT_2 link is based on the fact that there are just no a priori grounds to think that, just because evolutionary-developmental biology must appeal to entire systems (or large parts thereof), so does cognitive science. Even if we did find it useful to theorise in terms of large numbers of (extended) developmental systems in evolutionary-developmental biology, that does not, as such, imply that we would also find it useful to theorise in terms of large numbers of (extended) cognitive systems in cognitive science. Irrespective of how many developmental systems there are, what is key is that DST and the EMT_2 rely on argumentative support that is unique to them – they are tied to their respective sciences quite closely, and thus cannot so easily be removed from these sciences.

At this point, an objection might be raised. This objection is based on the idea that extended cognitive systems are just proper parts or aspects of (extended) developmental systems (see e.g. Stotz 2010).[12] In particular, it might be thought that: (a) What DST shows is that organismic

[12] Note that the appeal to proper parts is key here, and distinguishes it from Rupert's (2009) concern with the width of the category of (extended) developmental systems noted earlier. The point here is not that cognitive systems are

development and evolution depend crucially on various kinds of non-genetic – such as external – features, and must therefore be analysed by considering the developmental system as a whole (or at least in large part). (b) In many cases, developmental systems have cognitive aspects (after all, cognition is a biological phenomenon). (c) Given (a), the cognitive aspects of developmental systems will include features external to the organism in question; among these might be the culture the organism lives in or the specially 'engineered' epistemic environment it grows up in (Sterelny 2003, 2012; Stotz 2010). (d) Hence, (human) cognition must be analysed by considering the cognitive system as a whole (or at least in large part), as suggested by the EMT_2. In other words, one would seem to get the EMT_2 for free by accepting DST: if we agree that understanding (human) evolution and development requires analysing it by reference to the entire developmental system, we must also accept the EMT_2 – for cognitive development and evolution are parts of overall (human) development and evolution.

However, while very interesting, this argument ultimately does not appear to be helpful here. The reason for this is that premiss (c) either begs the question or does not allow the derivation of conclusion (d).[13] In particular, if the claim in (c) is understood as saying that (human) cognition *somehow* involves various external processes, then this is undeniably true – but it does not underwrite the EMT_2. This is so, as, understood like this, claim (c) does not prohibit traditional cognitive scientists from claiming that the external factors at stake merely concern various subsidiary features of (human) cognition that it is not necessary to take into account when trying to understand (human) cognition as such (Rupert 2009a, chaps 6-7). In other words: on a weak

simply one type of developmental system; the claim is that cognitive systems are just *parts* of some developmental systems. Hence, the *number* of the latter is not so relevant here: what matters is their *nature*.

[13] This is not to say that this objection is completely without interest – on the contrary. In particular, the argument it is based on shows that, if the EMT_2 is accepted as true, then the cognitive systems the latter identifies might be fruitfully seen as elements of the developmental systems identified by DST. This is an important insight – it just does not help us determine whether the EMT_2 is in fact to be accepted as true.

reading of (c), cognitive scientists are free to say that, while a *full* picture of organismic evolution and development needs to situate the organism in its (partly self-constructed) environment, a picture only of the way *cognition* works *now* can focus largely on the way brains process information (or some such).[14]

By contrast, if the claim in (c) is understood more strongly as saying that (human) cognition involves various external processes *in the same way that the entire (human) developmental system involves external features*, then this does underwrite the EMT_2 – but only by begging the question in favour of the latter thesis. In particular, whether (human) cognition is embedded in its environment in the same way that the entire (human) developmental system is embedded in its environment is precisely what the debate surrounding the EMT_2 is all about. For this reason, this claim cannot be an assumption in an argument that seeks to favour one side of this debate – it needs to be its conclusion.

Overall, therefore, it becomes clear that Rupert (2009) is right in being sceptical of the support DST can give to the EMT_2. While the particular arguments he gives might face objections, the general idea behind the concerns expressed by him is cogent, and shows that this way of underwriting the truth of the EMT_2 is not compelling.

What about the EMT_1 – can DST be used to support this version of the EMT further? Alas, the answer here seems to be no as well, and that for reasons that have already been laid out in the context of the discussion surrounding the EPT. These two theses make very different kinds of claims: the EMT_1 is an inherently metaphysical / ontological thesis about which features of the world can realise mental states; DST is a theory about what kind of methodology is useful for

[14] Indeed, this sort of stance can even be found among strong champions of the importance of 'external' features for (human) evolution and development. For example, Boyd & Richerson (2005) seem quite comfortable with the traditional approach towards cognitive science, despite being adamant about stressing the importance of cultural influences on human cognition.

doing biology. As made clear earlier, it is hard to even begin to see how these two can be connected in any kind of meaningful way. Hence, there is little to be gained by appealing to DST to support the EMT_1 further.

V. Conclusion

I have tried to show that appealing to evolutionary biological considerations derived either from the EPT or from DST fails to support the EMT – either if spelled out as EMT_1 or as EMT_2. Importantly, this conclusion also reveals two more general lessons that go beyond the plausibility of the EMT, the EPT, and DST directly at stake here: firstly, not all ways of externalising traditionally internal features of an organism are created equal – in particular, the way in which phenotypes might be said to be extended turns out to be fundamentally different from the way in which mental states or cognitive systems might be seen to be extended, and both of these are (or might be) different from the way in which developmental systems might be said to be extended. Secondly, while the introduction of evolutionary biological considerations into psychological or philosophical debates sometimes allows for significant progress to be made in these debates (see e.g. Sober & Wilson 1998), this is not always the case. Everything here depends on the details, and a close analysis of the relevant arguments is necessary to determine what the situation is in each particular case.

Bibliography

Adams, Frederick & Aizawa, Kenneth (2001). 'The Bounds of Cognition'. *Philosophical Psychology*, 14, 43-64.

Adams, Frederick & Aizawa, Kenneth (2008). *The Bounds of Cognition*. Oxford: Blackwell Publishing.

Clark, Andy (1997). *Being There*. Cambridge, MA: MIT Press.

Clark, Andy (2001). 'Reason, Robots, and the Extended Mind'. *Mind and Language* 16: 121-145.

Clark, Andy (2007). 'Curing Cognitive Hiccups'. *Journal of Philosophy* 104: 163-192.

Clark, Andy (2008). *Supersizing the Mind*. Oxford: Oxford University Press.

Clark, Andy, and Chalmers, David (1998). 'The Extended Mind'. *Analysis* 58: 7-19.

Clark, Andy, and Wheeler, Michael (1999). 'Genic Representation'. *British Journal for the Philosophy of Science* 50: 103-135.

Damasio, Antonio (1994). *Descartes' Error*. New York: Grossett / Putnam.

Dawkins, Richard (1982). *The Extended Phenotype*. Oxford: Oxford University Press.

Dawkins, Richard (2004). 'Extended Phenotype – But Not Too Extended: A Reply to Laland, Turner, and Jablonka'. *Biology and Philosophy* 19: 377-396.

Fodor, Jerry (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.

Fodor, Jerry (2005). *The Mind Doesn't Work This Way*. Cambridge, MA: MIT Press.

Gilbert, Scott, and Epel, David (2009). *Ecological Developmental Biology: Integrating Epigenetics, Medicine, and Evolution.* Sunderland, MA: Sinauer Associates.

Gissis, Snait, and Jablonka, Eva (eds.) (2011). *Transformations of Lamarckism: From Subtle Fluids to Molecular Biology*. Cambridge, MA: MIT Press.

Griffiths, Anthony; Wessler, Susan; Lewontin, Richard, and Carrol, Sean (2000). *An Introduction to Genetic Analysis*. Ninth Edition. New York: Freeman.

Griffiths, Paul, and Gray, Russell (1994). 'Developmental Systems and Evolutionary Explanation'. *The Journal of Philosophy* 91: 277-304.

Griffiths, Paul, and Gray, Russell (2004). 'The Developmental Systems Perspective: Organism-Environment Systems as Units of Evolution.' In K. Preston and M. Pigliucci (eds.). *Phenotypic Integration: Studying the Ecology and Evolution of Complex Phenotypes*. Oxford: Oxford University Press, pp. 409-431.

Griffiths, Paul, and Gray, Russell (2005). 'Discussion: Three Ways to Misunderstand Developmental Systems'. *Biology and Philosophy* 20: 417-425.

Griffiths, Paul, and Stotz, Karola (2000). 'How the Mind Grows'. *Synthese* 122: 29-51.

Hallgrimsson, Benedikt, and Hall, Brian (eds.) (2011). *Epigenetics: Linking Genotype and Phenotype in Development and Evolution*. Berkeley: University of California Press.

Hesse, Mary (1966). *Models and Analogies in Science*. Notre Dame: University of Notre Dame Press.

Jablonka, Eva, and Lamb, Marion (2005). *Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life*. Cambridge, MA: MIT Press.

Jablonka, Eva, and Lamb, Marion (2010). 'Transgenerational Epigenetic Inheritance'. In M. Pigliucci & G. Muller (eds.). *Evolution – The Extended Synthesis*. Cambridge, MA: MIT Press, pp. 137-174.

Lewontin, Richard (1992). 'Genotype and Phenotype'. In E. F. Keller & E. Lloyd (eds.). *Key Terms in Evolutionary Biology*. Cambridge, MA: Harvard University Press, pp. 137-144.

Maestripieri, Dario, and Mateo, Jill (eds.) (2009). *Maternal Effects in Mammals*. Chicago: The University of Chicago Press

Menary, Richard (2007). *Cognitive Integration*. New York: Palgrave Macmillan.

Neumann-Held, Eva (1999). 'The Gene is Dead – Long Live the Gene'. In P. Koslowski (ed.). *Sociobiology and Bioeconomics*. Berlin: Springer, pp. 105-138.

Nijhout, H. Frederik (2001). 'The Ontogeny of Phenotypes'. In S. Oyama, P. Griffiths, and R. Grey (eds.). *Cycles of Contingency*. Cambridge, MA: MIT Press, pp. 129-140.

Odling-Smee, John; Laland, Kevin, and Feldman, Marcus (2003). *Niche Construction*. Princeton: Princeton University Press.

Oyama, Susan (1985). *The Ontogeny of Information*. Cambridge: Cambridge University Press.

Oyama, Susan (2000). *Evolution's Eye*. Durgham, NC: Duke University Press.

Oyama, Susan; Griffiths, Paul, and Grey, Russell (2001) (eds.). *Cycles of Contingency: Developmental Systems and Evolution*. Cambridge, MA: MIT Press.

Purugganan, Michael (2010). 'Complexities in Genome Structure and Evolution'. In M. Pigliucci & G. Muller (eds.). *Evolution – The Extended Synthesis*. Cambridge, MA: MIT Press, pp. 117-134.

Pylyshyn, Zenon (1984). *Computation and Cognition*. Cambridge, MA: MIT Press.

Richerson, Peter, and Boyd, Robert (2005). *Not By Genes Alone*. Chicago: University of Chicago Press.

Robert, Jason Scott (2004). *Embryology, Epigenesis, and Evolution*. Cambridge: Cambridge University Press.

Rosenberg, Alexander (2000). 'What Happens to Genetics When Holism Runs Amok?'. In *Darwinism in Philosophy, Social Science and Policy*. Cambridge: Cambridge University Press, pp. 97-117.

Rowlands, Mark (1999). *The Body in Mind*. Cambridge: Cambridge University Press.

Rupert, Robert (2004). 'Challenges to the Hypothesis of Extended Cognition'. *Journal of Philosophy* 101: 1-40.

Rupert, Robert (2009a). *Cognitive Systems and the Extended Mind*. Oxford: Oxford University Press.

Rupert, Robert (2009b). 'Innateness and the Situated Mind'. In P. Robbins and M. Aydede (eds.). *The Cambridge Handbook of Situated Cognition*. Cambridge: Cambridge University Press, pp. 96-116.

Shapiro, Lawrence (2004). *The Mind Incarnate*. Cambridge, MA: MIT Press.

Shapiro, Lawrence (2010). *Embodied Cognition*. London: Routledge.

Shapiro, Lawrence (2011). 'James Bond and the Barking Dog: Evolution and Extended Cognition'. *Philosophy of Science* 77: 410-418.

Sober, Elliott, and Wilson, David Sloan (1998). *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge, MA: Harvard University Press.

Sprevak, Mark (2009). 'Extended Cognition and Functionalism'. *Journal of Philosophy* 106: 503-527.

Sterelny, Kim (2000). 'Roboroach: The Extended Phenotype Meets Cognitive Science'. *Philosophy and Phenomenological Research* 61: 207-215.

Sterelny, Kim (2003). *Thought in a Hostile World*. Oxford: Blackwells.

Sterelny, Kim (2012). *The Evolved Apprentice*. Cambridge, MA: MIT Press.

Sterelny, Kim, and Griffiths, Paul (1999). *Sex and Death*. Chicago: University of Chicago Press.

Sterelny, Kim; Smith, Kelly, and Dickison, Mike (2001). 'The Extended Replicator'. In K. Sterelny. *The Evolution of Agency and Other Essays*. Cambridge: Cambridge University Press, pp. 53-80.

Stotz, Karola (2006). 'Molecular Epigenesis: Distributed Specificity as a Break in the Central Dogma'. *History and Philosophy of the Life Sciences* 28: 527-544.

Stotz, Karola (2010). 'Human Nature and Cognitive–Developmental Niche Construction'. *Phenomenology and Cognitive Science* 9: 483–501.

Stotz, Karola, and Allen, Colin (2012). 'From Cell-Surface Receptors to Higher Learning: A Whole World of Experience' In K. Plaisance and T. Reydon (eds.). *Philosophy of Behavioral Biology*. Dordrecht: Springer, pp. 85-123.

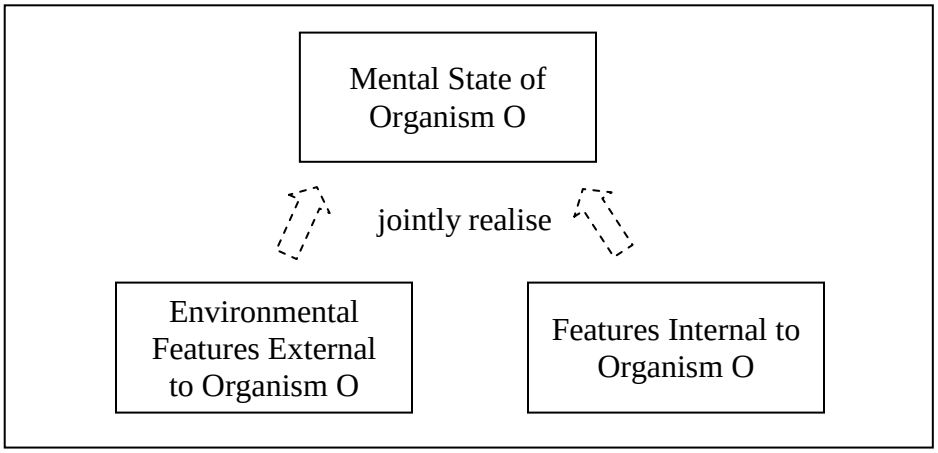
Wilson, Robert (2004). *Boundaries of the Mind*. Cambridge: Cambridge University Press.

Wilson, Robert (2005). *Genes and the Agents of Life: The Individual in the Fragile Sciences*. Oxford: Oxford University Press.

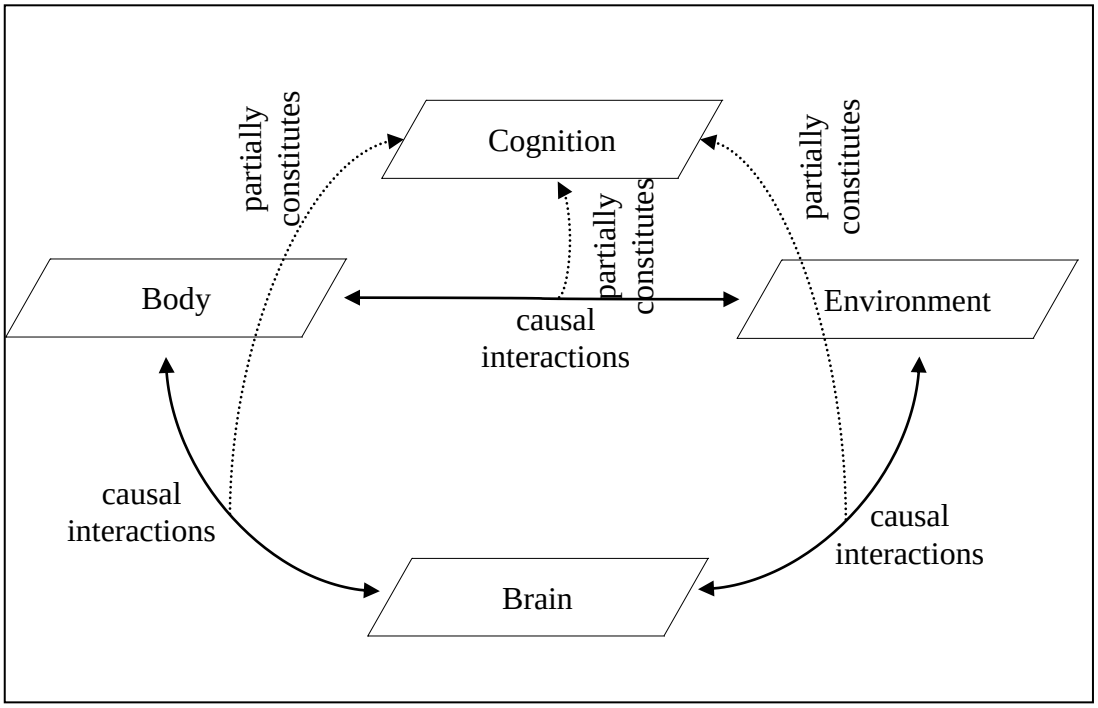
Wilson, Robert (2010a). 'Meaning Making and the Mind of the Externalist'. In R. Menary (ed.). *The Extended Mind*. Cambridge, MA: MIT Press, pp. 167-188.

Wilson, Robert (2010b). Review of Robert Rupert's 'Cognitive Systems and the Extended Mind'. *Notre Dame Philosophical Reviews*.

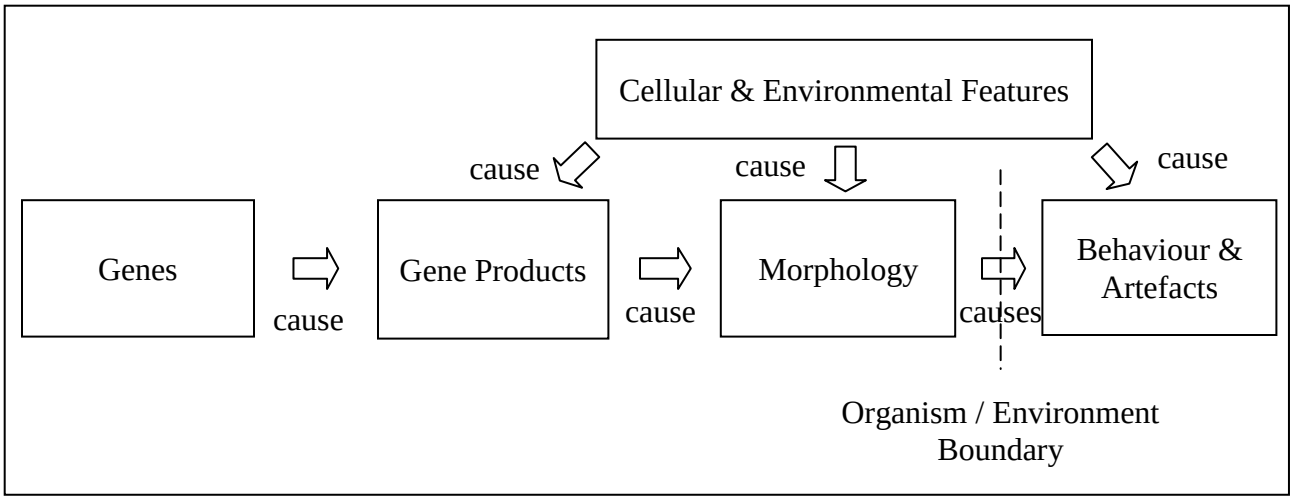
Figures



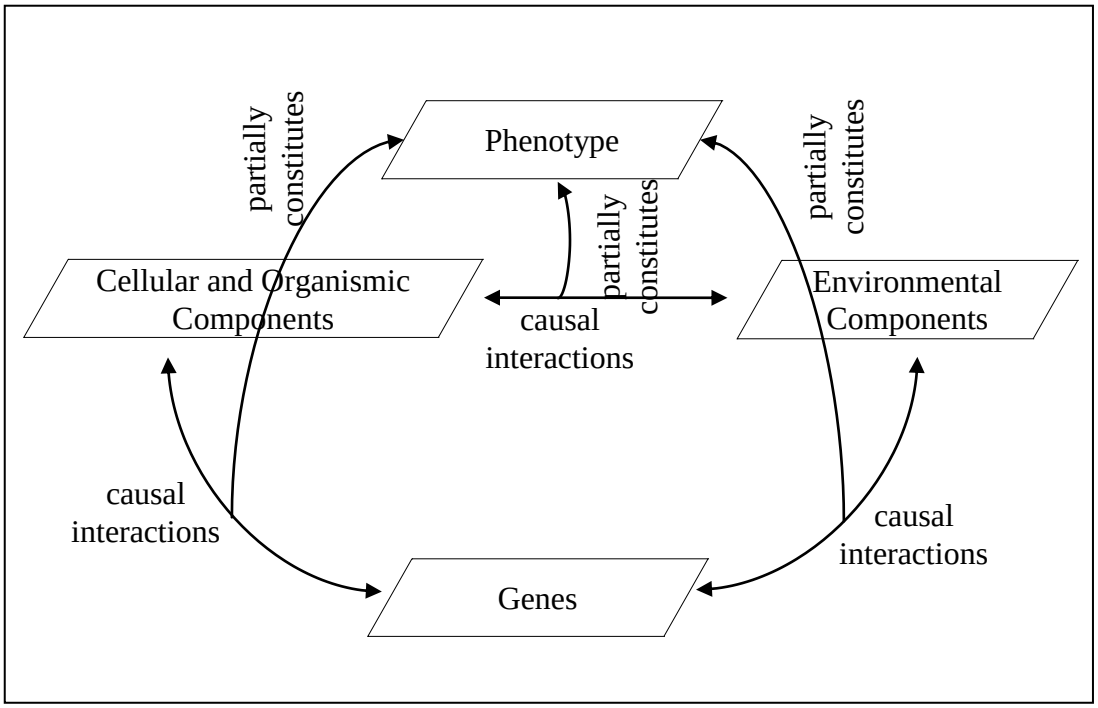
[Figure 1: A Partially Extended Mental State]



[Figure 2: The Causal Interaction among the Different Components of a Human Cognitive System Constitutes Cognition]



[Figure 3: Three Stages in the Genotype / Phenotype Relation]



[Figure 4: The Causal Interaction among the Different Components of a Developmental System Constitutes its Phenotype]