ACTIVE PERCEPTION AND VEHICLE EXTERNALISM
Susan Hurley

A revision of ch. 8 of *Consciousness in Action* (Harvard University Press 1998), which resolves an ambiguity in the published text.

1. Some cases.

Certain empirical results suggest a way of challenging two natural and widespread assumptions about the mind. One assumption is about the relations between perception and action. This shows up in the widespread conception of perception and action in terms of input and output, respectively. Perception is conceived as input from world to mind and action is conceived as output from mind to world. The other assumption is about the relations between mind and world. It influences various opposed views about whether the contents of the mind are in principle independent of the outside world.

We'll return to these assumptions shortly and spell them out further. But first consider two cases.

A. The first is a case of perceptual adaptation to long-term wearing of left-right reversing goggles, described by James Taylor. Such goggles can be made by mounting a "prism with its base parallel to the median plane of the head" (Taylor 1962, p. 192). (If you don't have a prism handy, you can get a similar effect by looking into a small mirror held up to the side of your eye like a blinker.) The earlier goggle experimenters had worn goggles continuously except when wearing a blindfold. But Taylor's volunteer subject , the mathematician Seymour Papert, wore his goggles every morning, and took them off every afternoon. Papert went through a torturous training program during his goggle periods, supervised by a Mrs. Schermann. She issued instructions such as: "Take the book on the right with your left hand and place it on the right-hand chair, at the same time putting your left foot on the left chair". Failure in some of these exercises met with sharp taps on the hand and blows with a walking stick. Taylor reports that, during one of his exercises, Papert

> ...had his first experience of perceiving an object in its true position. But it was a very strange experience, in that he perceived the chair as being both on the side where it was in contact with his body and on the opposite side. And by this he meant not just that he knew that the chair he saw on his left was actually on his right. He had that knowledge from the beginning of the experiment. The experience was more like the simultaneous perception of an object and its mirror image, although in this case the chair on the right was rather ghost-like. (Taylor 1962, p. 202; see also p. 206; cf. Harris 1980, pp. 109-113)

Taylor goes on to report the final result of the experiment:

> As predicted, adaptation was not blocked by the daily periods of normal vision. On the contrary, it was more rapid than it had been in the two previous experiments ..., and this can doubtless be attributed to the systematic training.... Also there was no disruption of behavior when the spectacles were put on or taken off. And finally, the prediction that the right-left ordering of the perceptual field would remain unchanged when the spectacles were put on or removed was confirmed. This was strikingly illustrated when the subject

> rode a bicycle while wearing the spectacles, and took them off and replaced them without changing course or wobbling or showing any other signs of disruption. Objects that he perceived as being on his left while wearing the spectacles were still on his left when he took them off. (Taylor 1962, p. 204)[1]

B. The second case, described by Ivo Kohler, also involves perceptual adaptation, this time to blue-left yellow-right goggles. Each lens of a pair of goggles is vertically divided. The left side of each lens is tinted blue and creates an apparently blue-tinted world when the wearer looks to the left. The right side of each lens is tinted yellow and creates an apparently yellow-tinted world when the wearer looks to the right. If the goggles are worn for several weeks, the eye adapts and the color distortions tend to disappear. Somehow the visual system learns to introduce the proper correction according to whether the eyes are turned to the left or right. Kohler writes:

> When the subject eventually removes his two-color goggles after wearing them continuously for 60 days, there is no doubt that his visual world is tinged distinctly yellow when he looks in the direction that his goggles had been blue and blue in the direction that his goggles had been yellow. The movement of the eyes either to right or left seems to act as a signal for the foveal area to switch over in its color response, compensating for a yellow image in one case and a blue image in the other. (Kohler 1961, p. 68; see also p. 65)[2]

In both these cases, perceptual experience seems to depend on feedback from motor factors, in surprising ways.[3] This point can be used as the basis of thought experiments that challenge the two natural and widespread assumptions. Consider each assumption and how it can be challenged.

2. The Input-output picture.

A. First, these cases disturb what can be called the *Input-output picture.* This is widely taken for granted as a general conceptual framework for thinking about perception and action, despite occasional disclaimers. It takes the common talk of 'perceptual input' and 'behavioral output' seriously. It conceives of perception as input from the world to the mind and action as output from the mind to the world.

Now there are two perfectly good distinctions we recognize: one is the distinction between perception and action. We use this distinction when we are talking at the personal level, about the contents of people's experiences or intentions. The other is the distinction between causal input and causal output. We use this distinction when we are talking at the subpersonal level, about functions or causal mechanisms: for example, when we distinguish afference and efference. The Input-output picture maps these two distinctions onto each other; it assumes that they coincide.

Cases like those just described undermine the status of this picture as a general conceptual framework for perception and action (see Hurley 1998 essay 9 for other examples and detailed arguments). They suggest by contrast that perception and action may be more intimately interdependent than that picture allows. Obviously, what we perceive will depend on what we do in a certain uncontroversial way: for example, what I see depends on whether I move around the corner or not. We can call this *instrumental dependence*: when perceptual experience depends on motor output, but only as a means to a different sensory input. This is allowed by the Input-

output picture. But these cases, and other related cases, raise the empirical possibility that perception may depend on motor factors in a different, more intimate and surprising way, that goes beyond instrumental dependence. They suggest that perceptual experience might vary with motor output even when sensory input doesn't change.[4]

Notice that challenging the Input-output picture does not require that the picture never apply. Indeed, it is about right for many cases. What should be challenged is the status of the picture as a general conceptual framework. To do this, we only need to show that *whether* it applies in particular cases is an empirical question, not that it *never* applies. The best empirical view of some cases may not fit this framework. The cases that fail to fit it, such as some involving spatial perception, may play so fundamental role in perceptual experience as to give these failures more general significance.

B. To illustrate these points, let's do a two-stage thought-experiment with the subject who has adapted to the color-divided goggles. Suppose we place this subject in front of a uniform white field, lit so that the same white light reaches his eyes whichever way he moves them.

While wearing the goggles, suppose he moves his eyes from left to right. There will be variation in retinal input as he moves his eyes from the blue to the yellow side. But we can suppose he would perceive the field to be white in both directions, because he has adapted to the goggles and has reestablished the color constancies that were at first disturbed by the goggles. So, while wearing the goggles he sees a uniformly white field as uniformly white, despite the variation in retinal inputs that results from his movements.

Now we can normally recognize objects of the same color in very different lighting conditions, which produce different visual inputs relevant to color--wavelength, reflectance, etc. (for various accounts, see Hardin 1988). There are various detailed views about how our brains achieve normal color constancy, but presumably it is by finding some invariant: some complex, higher-order relationship between varying visual inputs that doesn't itself vary.

The goggles introduce what at first seems to be noise into this process: eye movement in one direction or the other seem to produce arbitrary changes in color-related visual inputs. Whatever higher-order invariant among visual inputs had been the basis for recognition of a given color has been disturbed. However, no information has in fact been lost; it has merely been transformed by motor output. The information is still recoverable. A new invariant might now correspond to a given color, one that allows for the new relationship of motor output to visual input (cf. Rock 1966, p. 261ff; Taylor 1962, p. 258). But the relevant invariant might now be a function of both input and output signals--of their relations--, rather than just of inputs.

Note that the claim is not that this *must* be the right way to interpret this example. For example, a different interpretation could appeal to possible proprioceptive input, fed back from eye movements, rather than to feedback of motor output or efferent signals. Which interpretation is best is an empirical question. The thought experiment assumes one resolution of that empirical question. But, again, all that is needed to challenge the Input-output picture is that it *be* an empirical question.

Consider next what happens when the subject removes the goggles, still in front of the uniform white field. Now movements of his eyes bring about no change in retinal inputs. Yet

from what Kohler says about the aftereffects of adaptation to the goggles, presumably when he moves his eyes to the left he would experience yellow and when he moves his eyes to the right he would experience blue. So there would be a change in color experience with motor factors. This aftereffect of adaptation would be transient, since the subject would readapt to the absence of goggles. But even so it illustrates the empirical possibility of motor-dependence of perceptual content that would go beyond instrumental dependence, by obtaining even though sensory input is unchanged. So the motor factors here may not simply be means to different sensory inputs, either visual or proprioceptive. Motor or efferent signals may be playing more than a merely instrumental role in color perception. It is an empirical question about the case whether or not this is so. The fact that this is an empirical question challenges the Input-output picture's status as a general conceptual framework for thinking about perception and action.[5]

Notice that the kind of motor-dependence we are here concerned with is quite independent of behaviorism. It derives from the role of motor factors as *causes* within dynamic systems. It does not derive from any behavioristic collapse of the motor *effects* of perception into the perceptual states they are evidence for (see Hurley 1998, essay 10).

Notice also that the noninstrumental motor-dependence of perceptual content can be both a constitutive and a causal relation, as is the relation invoked by the externalist who holds that the contents of mental states depend constitutively on the causal relations between mental states and the world.[1]

C. Taylor distinguishes types of sensory information that normally stand in systematic relationships to motor output, like much spatial information, from types of sensory information that apparently do not normally stand in systematic relationships to motor output (other than verbal), like color information (Taylor 1962, p. 239-40; but cf. Noë; Myin on the motor-embeddedness of color perception). Motor-dependent perception may be less surprising where there are evident systematic sensory-motor relationships, as in spatial perception. It is more surprising for color.[6]

What Kohler's left-blue right-yellow goggles do is to impose artificially a crude but systematic relationship between sensory information about color and motor output--eye movements. This artificial sensory-motor correlation transforms but does not eliminate information about true color constancy, which is now a new function of, among other things, motor signals. Kohler's experiment suggests that color perception may be able to adapt to 'find' the new motor-mediated form of such information, despite the fact that there is presumably no natural or evolutionary basis for this particular kind of motor-mediation of color information. In the first stage of our thought experiment, perception of a given color came to depend on a higher-order neural invariant that presumably included an efferent or motor signal, even though normal color constancy presumably does not depend on such a 'mixed' invariant. After all, if color perception normally depends on invariant relationships between different types of sensory signals, there should be no reason in principle that it could not depend on invariant relationships of sensory signals to motor signals as well: there is no impenetrable boundary in the brain that rules out such mixed invariants. So there is nothing about color perception that makes it essentially resistant to motor-dependence. Motor-dependent adaptation is not in principle limited either to spatial perceptions in general (or to what is sometimes called the 'where' system).[7]

3. Duplicability Assumptions

A. The second challenge raised by the cases I started with is to assumptions of the duplicability in principle of internal physical states in different environments. Here is it important to distinguish general from specific assumptions of duplicability. A traditional view of the mind assumes both internalism and that it makes sense *in general* to suppose that internal physical states can be duplicated in radically different environments, or *general duplicability*. By contrast, when internalist and externalists describe Twin Earth cases and proceed to disagree about their implications, they make specific shared assumptions about the duplicability of internal physical states in particular pairs of environments, which are inverted or distorted relative to each other in specific ways.

Consider first how general duplicabilitiy assumptions feature in a traditional view of the mind. This view conceives the mind as inner and the world as outer, and supposes that the contents of the mind are in principle independent of the world. As a result, we might be mistaken about what the world is like--not just in the ordinary work-a-day way, but systematically deluded. For example, we might be brains in vats being manipulated by mad scientists, for inscrutable purposes of their own, to experience and believe certain things, regardless of what the world is really like. The world could be wildly different in many ways from how it seems to us, and we'd be none the wiser. This supposition of course generates skepticism. But whether that skepticism can be answered is not our topic here. It is rather the assumptions that underlie it.

Suppose that we are not dualists: we do not believe that there are separate mental and physical substances. So somehow the inner, mental realm, our experiences and beliefs, are embodied in physical states or processes, their contents carried by physical vehicles. Then, if we are at all prone to the thought that the inner realm of mental content is in principle independent of the world so that we could be systematically deluded, it will natural for that thought to involve two assumptions. These two assumptions characterize a traditional view of the mind.

The *first* is *internalism* about mental content: that our intrinsic internal physical states fix the contents of our beliefs and experiences. I couldn't have beliefs or experiences with different content if I were in the very same intrinsic body and brain states. If someone in another environment had intrinsic body and brain states that duplicate mine, she must believe and experience the same things I do. Mental content can only depend on differences in the world outside *instrumentally*: when such external differences are a means to altering physical states inside some boundary between the inner and the outer. An internalist holds that content is in this way autonomous with respect to the external world. Internalism is of course highly controversial and many philosophers reject it.

The *second* assumption is that it makes sense in principle to suppose quite generally that a person's internal physical states could be duplicated in systematically different environments. There must be enough discontinuity or slack in the causal relations between the outside world and our bodies that our bodies could, in principle, be physically duplicated in very different worlds--though of course there may be technical difficulties. This generalized assumption of duplicability made by a traditional view of the mind will be called the general duplicability assumption. Consider Earth and some different possible world, Inverted Earth, which is systematically different from Earth. The general duplicability assumption implies that despite those differences there could be twins going about their business in the same way in those different worlds: an

Earthling and an Inverted Earthling who are physical duplicates, with the same internal physical states and processes. The general duplicability assumption made by traditional views of the mind, unlike internalism, tends to be ignored.

TRADITIONAL VIEW:
---mind as inner, world as outer
---mind in principle independent of world
---systematic delusion possible

STANDARD ASSUMPTIONS OF NONDUALISTIC VERSION:
---*internalism about content*: internal physical states fix mental contents
(externalism about content denies this)
---*general duplicability assumption*: in general, duplication of internal physical states in systematically different environments makes sense in principle
(Cf. duplicability for specific Twin Earth scenarios, assumed by both internalists and externalists)

The thought that systematic delusion is possible emerges naturally from those two assumptions. The *general duplicability assumption* says that two worlds could be very different in many ways, while the twins' internal physical states were still duplicable. But *internalism* says that people's internal physical states fix the content of their beliefs and experiences. So the Earthling and the duplicate Inverted Earthling would have the same beliefs and experiences on Earth and on a wide variety of Inverted Earths, respectively. If their beliefs are about those aspects of the world that differ, then one of them will be systematically deluded.

An aside: Exactly how should the boundary around 'internal' states be drawn for purposes of assumptions about duplicability? This question will loom large in the following discussion. We will gerrymander it in various ways in evaluating the possibility of duplication in various cases. But notice that it will not do to show that duplication is possible within just any old boundary. The boundary around internal physical states and processes is constrained not just by the way a traditional view of the mind assumes the general possibility of duplication, but also by the way the debate about content assumes, more conservatively, the possibility of duplication in specific Twin Earth and Inverted Earth examples. The boundary within which duplication is supposedly possible must be of some philosophical significance in order to have a plausible role in the debate about content. We'll return to these points.

B. Many philosophers have assumed that the general possibility of duplication and of this kind of delusion makes sense. Some have tried to resist this possibility by denying internalism. Externalists allow that someone's mental contents can depend directly on the world outside, not just instrumentally via the effects of the outside world on someone's internal physical states. The contents of the mind, on this view, are not in principle independent of the world, but are context-

dependent. Two people in two different worlds might have different mental contents, even though their internal physical states were duplicated.[8]

Philosophers who resist the possibility of delusion by denying internalism do not deny duplicability in the specific cases they appeal to. Internalists and externalists share an assumption of duplicability for specific Inverted Earth scenarios; both depend on it to argue their positions (see Davies 1992, p. 27). [9] The assumption of duplicability in specific cases gives substance to their disagreement about whether content depends on the external world directly or merely instrumentally, since it's uncontroversial that content can vary with external states if internal physical states vary also.

Content externalists assume duplicability across specific Earth/Inverted Earth pairs. They do not depend on the general assumption of duplicability. But neither do they do explicitly reject it. Given the dialectical starting point in a traditional view of the mind, the ignored general duplicability assumption sets up a kind of default presupposition of duplicability in specific cases, and protects the latter from critical scrutiny.

A closely related point is this. Even if content is determined by external factors, it is normally assumed by both internalist and externalist parties to Inverted Earth arguments that the *vehicles* whose content is in question must be internal: even *content externalists* tend to be *vehicle internalists*. At the end of this article I'll consider the idea of *vehicle externalism* and its connection to the possibility of duplication.

To summarize: One way of rejecting the traditional view of the mind is to deny internalism while retaining the general duplicability assumption. A different way of rejecting it would deny the general duplicability assumption. Though generally ignored, this second aspect of the traditional view of the mind is challenged by the cases we began with. As just explained, setting up specific Inverted Earth cases requires only duplicability in those cases, not general duplicability. But to the extent the traditional view of the mind is the stepping off point for debates about content externalism, those debates inherit a presupposition of duplicability in specific cases, which has seemed natural and uncontroversial to both sides. Challenges to the general duplicability assumption bring the presupposition of duplicability in specific cases under critical scrutiny, and suggest a different kind of externalism, about vehicles of content, which pursues further the second way of rejecting a traditional view of the mind.

C. Natural as it is to make, the presupposition of duplicability needs scrutiny. Duplication is duplication within some internal/external boundary. But in dynamic cases, involving active subjects, the needed boundary can be harder to make out than we might suppose. Moreover, defending the possibility of duplication may require an unjustified assumption of temporal atomism about vehicles of content. I'll argue these points by comparing an Inverted Earth thought experiment where duplicability appears unproblematic to one in which there are difficulties in principle, not merely technical difficulties, for duplicability. These difficulties for duplicability stem from systematic patterns of sensorimotor contingencies and interdependencies between perception and action. [10] Such interdependencies can cross internal/external boundaries in complex ways that defeat the possibility of duplication.

The reasons duplicability assumptions have not been controversial include a tendency by philosophers to simplify cases of the problematic kind, to turn them into toy cases, in part by

ignoring complex dynamic feedback. These simplifications may seem innocent, but are not really. So it is important that we consider a realistically complex example to understand how the difficulties in principle for duplication arise.

It turns out that the points that tell against the input-output picture and against the general duplicability assumption are connected. The input-output picture makes it natural to assume general duplicability: if perception is causal input from world to mind and action is causal output from mind to world, it is natural to assume that mind is embodied within some boundary around inner states and processes, which should in principle be duplicable irrespective of what is going on outside that boundary. But the same factors that tell against the input-output picture, namely, the dependence of content on complex but systematic input-output relations and dynamic feedback, can also make duplication problematic.[11] In the problematic cases, in order to explain what is going on at the body surface, we need to consider not just the internal function that takes us from causal input to causal output, but also various functions from output back to input, some of which operate internally, such as efferent and proprioceptive feedback, and some of which operate externally, via the environment, such as visual feedback from movement. Such complex sensorimotor contingencies and dynamic feedback relations are glossed over when philosophers make their simplifying assumptions. But they can make difficulties of principle for duplication.

In the rest of this article I shall argue that the general duplicability assumption made by a traditional view of the mind is problematic. I'll also the implications of its demise for externalism. If general duplicability fails, then externalists should no longer presuppose that duplicability is unproblematic when they consider specific Earth/Inverted Earth scenarios. Each pair must be evaluated for duplicability. This could in effect broaden the externalist's argumentative base. Specific Earth/Inverted Earth pairs where duplicability does hold could support externalism about content, while the very failure of duplicabililty in other cases could support another brand of externalism, about the vehicles of content.

4. Red/green inversion.

A. Begin with a color inversion case in which duplicability seems unproblematic. Suppose that everything red on Earth is green on Inverted Earth and vice versa (see Block 1990). Moreover, people on Inverted Earth call green things 'red' and vice versa. The inversion is symmetrical. Otherwise, Earth and Inverted Earth are twin possible worlds. The specific duplicability assumption is that it makes sense here to suppose that a particular Earthling can have a twin or physical duplicate on Inverted Earth.

We assume that the Earthling and her twin Inverted Earthling go about their normal business in the same way and go through the same motions in their counterpart environments. It wouldn't be interesting to suppose, for example, that the Inverted Earthling's duplicate internal physical states have been rigged and she has then been put into a lead box to keep them from changing. That wouldn't serve the purposes of the debate about the role of the outside world in fixing content. Rather, the Inverted Earthling should interact causally with her inverted environment, establish functionally normal relationships with it. But then why aren't the intrinsic physical states of the Inverted Earthling affected by the differences on Inverted Earth? For all we've said so far, the color differences on Inverted Earth will make for differences in the internal physical states of the Inverted Earthling through normal sensory-motor channels. Her internal physical states when looking at a green object on Inverted Earth will not duplicate the Earthling's

internal physical states when looking at a red counterpart object on Earth. Different light, for example, will fall on the Inverted Earthling's retina, so correspondingly different neural signals will be sent to her brain.

This is where the causal slack between outside world and internal physical states is exploited. We next suppose that a change can in principle be made in the causal pathways that connect the outside world with the Inverted Earthling's internal physical states. This change just compensates for or cancels out the differences on Inverted Earth, so it vindicates the assumption of duplicability. Here, we need to add a reversing device to the causal pathways from red and green objects to the Inverted Earthling's internal physical states: for example, red/green reversing lenses are inserted into her eyes while she sleeps. These lenses themselves obviously cannot count as part of the internal physical state that is duplicated between Earth and Inverted Earth. The lenses must be outside the relevant internal/external boundary. (This looks like a harmless enough piece of boundary-drawing, but the gerrymandering that begins here ultimately turns problematic.) The causal pattern that held between certain internal physical states and red objects on Earth now holds between those same internal physical states and green objects on Inverted Earth. And vice versa. Once the Inverted Earthling has got these reversing lenses installed, her internal physical states will duplicate those of the Earthling, despite the color differences in the outside world.

At this point we could pursue the usual debate about whether the contents of experience can depend on the outside world directly, if internal physical states are duplicated, and not only instrumentally, via the world's effects on internal physical states. Does the Inverted Earthling perceive that limes are red on Inverted Earth, or suffer from the illusion that they are green? Must the content of her experience be the same as that of the Earthling's, given that their internal physical states are the same but their environments are different? This essay depends on no position either way on this question about content. Rather, it scrutinizes the assumptions of duplicability the opposed view about content share in specific cases. Now in this case, and others, the assumption of duplicabililty seems acceptable. At most it raises technical issues--how would the color-reversing lenses work?--not difficulties of principle. So the Internalist and Externalist about content can disagree about whether the duplicated internal states fix color experience, about whether systematic delusion is possible and skepticism makes sense.

B. But now return to the symmetrical character of the inversion. Notice that this is playing an important role in the possibility of duplication. Duplication may become problematic for nonsymmetrical inversions even in cases where action plays no particular role.

To see this, consider the search-and-replace fallacy. Most users of the search-and-replace facility on word processing programs learn the hard way that it is not always possible to undo the effects a document-wide search and replace command. Suppose I tell my processor to replace every occurrence of the word "red" with the word "green". Suppose I then decide I've made a mistake and tell it to replace every occurrence of the word "green" with the word "red". Will I be back where I started? Certainly not. I initially only replaced "red" with "green", not vice versa. This nonsymmetrical substitution lost information. It created extra members of the class of words "green" in my document, which were not distinguished from the preexisting members. If I now try to reverse the substitution, and replace "green" with "red", not just the extra instances but all the preexisting instances of "green" will be picked up, and I will end up with only "red" and the loss of the initial red/green distinction. We need to tag the preexisting instances in some way so

that we can keep track of them. So we learn to perform such search and replace exercises in stages that preserve the distinction between extra and preexisting instances, in case we want to restore the status quo ante.

The moral is that there is a source of trouble for the General duplicability assumption even before we reach the issues about action and dynamic feedback.[12] Consider a world in which all things that are red on Earth are instead green, but not vice versa. The stipulated environmental difference is such that earthly information about the red/green distinction is not simply transformed, but is lost. No reversing device can then cancel out the environmental difference, because the distinction it needs to work with in order to duplicate Earthly internal physical states is absent from the twin environment. So even apart from action, duplication may be restricted to environmental alterations that do not lose information in this way. The skeptical claim that our environment may be systematically different from the way we perceive it to be is less striking when qualified thus: "... so long as the environmental difference in question does not miss distinctions and hence does support the possibility of duplication." Mad scientists in the duplication business may be out of work for a wide range of nonsymmetrical environmental differences.

The search-and-replace problem for duplication is mentioned in order to distinguish it from the action and dynamic feedback problems for duplication, which are our main focus here. The latter appear to be independent of the search-and-replace problem, although action can give rise to versions of the search-and-replace problem, as we shall see (sect. 6 below). But little weight will here be put on the search-and-replace problem per se.

C. In some cases, duplication is typically problematic for reasons that involve action and dynamic feedback. To see this, let's switch from a 'what'-system case to some cases involving the 'where'-system. The suggestion is not that duplication is always OK for 'what'-system cases and never for 'where'-system cases; the distinction won't bear that much weight. But there is something typical of 'where'-system cases that makes trouble for duplication. This is the systematic interdependence of perception and action, with multiple channels of motor-to-sensory feedback, which is characteristic of cases involving orientation and localization, but not of normal color perceptions. The problem in a nutshell is: how do you, even in principle, get all the needed reversing devices outside the boundary of the states you're trying to duplicate? We will first consider the skin as boundary, and then retreat to the central nervous system. But if the reversing devices needed depend on multimodal information, in particular on the relationships between sensory and motor information, they are themselves central in character. So the source of the trouble is: given action and dynamic feedback, to renormalize functions for some environmental distortions may require essentially multimodal, central tinkering, and so be incompatible with the duplication of even central states. That is the hypothesis we here pursue, though we do not fully resolve it. Call this the antiduplicationist hypothesis.

5. Left/right inversion and dynamic duplication.

A. Orientation and localization are paradigmatic 'where' system functions. So, what would happen if, for example, we try to model a left-right reversal case on the red-green Inverted Earth case, but take action and dynamic feedback into account? Consider Seymour on Earth and Twin Seymour on Mirror Earth. Seymour rides his bicycle along down the street, wearing his

favorite biking goggles to keep the wind out of his eyes. These have normal lenses, not left-right reversing lenses. Every once in a while he takes them off in order to wipe his brow.

Can we establish for Twin Seymour normal functioning parallel to Seymour's while also duplicating Seymour's internal physical states in Twin Seymour? Instead of supposing that red things on Earth are green on Inverted Earth and vice versa, we postulate a symmetrical global left-right reversal of the Seymour's familiar environment. You can think of this as the Twin Seymour inhabiting the environment represented by the mirror image of Seymour's environment. If Seymour's environment were a glove, it would be turned inside-out on Mirror Earth. Further, on Mirror Earth, people use the word 'right' to mean left and vice versa. These are the stipulated initial environmental inversions.

The idea of a left-right reversed world needs some attention. Consider the vertical plane that is perpendicular to your face at your nose (the sagittal plane or 'nose plane'). If you mapped every point to the right of that plane onto the corresponding point to the left of it, and vice versa, you would have performed a structural reversal: made a left glove out of a right glove. Consider for example a very simple world, consisting of a rectangular array of four differently shaped figures, and its mirror reversal.

Now it might seem that the plane across which the mapping occurs has an important effect on the result. But in fact the plane across which you carry out the point to point mapping disappears in the final result: the result is seamless. The plane can intersect the world or object at any angle, or can be outside the object entirely. The effect of different choices of plane is simply to displace or rotate the resulting object or world, though all the results will be left-right reversed and, for a given object, can be superimposed on one another by turning them around or moving them until they align (see Locke 1977, pp. 4-6). Our Mirror Earth, then, is the result of structural reversal of Seymour's environment, aligned as it would be if the mapping had been carried out across Seymour's nose plane at some given time.

Now what do we need to do to establish duplication between Earth and Mirror Earth? Again, the would-be duplicates should both interact in functionally normal ways with their environments. And again, if we simply suppose that Seymour and Twin Seymour are located at the corresponding points on Earth and Mirror Earth respectively, looking the same way, their internal physical states will not be the same. Suppose a familiar spherical object is to Seymour's left when he is in a particular location and orientation, and so sends light to his eyes from the left side. The counterpart sphere on Mirror Earth will be on Twin Seymour's right when he occupies the corresponding location and orientation on Mirror Earth, and so will send light to his eyes from the right hand side. As a result, correspondingly different neural signals are sent by Twin Seymour's eyes to his brain.

B. So to get duplication, we again need to imagine some cancelling or compensating changes or reversing devices. Let's begin by introducing an input-side switch, left-right reversing lenses. Instead of wearing biking goggles with normal lenses, like Seymour, Twin Seymour wears biking goggles with left-right reversing lenses like those in the real Papert's goggles. These goggles must of course be outside the boundary within which we are trying to establish duplication. Otherwise they would immediately defeat duplication, since Twin Seymour's goggles are physically different from Seymour's.

Does this get us duplication? Does the sphere to the right on Mirror Earth produce the same type of internal states in Twin Seymour as the sphere to the left on Earth produces in Seymour? Well maybe, so as long as Seymour and his Twin just look and don't move. But a stationary perceptual system is a truncated and denatured one; it's not significant here that you can trick it. Once they get on their bikes, the complex yet systematic relationships between various types of feedback from movement make trouble for duplication. Further tinkering is required. There was nothing similar to make trouble for duplication in the color inversion case, since there is no normal, systematic feedback relationship between movement and visual signals for color. No input-side only switch will establish duplication in our left-right case, just because it involves the realistically complex feedback relationships that are typical of 'where'-system cases.

In this Mirror Reversal case we're going to need several input-side switches plus an output-side switch in order to reestablish normal input-output and intermodal harmony. Moreover, to do so here we will have to contract the boundary within which we are trying to achieve duplication from the skin to the boundary of the central nervous system. In this case we can achieve duplication by confining our reversal devices to the peripheral nervous system, where they are outside the boundary of duplication but safe from active interference by the agent. That is, though we need to apply reversals to several different modalities of information, we can tamper with each modality separately. The needed transformations are not themselves essentially multimodal, or central. But considering how to make duplication possible in this Mirror Reversal case will give us a clue to why duplication may not be possible in yet other cases, even for only central states. So the strategy is this: the Mirror Reversal case forces us to retreat to central states to make duplication work, but also suggests why this retreat will not always work: functional normality for an active Twin in an altered environment may demand changes even in the very central states we are trying to duplicate. This suggestion will be illustrated in the following El Greco World examples.

C. To give specific assumptions of duplicability in Inverted Earth arguments a fair run, we need to do as much as possible to reestablish intermodal harmony, including dynamic input-output harmony, in the altered environment. So our efforts below to make sense of motor reversal and so on are not mere epicycles but are essential to the argument. The point that will emerge is that even when we do our best, the requirement of input-output harmony may not be met, for reasons that relate to complex dynamic feedback.[13]

We can adopt slightly simplifying assumptions that abstract from tactile considerations, before considering the effect of the mirror-reversing goggles in relation to motor output and visual and proprioceptive feedback. Suppose both twins reach leftward. Seymour's hand will touch the sphere; Twin Seymour's will not. Even when they stand still in this position, different tactile stimulation will result in different neural signals, hence different internal physical states. Now we could at this point add a reversal device for tactile inputs, in order to duplicate static relations between tactile and visual inputs. There is no difficulty in principle with this for such a static multimodal case, so we won't pursue this route. The basic difficulty concerns dynamic feedback, and arises without taking account of tactile feedback in particular. In order to reach the more basic point without extra complications, we can abstract from the tactile dimension. Let's suppose the sphere distant enough to be out of reach, and the Seymour is only pointing to it. We will focus on relations between visual and proprioceptive inputs and motor outputs.

The important point here is this. Twin Seymour's left-right reversing goggles will alter the normal relationship between motor output signals that go with reaching left and visual input signals received from the hand moving leftward out there in the environment: between motor output and visual feedback. The reversing lenses will also alter the normal relationship between visual input signals received from the hand moving leftward through space and whatever innervation states go with felt position sense or proprioception: between visual feedback and proprioceptive feedback. Both these types of higher-order relationship between types of physiological states will differ between the twins, so they will not yet embody the same internal physical states.

Notice a pattern instantiated by this example that will turn out to be important: there are normal relationships between motor output and sensory feedback of different types, and there are normal relationships between different types of sensory feedback. Some channels of motor feedback have wider orbits, which may go through the environment (as does visual feedback from a hand moving through space, or from locomotion, as Gibson emphasized), while some have smaller orbits, which may stay internal (as with proprioception or copies of efferent signals inside the central nervous system). This type of complex structure of causal relationships is characteristic of what we can call complex dynamic feedback systems: systems in which causal feedback plays multiple roles and is both external and internal. If we track the causal arrow through time, a complex dynamic feedback system looks like a tangle or knot, centered on the organism and moving with it: a singularity in the field of causal flows.[14] The idea of such a dynamic singularity can be used to provide one alternative to the Input-output picture as a general conceptual framework for thinking about perception and action, and about how the personal and subpersonal levels are related. That is, the contents of both perceptions and intentions may in general be carried by the complex dynamic relations between inputs and outputs in such a system. The dynamic singularity conception would allow the Input-output picture to hold, as an empirical matter, in particular cases, but it does not assume it must hold.

Notice that on the dynamic singularity view, we cannot assume temporal atomism (see Consciousness in Action, essay 1): we cannot assume that perception over time, for example, must be like a cine projection in which each snapshot carries determinate content and the various momentary contents are strung together in a sequence. The subpersonal processes that carry content may be essentially continuous and dynamic.

D. So far, the reversing goggles plus linguistic switch have not cancelled out or compensated for the worldly mirror reversal. So we have not yet got duplication of internal physical states between Seymour and Twin Seymour.

Can we introduce further compensating output-side switches to get duplication? Let us begin by assuming the boundary within which we are trying to duplicate physical states is the skin. Enough of the debate about Content Externalism has assumed, either explicitly or tacitly, the skin as the relevant boundary to make it worth spelling out why it will not work for this case. We will then go on to assume instead that the relevant boundary is around the central nervous system.

Suppose that Seymour and Twin Seymour ride their bikes but keep their goggles on for the time being. Perhaps we can imagine a device that somehow creates a motor reversal, so that the very same type of motor output signal that on Earth goes with reaching left, on Mirror Earth goes

instead with reaching right.[15] No doubt there are technical difficulties, but we're used to those in Twin Earth cases. Suppose, for example, the motor reversal device is a force field of some kind that transforms and instantly inverts the bodily trajectory generated by internal physical states, without changing or distorting anything else, including visual feedback from movement. So the motor nerve firings and muscle contractions stay the same, but somehow the hand moves to the right instead of the left.

But there are various difficulties in understanding how this could happen. The result must produce visual feedback from movement in Twin Seymour that duplicates that in Seymour. If we just make Twin Seymour's left hand move rightward, his goggles will yield signals as from a right hand moving leftward on Earth, which is not what we need for duplication. Rather, we need signals like those of a left hand moving left, which the goggles will produce given a right hand moving right.

So a better way to proceed, given the reversing goggles, would be find a motor reversing device such that Twin Seymour reaches right with his right hand while Seymour reaches left with his left hand. But that means that not only the path the hand takes, but also which hand is used, has to be different between the twins. Then, when Seymour touches an object to the left, Twin Seymour will touch a counterpart object to the right; when Seymour puts on his left glove, Twin Seymour will put on his right glove.

Now if these differences count as differences in internal physical states, we are defeated already in our efforts to set up duplication. So some careful boundary-drawing is needed here: these differences in, first, direction of reaching and, second, which hand is used, must count as external, just as the presence of reversing goggles does.

But, third, we must get the motor-reversal device itself, the cause of these differences, outside the boundary around the internal states we are trying to duplicate. The motor-reversal device makes it the case that the same type of motor signal that goes with reaching left with the left hand on Earth goes with reaching right with the right hand on Mirror Earth. Could a force field outside the body, which leaves muscle contractions and motor nerve firings the same, change not just the direction of movement but also which hand is used? There's a dilemma: if we don't switch from one hand to the other, we don't duplicate visual feedback from hand movement, but if we do switch from one hand to the other, we don't duplicate muscle contractions and motor nerve firings.

Now let's be generous and assume for the sake of argument that this is just a technical problem. Perhaps we could, in principle, add another external force field to correct visual feedback. In a moment we'll return to the problem of duplicating normal internal relations between motor output signals and visual feedback signals. But note first that similar questions arise for relations between motor output signals and proprioceptive feedback signals, and for relations between visual feedback and proprioceptive feedback signals! So, even having abstracted from tactile considerations, we've got three sets of relationships to restore. Suppose we've solved the duplication problem for motor-visual relations: we have preserved normal internal motor-visual relations by adding the motor reversal device, etc., to the reversing goggles on Mirror Earth. This will still not preserve normal relations between visual and proprioceptive feedback on Mirror Earth, and moreover it alters the normal relation between motor neural signals and proprioceptive feedback. Twin Seymour would really be moving his hand to the right, not to

the left, so that proprioceptive signals would be like those that on Earth would go with movement to the right. But Seymour would really be moving his hand to the left, so proprioceptive signals would be those normal to reaching left. So proprioceptive signals also have to be reversed on Mirror Earth if we are to duplicate their Earthly relationships to motor neural signals and to visual feedback signals from movement. Surely we cannot coherently suppose this to be done wholly outside the body.

Now in real life, when reversing goggles introduce a conflict between visual and proprioceptive information, there is a normal tendency for vision to dominate, and for proprioceptive content or felt position to adapt to the content of visual perceptions (Harris 1980).[16] But that is no help to us in establishing duplication for proprioceptive signals: proprioception cannot be assumed to adapt in this way without physical states of the central nervous system altering. The proprioceptive reversal we need for duplication must be accomplished in a way that is external to the internal physical states we are trying to duplicate.

To summarize where we we've got to: to make out duplication from the skin inward for the left-right reversal case we need to do at least two things, in addition to inserting reversing lenses. First, we need a motor reversal device, and second, we need a proprioceptive reversal device. It is difficult to see in either case, but especially in the second, how this prescription be filled except by tinkering inside the body. But if we do that, we are reaching inside the boundary we initially assumed, namely the skin, and changing the internal physical states we are supposed to be holding constant. And that would defeat the effort to show that duplication in this case makes sense.

At this point there are two possible escape routes to consider. The first is to suppose the laws of nature are different on Mirror Earth, and the second is to give up trying to establish duplication for the whole body, and retreat to trying to establish it only for the central nervous system. The first escape route fails. The second escape route leads to an obstacle, but it can be avoided for this case. However, we get a suggestion of how it might not be possible to avoid similar obstacles in other cases.

E. First, then, perhaps we should get round the problem by supposing the laws of nature differ on Mirror Earth. Suppose that the very same motor nerve firings and muscle contractions occur in Twin Seymour as in Seymour, but in one case the laws of nature are such that the right hand moves rightward, and in the other case the laws of nature are such that the left hand moves leftward. Similarly, we suppose that these two different movements produce, via different laws of nature, the very same set of proprioceptive signals in the two cases.

Of course, if the laws of nature are different on Mirror Earth, they are different inside the skin as well as outside the skin, so internal physical dispositions are different. Isn't this an internal difference itself, a difference in what we're supposed to be holding constant? Someone might reply: "Well, the laws may differ between Seymour and Twin Seymour, but their internal physical states may still be duplicated. Only the bare physical states, not the governing laws of nature, count as within the duplication boundary. We exclude the laws, just as we excluded the reversing goggles, from what we're trying to duplicate." But the gerrymandering has now gone too far--at least if you hold the plausible view that a series of physical states can't hold constant while the governing laws of nature vary, because the individuation of physical properties and

events depends on the laws they figure in (see Hornsby 1985, p. 454). So this way of defending duplicability doesn't work.

F. The second way of defending duplicability is to move the boundary inward, give up the aim to duplicate except inside the central nervous system (CNS). We will return to the question of exactly how the boundary around the CNS should be determined for these purposes. Perhaps we could assume that the reversal devices on Mirror Earth operate only through the peripheral nervous system, outside the boundary of the CNS, so that all motor nerve firings and muscle contractions would not need to be duplicated. So, for example, Twin Seymour's motor reversal device takes motor signals that emerge from the central internal states and switches them across to the other side on the way out, with the result that the right hand moves rightward instead of the left hand moving leftward: undoes the brain's normal contralateral control. And his proprioceptive reversal device takes proprioceptive signals coming in from the right arm and switches them across to the other side on the way into but just outside the CNS. Similar left-right motor and proprioceptive switches are made for all other peripheral motor and proprioceptive pathways. Notice that while various modalities of information are transformed, each transformation is unimodal. Now the relevant inner-outer boundary is not around the person, is not the skin. We are no longer trying to duplicate all nerve firings and muscle contractions, but only physical states inside the boundary of the CNS.

So far so good. A simple, uniform transformation that is peripheral rather than central in character is applied to information in each modality. Together, these various peripheral transformations should work to permit functional normality with central duplication, even for bike riding and many other motor activities. For example, if both twins stand up and tilt their heads, one to the right and the other to the left, central motor signals will be duplicated (because of the left-right reversal of peripheral motor nerves after they leave the CNS). If Seymour tilts his head to the right, the CNS will receive proprioceptive information as of a head tilting to the right. And light from Seymour's feet will enter his two eyes from the direction of their right corners. If Twin Seymour tilts his head to the left, the CNS will receive proprioceptive information as of a head tilting to the right (because of the left-right reversal of peripheral proprioceptive nerves before they enter the CNS), so central proprioceptive signals will be duplicated. And light from Twin Seymour's feet will enter his two eyes from the direction of their right corners (because of the left-right reversing goggles), so central visual signals will also be duplicated.[17]

But at this point recall that it is Seymour's habit to remove his goggles occasionally as he bicycles in order to wipe his brow, and then to put them back on. Let us remove the constraint on his activity that prevented this. Twin Seymour must of course do the same (using the opposite hand) or duplication is defeated. But is duplication preserved if he does remove his goggles? No.

Twin Seymour will, when he first removes his goggles, be in the position of someone who puts left-right reversing goggles on. Normal central relations between his motor, visual feedback and proprioceptive feedback signals will now be altered and intermodal and input-output harmony will be lost. Of course, he may adapt to the new set of relations and reachieve functional normality, but the complex multimodal and input-output recalibration this would involve is paradigmatically a central process, so would not preserve central duplication.

Indeed, Twin Seymour may even adapt to two sets of such relations, as the real Papert did. So Twin Seymour might be able to achieve functional normality relative to Seymour on Earth: he

might be able to ride his bike along in Mirror Earth, taking his reversing goggles off and on and not noticing any difference. But this even more complex neural achievement would involve even more complex multimodal central processes, which there would be no call for in Seymour's case, since his goggles do not reverse anything. The switch between the two sets of relationships cannot be performed peripherally, that is, by applying separate unimodal transforms to each modality of information; it depends essentially on intermodal and sensory-motor relationships. So again central duplication would fail.

Now there are two reactions to this scenario that must be considered. First, isn't allowing Twin Seymour's actions to interfere with one of the artificial peripheral reversing devices in some sense cheating? Well, why should it count as cheating? After all, the action parallels Seymour's normal action, and its intended object occurs well outside the boundary around the CNS within which we are trying to achieve duplication. And the hypothesis we are considering is that feedback from action may defeat duplication, even central duplication.

But we don't have to argue this first point because it won't save the example from the second point. This is that we can easily avoid this problem here. Simply move the reversing lenses into the peripheral nervous system: put them surgically into Twin Seymour's eyes, so that they are safe from his interventions and are not affected if he puts his nonreversing goggles off and on just as Seymour does. That way the reversing lenses are in a safe zone: out of the reach of anything that might count as a normal action or normal functioning, but not inside the boundary around central states. Now separate transformations in each modality will again to the needed work. The multimodal sensory-motor recalibration demanded by putting the goggles off and on will not be required. Central relations among motor, visual feedback, and proprioceptive feedbacks signals will again be duplicated.

The conclusion of the discussion of Mirror Earth is that the retreat to the CNS does preserve the possibility of dynamic duplication in this case, so long as it insulates the various peripheral neural reversing devices from interference by normal actions that requires central processing to neutralize. But the discussion also raises the question of whether such insulation from action is always possible. Can we find a case in which the safe zone strategy will not work? Programmatically, what would be needed would be a situation in which something analogous to putting the reversing goggles off and on is a normal activity and so cannot be avoided by the safe zone strategy.

6. El Greco Worlds.

A. Shape is a 'what'-system property, but shape distortion can have consequences for orientation and localization and hence for the 'where' system. Consider a mixed 'what'/'where' version of Inverted Earth in which all shapes are stretched vertically, in the direction opposite to the pull of gravity--an El Greco world (compare the circle/ellipse case in Davies 1992, 1993, which prompted my consideration of El Greco worlds). Our twins are Domenikos on Earth and Twin Domenikos on El Greco Earth: Dom and Twin Dom for short. Suppose that the duplication boundary is around the CNS, not the whole person. Since your body is among the things you perceive, and since we're not bothered about noncentral bodily differences between our twins, let's apply the El Greco distortion to the body itself, but excluding the CNS. Our strategy for pursuing the antiduplicationist hypothesis will now be simply to consider various types of activity by Dom and Twin Dom for which assumptions of duplicability need hard scrutiny.

There are at least two ways of specifying the El Greco distortion. First, it could be rigid. At a given moment t0 when Earth and Dom physically duplicate El Greco Earth and Twin Dom, a one-off proportionate vertical stretch is applied to all objects on El Greco Earth including bodies (those of others on El Greco Earth as well as of Twin Dom), but excluding the CNS. The immediate result at t1 is that everything, say, is twice as tall on El Greco Earth as their counterparts are on Earth. But objects keep their shape even when their orientation relative to gravity is subsequently changed. So the counterpart of a blue earthly sphere is a taller blue eggish thing set up on its end. But if Twin Dom tips the egg over so that is lies lengthwise on the ground, its shape does not change.

Will the one-off stretch of environment and body automatically provide Twin Dom with central relations between visual, proprioceptive, and motor relations that duplicate those of Dom? Here are several of the many scenarios that need scrutiny.

First, suppose that the relative orientations of Twin Dom's body and some other objects, such as the blue eggs, have not changed since t1. Suppose at t1 each Dom faces three spheres arrayed in front of him at equal horizontal distances. One is the same height as his eyes, one is three times their height, and one is eight times their height. Each has a red spot on top. Twin Dom faces three eggish things, which are the same horizontal distance from him as Dom's spheres are from him, since no horizontal distortion is applied. All the eggish counterpart objects on El Greco Earth are twice as high as their counterpart objects on Earth, but their heights relative to Twin Dom's eyes are the same: 1 to 1, 1 to 3 and 1 to 8. (Illustrations are omitted here; see Consciousness in Action, ch. 8, for figures.)

This means that the line from Dom's eyes to the top of the 1 to 1 object will parallel ground level, and this also be true for Twin Dom. But the line from Dom's eyes to the top of his 1 to 3 sphere will not be at the same angle as the line from Twin Dom's eyes to the top of his 1 to 3 egg: the angle of Twin Dom's line will be steeper by some percentage p.

Moreover, the line from Dom's eyes to the top of his 1 to 8 sphere will not be at the same angle as the line from Twin Dom's eyes to the top of his 1 to 8 egg: the angle of Twin Dom's line will again be steeper, by some percentage q.

Furthermore, the percentage differences in angles will not be the same: q will not equal p. The percentage differences in angles will be smaller when we compare the 1 to 8 objects than when we compare the 1 to 3 objects (approximately 70% vs. 100% in these sketches). So a different percentage transformation of angle is needed for objects whose twin-relative heights differ.

Suppose Dom points to each red spot in turn, and Twin Dom must do the same. Can central states be duplicated? Consider visual input first. The angle of incidence of light coming into Twin Dom's eyes needs to be transformed in a way that is a function of the twin-relative heights of the red spots. Will the vertical stretch applied to Twin Dom's retina itself provide just the needed variable transformation of angle of incidence of light from the various red spots? Not clear. It may depend on whether Dom is allowed to move in ways that alter the relative orientations of the vertical axes of his body and the various spheres.[1] Or can some other noncentral transformation based purely on visual input do this work? Not clear either. Relative height can be figured as a function of actual angle of entry of light plus distance, which is in turn a

function of visual vergence or stereoptic disparity information. That sounds a bit central. Here are several points at which duplication looks problematic and needs further scrutiny.

But suppose this the needed noncentral visual transformation is possible, at least for initial relative heights. Visual feedback from the pointing arm will automatically be corrected along with the angle of incidence of light from the red spots, so visual matching of arm and spot positions will automatically duplicate the angles of the Twin's pointing arms, hence will duplicate central relations among visual, motor and proprioceptive information. However, more action makes more trouble. If at t2 Dom turns his spheres sideways and Twin Dom turns his eggs onto their sides, so as to change their orientation, a different transformation will be needed on Rigid El Greco World. The girth of the objects has not been distorted, but it now determines how far they and their red spots extend up from the ground, since they are placed sideways. So we don't want a correction applied here. Moreover, the initial height distortion of the objects now makes for horizontal distortions in Rigid El Greco World, which do need to be corrected. But we don't want to apply the horizontal correction to all horizontal dimensions, since objects whose orientations have not changed will not be horizontally distorted. Otherwise, we run into the search-and-replace problem: we lose distinctions. So action can generate instances of the search-and-replace problem. The needed distinction-sensitive correction at t2 must be a function of initial orientation at t1, so will involve memory. That also sounds a bit central. So here is another point at which duplication looks problematic and needs scrutiny.

Can we imagine a device that keeps track of changes in orientation between Twin Dom and other objects by monitoring the relations between his visual, proprioceptive and motor signals over time and, based on this multimodal stored information, then makes selective alterations in these various signals at each point so as to duplicate Dom's central states at that time? Of course, it depends on what we mean by 'central'. Suppose a mad scientist trains an artificial neural network to make and remember these multimodal comparisons and to make the needed selective alterations. If he inserts it and connects it up inside Twin Dom's meninges (the membranes around the brain and spinal cord), his physical and functional states within this anatomical boundary have been altered, so duplication fails for this interpretation of the central boundary. But if the network is instead inserted into Twin Dom's elbow, does that make a difference? Would a bit of neural circuitry normally inside the brain that somehow became extruded from the meninges but continued to function no longer count as central? Surely the philosophically relevant point is that the network is performing a central function by selectively altering information based on the relationship between other, convergent and recurrent streams of information from various modalities. So the network itself constitutes an alteration in Twin Dom's central states. Sheer location relative to an internal bodily membrane cannot provide a significant criterion of centrality for philosophical purposes.

So again it is not clear that duplicability of central states does make sense in principle for the Rigid El Greco case, once a normal range of actions are permitted. At least the answer to the question is far from obvious, when we consider what is involved.

B. Perhaps by making the distortion rigid we make the antiduplicationist's work too easy. Consider instead a plastic El Greco distortion. Suppose a kind of anti-gravity force field is installed on El Greco Earth, just outside the CNS (though the laws of nature are the same). Things, including the body outside the CNS, are stretched vertically from a baseline, the surface of the earth, whatever their orientation and as they move, so that they change shape as they change

orientation. The plastic distortion of the body and other objects as they move may deal with the some of the problems for duplication that arose for in the Rigid El Greco case. But plasticity will give rise to further problems of its own.

Recall the tactic suggested by Papert's goggles. We set out to find whether some functionally normal action might naturally play a switching role, similar to the role of Twin Seymour's putting his reversing goggles on and off. If so, we could not protect central states from interference by this action by relocating the switch in a safe zone, out of reach, while preserving functional normality.

Since normal bodily actions might be significant in relation to the speed of sound, but not in relation to the speed of light, let's consider audition instead of vision. Suppose Dom on Plastic El Greco World is playing with a whirligig. A whirligig is a kind of yo-yo that incorporates an electronic device that emits a musical tone at a certain fixed frequency when it is switched on. This device is at the end of a string. Dom is too young to make the yo-yo work properly by winding and unwinding the string around its spool. But he wants to play with his yo-yo, because his older sister is playing with hers. So he simply extends his hand out in front of himself and starts the tone-emitting device swinging round and round on the end its string. It traces a close-to circular path, all the while emitting a tone of a fixed frequency.

Meanwhile, what is happening on Plastic El Greco World? Twin Dom's whirligig is subject to the plastic vertical distortion. So as the string swings around, it gets longer as it approaches a vertical alignment and shorter as it approaches a horizontal alignment. It therefore traces an elliptical path. Moreover, the sound emitted by the tone-emitting device on the end of the string is subject to a doppler effect, as sound waves from the moving object are themselves subject to vertical but not horizontal distortion. Frequency and pitch of sounds reaching Twin Dom's ears will be distorted dynamically, as a function of the direction of movement of the whirligig. In effect, playing with the yo-yo in this way turns the need for auditory correction of the environmental distortion on and off, just as putting the left-right reversing goggles on and off turned the need for visual correction of the environmental mirror-reversal on and off. But in this case the switch cannot be hidden in a safe zone; it is built into Twin Dom's normal play behavior, given the environmental inversion.

The needed correction in auditory signal will be a function of direction of movement of the whirligig. It might be made be keeping track of motor signals and making selective auditory adjustments. But that sounds like a pretty central function. Could the selective transformation in auditory signal needed for duplication be achieved by a unimodal transform, based on auditory information alone, and without running into search-and-replace problems? Again, not clear. Also, acceleration and deceleration effects on the hand require motor and proprioceptive compensation. Duplicability here asks a great deal of noncentral, unimodal processes. Does it demand too much in principle?[2]

At this point it may be objected by a defender of duplicability that we should not after all be aiming to duplicate internal states and processes identified <u>physically</u>, but rather should be aiming to duplicate internal processes identified <u>functionally</u>.[3] But this objection faces a dilemma. To duplicate functions in the whirligig case, you must identify functions externally. For example, it might be the case that Twin Dom can, by moving his arm in a horizontally aligned ellipse while he swings the whirligig, produce a circular orbit, thus avoiding doppler effects. In doing so, his

internal physical processes might differ from those of Dom, but there is a sense in which they would be functionally the same. Both sets of movements would function to give the whirligig a nice circular orbit. But the trouble here is that the identity of function is determined externally, by the circular shape of the whirligig path produced on both worlds. On the other hand, if you identify functions strictly internally in the whirligig case, they are not duplicated between Dom and Twin Dom for reasons we have already considered. For the antiduplicationist arguments above have already assumed a functional understanding of central processes. It is precisely central processes understood as identified by internal functions that must differ between Dom and Twin Dom to accommodate dynamics. (The functional reading of centrality is considered explicitly in the next section.)

Notice the role of action in bringing duplicability into question in these examples. For a static subject most of the issues raised would not arise: a fixed set of devices, some altering inputs from various senses, some altering output, could in principle achieve duplication (though this remark would not apply to static versions of the search-and-replace problem). This indicates that the problem for duplicability does not derive merely from the relations between the inputs to different senses considered statically, or from primary qualities considered as statically accessible to more than one sense. It is not just multimodal harmony in general that is problematic for duplicability, but input-output harmony in the presence of dynamic feedback in particular.

Notice also that the difficulties for dynamic internal duplication are closely related to difficulties for temporal atomism (see Consciousness in Action, essay 1). The assumptions we make about the temporal boundaries of subpersonal processes that carry content have implications for the spatial boundary around duplicable processes. If we assume temporal atomism and suppose that content must be carried by momentary time slices of subpersonal processes, then we may be tempted to view dynamic feedback as merely creating extraneous interference with internal duplication at the next point in time. On such a temporally atomistic view, only moment-by-moment duplication is required. Each subpersonal snapshot carries determinate content. As long as duplication is possible in principle for each snapshot, considered separately, there is assumed to be no difficulty in principle with supposing the separately duplicated snapshots to be strung together, as in cine film, to capture dynamics. The mad scientist can do the stringing together of snapshots by hand; the dynamics are not intrinsic to the vehicles of content. The possibility that the very subpersonal processes that carry content might themselves be essentially continuous and dynamic, and so not amenable to snapshot duplication, is simply ignored. So the threat to dynamic duplication fails to register. But once made explicit, these temporally atomistic assumptions cannot be justified.[4]

Virtual reality helmets can keep track of some movements and alter a visual scene accordingly, creating the illusion of presence in another environment. Virtual reality is created, according to Richard Held, by exploiting ambiguities or many-one relationships between outer and inner states: that is, by duplicating the internal physical states that would occur if the simulated environment were real. But this becomes increasingly difficult as freedom of movement is increased: ultimately, it requires the creation of a duplicate environment. Freedom of movement progressively eliminates ambiguity, to a vanishing point, and thereby sets limits in principle to the creation of virtual reality.[5] These are the same limits in principle that make the General duplicability assumption problematic.

7. Centrality.

A. Does the assumption of duplicability or the antiduplicationist hypothesis have the burden of proof in particular cases? Our examples suggest that this question is bound up with the way the central physical processes whose duplication is in question are understood. How is the boundary around such processes determined? Our worries about gerrymandering now come home to roost with a vengeance.

The duplicationist assumes the duplicability in principle of internal physical processes within some philosophically significant boundary, in various distorted environments. For some environmental distortions, as we've seen, this boundary cannot be the skin and the duplicationist must retreat to central processes. Now dynamic duplication requires that the Earthling's normal actions and functions be restored to the Twin, despite the environmental distortion. If one acts normally and the other idles or malfunctions, duplication is threatened. But given an active agent, why suppose that functional normality can be achieved without intervention into central physical processes? The adaptation required for the Twin to achieve functional normality may also defeat duplication. Nothing guarantees that the needed complex adaptation can in principle be exported from central processes? After all, it seems no accident that the plasticity of active human beings in adapting to distorting goggles and so on involve central processes. The thought that such adaptation can necessarily be achieved through peripheral processes is prima facie puzzling.

The duplicationist needs some independently motivated and philosophically significant conception of the central physical processes he is trying to duplicate that allows him to mount an argument to the effect that noncentral transformations can in principle do the needed work in particular cases. The antiduplicationist needs to show that a conception of centrality that would support such an argument is not independently motivated, lacks philosophical significance, or is ad hoc. Because the notion of centrality that will serve here is constrained, the duplicationist carries a burden of proof, case by case. He cannot just wave away difficulties as merely technical. Some technical difficulties may in principle require central resources to address. We need to know why we should believe this is not true in particular cases.

B. These considerations can be illustrated by several proposals about centrality. The first illustrates some difficulties with ad hoc boundaries. Suppose the boundary around central processes is fixed by stipulation. The duplicationist could stipulate that the boundary goes around whatever bodily and brain states can be duplicated for a given environmental distortion. But if we apply this stipulation case by case, then the boundary will be relativized and may vary from case to case. For example, we didn't need to put the boundary as far in for the color reversal case as for the left-right reversal case. In any given case we might find some boundary within which internal physical states are constant, even despite adaptation. So we might find an area of the brain not at all involved in that particular adaptation and put a boundary around it. But there would be no motivation for doing so apart from the wish to save duplication. The resulting boundary would have no general significance or interest. There is no reason in principle to suppose the boundaries arrived at in this way for different environmental inversions would tend to coincide, or even nest.

However, we could instead stipulate that the boundary goes around whatever bodily and brain states and processes can be duplicated for any environmental inversion. Take Earth 1 and Inverted Earth 1; find the states that can be duplicated. Do the same for Earth 2 and Inverted Earth 2, and so on. Then take the intersection of the states that can be duplicated in each case, and

put the boundary around the states in this intersection. But this stipulation fares no better than the relativized version. First, why should we assume that there must in principle be any states in this intersection? Second, any states that were identified by this stipulation would not be significant for purposes of the debate about content. Consider an Inverted Earth thought experiment in which these states are constant for some given environmental inversion, but perceptual contents vary. That would provide no support for Externalism. There would be no reason to assume the differences in content depend on environmental differences as opposed to differences in the body or the brain that just happen to be outside this boundary. For example, the different perceptual contents might supervene on distributed neural processes that fall partly inside and partly outside this boundary.

C. Secondly, we could identify central states and processes anatomically, by reference to the meninges, the membranes that surround the brain and spinal cord. There may be good empirical reasons why processes that have certain functions are normally contained within such membranes. But it is their function and not the membranes that gives these processes their significance as central. A chip bearing an artificial neural network that performed the function of some damaged area of the brain could certainly count as significantly central if it was implanted within the boundary of the meninges. But what accounts for its centrality, its function or its location? Surely its function. If it was more effective and functional to locate the chip somewhere in the body outside the meninges, that would not alter its centrality. Moreover, if a normal bit of the brain was for some reason extruded from the meninges but continued to perform its normal function, it would still count as central. If, for example, multimodal processing of signals from eye movement as well as from retinal input occurred in the eye rather than the brain, such processes would still count as central.

So we should place the boundary around central physical states functionally. Intuitively, we want our boundary to capture those physical states in which central processes occur, that is, in which various modalities of sensory and motor information converge and feed back to and transform one another, in such a way that the complex dynamic relations between them selectively generate further information and signals.[6] This functional conception of centrality in terms of multimodal convergence, recurrence and selective transformation of signals has intuitive motivation independent of duplicability assumptions. But this is the conception of centrality we employed in pressing the antiduplicationist hypothesis for the examples in previous sections. For some types of environmental distortion, duplication of the central states of active creatures may not in principle be possible, since the required transformations of information may themselves be central in this sense, and so defeat duplication. To defend duplicability assumptions in specific cases we'd have to consider whether for active creatures the solutions to the information processes problems posed by particular environmental distortions would amount to tinkering with central states.[7]

8. Some reactions.

The General duplicability assumption has been challenged, if not definitively disproved. Let's now consider briefly some possible reactions to this challenge. We'll consider several possible dismissive reactions and a diagnosis, before going on to speculate about what the implications might be if the challenge to the General duplicability assumption were to succeed.

A. First, someone might say: "Your cases are exceptional. Normally duplication is not problematic." But when we look at the details, the opposite view is more likely to be right. There are indeed cases for which duplication is not problematic, such as the color inversion case, where there are no systematic sensory-motor relationships. There are also cases for which at least central duplication can probably, though not obviously, be vindicated, such as the left-right reversal case. But action makes at least prima facie difficulty for duplication in cases where there is systematic motor to sensory feedback, as our various examples suggested. Cases involving localization and orientation, or the 'where' system', typically involve such feedback.

If this is right, is it possible to contain the doubts about duplication to pure 'where' system cases, or will they tend to spread to other cases? The type of 'where'-system spatial content we've appealed to is basic to having a unified perspective or a point of view as a perceiver and agent, a sense of being a self present in the world, and hence to having a mind. As a result, even if duplicability is safe in some cases, such as the pure color inversion case, and some "where"-system cases as well, problems with duplication may tend to spread. For example, shape is a "what"-system property, which was distorted in our El Greco examples. But there were consequences for localization and orientation, and these made duplication problematic.

B. Another dismissive reaction to the difficulties for duplication would be to insist, in traditional philosophical style, that these are merely technical difficulties, not philosophically significant. We can just blur our vision and imagine them away for philosophical purposes.

But there is no reason to believe this. The potential self-defeatingness of efforts at duplication is hidden by the complexity of realistic cases and the resulting tendency to simplify for philosophical purposes. For example, the examples of duplication involving input-output harmony in the literature tend to be toy examples, which involve merely ballistic movement and avoid the very issues about dynamic intermodal feedback relations, and the resulting temporal spread of vehicles of content, that cause the problems (see Davies 1993). But anything toy enough to make General Duplication unproblematic is difficult to take seriously as an agent with a contentful perspective. And once General Duplication falls, duplicability in particular cases needs scrutiny. Introduce enough dynamic intermodal complexity to make it plausible to attribute a unified perspective with characteristically interdependent perceptions and intentions, and you may well make trouble for duplication in particular cases, as in the examples we considered. Those were not toy, but were relatively realistic in their complexity: at least there was no occasion to doubt the creatures in question had perspectives. The change in complexity makes a difference of principle, rather than merely creating technical problems to be overcome by a little science fiction.[8]

In a nutshell, the point is this: If the Earthling we are trying to duplicate is toy enough, he may be duplicable without issue, but his mentality is threatened. So content will fail not just where the environment is disturbed, on Inverted Earth, but on Earth as well. It is not interesting to claim that a duplicate of a toy Earthling that lacks mental states also lacks mental states, and this claim does not support Externalism about content. If the Earthling's complexity supports mentality but also threatens the possibility of duplication in the Twin Earthling in particular cases, then we may not be able to restore normal activity and functioning to the Twin in certain inverted environment while also achieving central duplication. But if we don't restore normal activity and functioning to the Twin, the deviation in activity between the Twins will itself defeat duplication,

so long as we maintain a dynamic view.[9] Toy Twins of nontoy Earthlings are not dynamic duplicates.

C. A different reaction, which may have more merit, is in effect a diagnosis of the situation. As already indicated, the Input-output picture makes it natural to assume duplication. If perception is conceived as input from world to mind and action as output from mind to world, then it is natural to assume that the mind is embodied in whatever stands between such input and output: in the body, or at least the CNS. Now perhaps the Input-output picture should be replaced with a conception of both perceptual and intentional contents as functions of relations between inputs and outputs within a complex dynamic feedback system. Such a two-level interdependence view would provide an alternative general framework for conceiving of the interdependence of perception and action and of the relations between the personal and subpersonal levels. This alternative framework would invoke subpersonal dynamic singularities that can include external as well as internal feedback loops, and hence can spread across internal/external boundaries.[10]

But rejection of the Input-output picture, and recognition that perceptual content depends on complex input-output and dynamic feedback relations, are not easily reconciled with the usual assumptions about the significance of Twin Earth arguments for our conception of relations between mind and world. As we've seen, complex feedback brings into doubt the set-up condition of the Twin Earth strategy, namely, duplicability. When there are systematic interdependencies between perception and action, it may not make sense to suppose internal physical states can be duplicated despite normal interaction with a systematically altered environment--even if we retreat inwardly. The internal states of an active creature are heavily constrained by multiple channels of motor feedback, some of which go through states of the environment (as in Gibsonian visual feedback from movement). Such complex dynamic feedback takes up the explanatory slack that duplication exploits. An environmental reversal may well require a combination of compensating reversals on various channels: a compensating switch on one feedback channel, with a wider, external orbit, may make trouble on another channel, with a smaller, internal orbit. It may not make sense to suppose that all the needed compensating switches can be made simultaneously, without reaching right into the internal physical states we are trying to duplicate. In this way, complex dynamic feedback systems knit nervous systems causally into their environments.

9. Externalism and explanation.

A. The preceding diagnosis suggests some further speculations about the significance of the threat to the General duplicability assumption in the context of rejection of the Input-output picture. However, what has been said so far does not depend on these further speculations.

Is the Twin Earth or Inverted Earth strategy of argument, which depends case by case on assumptions of duplicability, the best way to focus issues about relations between mind and world? Maybe these assumptions are limited, and we need in addition a different, if complementary, way of challenging the traditional conception of the mind as inner and in principle independent of the world, and of expressing a broadly Externalist view of the mind.

The modal Externalist who relies on the Twin Earth strategy of argument may reply that he only needs environmental determination of content to be possible. So long as the issue arises

and goes his way for some cases, he will be vindicated, even if it doesn't arise for other cases. This reply is fine as far as it goes, but it doesn't pursue the point far enough. If duplication is in principle problematic in certain types of case that are basic to what it is to have a perspective, doesn't this itself suggest a more radical kind of Externalism, about the vehicles of content, which may be complementary to Content Externalism?

It may also be said: the failure of duplication deprives us of a way of testing the Externalist, world-involving view of the mind. But it still might be true that mental states depend in principle on the world. The explanatory work that mental states do, and so what it is for a creature to have such states, may essentially involve its environment. Perhaps the alternative formulation of Externalism should turn on the explanatory work done by external states rather than depending on the Inverted Earth strategy with its assumptions of duplicability. And perhaps the difficulties about duplication are symptomatic of deeper points about how we conceive of the explanatory role we want mental states to occupy.

B. The preceding considerations about the role of dynamic feedback relations in knitting nervous systems into environments suggest one such reformulation. We tend to take duplication for granted in arguing about Externalism about content. But the difficulties with duplication suggest what might be called Externalism about vehicles. These issues are at least on the surface independent of one another.

Externalism about content has cross-level implications. That is, it concerns the relationships between personal-level contents and subpersonal processes. It allows for the brutely relational determination of content by external factors, not merely for the instrumental determination of content by external factors via their effect on internal physical states and processes. It is this distinction that assumptions of duplicability serve. The Internalist about content does not allow for the brutely relational determination of content; the role of external factors in fixing content is merely instrumental.

By contrast, Externalism about vehicles is not a view about how content is determined at all, so not about whether it can be determined by external factors in a brutely relational way or merely instrumentally. Rather, Vehicle Externalism holds that the subpersonal vehicles of content need not be wholly internal, but can include external processes. That is, it is about the states or processes that carry content, rather than those that fix content. So in principle the two issues cut across one another. An Externalist about content may well assume that vehicles must be internal, even if the contents are determined in a brutely relational way. And someone who allows that vehicles might include external processes might nonetheless insist that their contents cannot be determined other than instrumentally by even more external events, that is, by the brute relations of these even further-flung events to vehicles of content.

C. Now we should be wary of supposing that an indiscriminate appeal to causal spread provides a basis for the possibility of Externalism about vehicles. Of course, causal processes cross back and forth between organism and environment. But not just any old causal spread would suggest the idea of Vehicle Externalism. Vehicles do not escape organisms quite so easily. The Vehicle Externalist needs to make a discriminating appeal to causal spread. To this end, we need to examine the explanatory role occupied by a vehicle of content and its relationship to assumptions of duplicability.

Consider first the suggestion that subpersonal vehicles of content have a token-explanatory relationship to mental states. This suggestion must be understood in the context of the kind of multimodal complexity and feedback that mind and perspective involve. While the generic distinction between token-explanatory and type-explanatory factors can be applied to states of toy systems and devices, these are not the kind of state we are here trying to explain. The suggestion is of course not that any token-explanatory factor is a vehicle of content. It is that in the context of the kind of dynamic complexity that minds involve, we can focus in on the vehicles of content in particular, as distinct from causal spread in general, by appealing to their token-explanatory role. Vehicles causally explain the obtaining of particular mental states, even if the contents of the mental states in question are determined by further type-explanatory factors. Particular mental states can extend through time: for example, particular perceptions or intentions with dynamic spatial content. Vehicles need not explain why the token mental states in question have the contents they have. Nor must the properties of vehicles fix the contents they carry. In addition to normative contributions to fixing content, external factors may contribute to fixing content by explaining why mental states of a certain functional type exist at all. For example, evolution might be appealed to in a teleological account of how a certain type of content is determined. But more than this is needed to explain the obtaining of a particular mental state with content of that type: we might need to know as well, for example, what neurons fired during a certain period. Vehicles are what does the differentially token-explanatory work: what is left when you subtract the explanation of why content of that type exists at all from the full explanation of why one particular mental state with that content obtains. ("Obtains" is here intended to be neutral between occurrent and nonoccurrent mental states.)

Holding the type of vehicle fixed between Earth and Inverted Earth does not necessarily hold the type of content fixed: the issue between Internalists and Externalists about content can be seen as an issue about whether vehicle-type fixes content-type. But even those who deny that vehicle-type fixes content type usually concede that the vehicle-type is fixed by internal physical properties and can in principle be duplicated in an altered environment (though its content may change).

Secondly, the idea of a vehicle of content can be connected to assumptions of duplicability via the token-explanatory characterization of vehicles.

The subpersonal states and processes that do include all token-explanatory factors, with nothing left out, should in principle be duplicable in a different environment. If certain subpersonal states or processes are not duplicable then they do not include everything that is doing token-explanatory work. Something playing a token-explanatory role has been left out; the boundary around the token-explanatory states, or vehicles, should be expanded. Token-explanatory factors are bounded so as to be duplicable in counterfactually altered environments. So, oversimplifying: if a vehicle, then duplicable, and if not duplicable then not the whole vehicle.

Conversely, the explanatory physical states and processes that can in principle be duplicated are those that do the differentially token-explanatory work of vehicles of content. If the relevant explanatory states or processes are duplicable in counterfactually altered environments, then nothing doing token-explanatory work has been left out (though external factors doing type-explanatory or content-fixing work may be, as the Content Externalist holds). Any token-explanatory factor that had been left out could prevent duplication. So, again oversimplifying: if duplicable, then a vehicle, and if not the whole vehicle, then not duplicable.

Finally, the idea of vehicles as token-explanatory and hence as duplicable provides us with a discriminating way to appeal to causal spread in support of the possibility of Vehicle Externalism. We can accept that token-explanatory factors, or vehicles, must in principle be duplicable, without accepting that they must be internal.[11] Perhaps vehicles can go external too. That is, perhaps not just what determines content, including type-explanatory factors, can go external, but also the token-explanatory work of vehicles. Then it might not always be the case that internal physical states can in principle be duplicated in a different environment.

On this view, the boundaries of duplication and of vehicles of content are nonaccidentally related. Issues about the boundaries for duplication can provide a handle on issues about the boundary around vehicles. The threats to assumptions of duplicability from dynamic considerations put outward pressure on the spatial as well as temporal boundaries around vehicles of content. They suggest that token-explanatory work and vehicular status may in some cases extend outward from central physical states and the body into the environment. The token-explanatory vehicle of a particular perception or intention with dynamic spatial content, for example, might include feedback loops that pass through the environment as the agent moves through it. The possibility that other particular mental states might have the same type of content in the absence of such external interaction, say in a dream, does not show that the vehicles must be the same in the two cases, and hence internal. Sameness of content does not guarantee sameness of vehicle; content of a given type can be carried by vehicles of different types. The token-explanatory conception of vehicles respects this point.

The question arises whether these speculations about Vehicle Externalism should be regarded as a reductio of the token-explanatory conception of vehicles, rather than a motivation for Vehicle Externalism. Token-explanatoriness itself does not obviously require any particular boundary that would keep it from leaking out into the world; organisms are causally permeable. Do we have an independently motivated conception of vehicles of content that does require such a boundary for all cases, that requires vehicles to be internal? Or is the usual assumption of Vehicle Internalism a prejudice we can rise above? These questions are raised but not settled here.

D. However, related though more general points can be made about the explanatory role we conceive mental states to occupy, which are not committed to a particular conception of vehicles of content.

The Input-output picture makes it natural to think of the role of mental states in a certain standard way, as mediating between input and output, and hence as internal.[12] But the dynamic considerations that challenge assumptions of duplicability suggest that internal factors may not always be sufficient to mediate input-output relations. This again suggests the possibility of a kind of Externalism about mental states that is not tied to the assumptions of duplicability.

Where duplication is not problematic, the explanatory relationship between external and internal states is slack, and relations between input and output can be mediated by internal states. But the dynamic factors that make duplication problematic also make it problematic to assume that there is an explanatory discontinuity or boundary such that the relations between inputs and outputs at that boundary can be explained solely in terms of states and processes internal to that boundary. We may rather need to explain input-output relations in terms of an entire dynamic singularity: a continuous and complex dynamic system centered on a an active organism, with

feedback loops that may have external as well as internal orbits. On this view, to explain input-output relations, we need to consider not just what function outputs are of inputs, but also what functions inputs are of outputs, where the latter functions may be determined in part by external states (again, consider the higher-order invariants carried by feedback from motion, classically described by Gibson).[13]

The standard view, which has seemed natural to many, ignores dynamic feedback and assumes that we should look at the processes leading from input to output for the explanations of input-output relations. But perhaps we should instead look at whole dynamic singularities to find explanations of input-output relations. Given the way external feedback loops may go through, and bounce off, the external world, such explanatory states may include external factors. Nevertheless, dynamic singularities are centered on and move around with organisms. That which mediates input-output relations can both spread into the environment and move around through environments with the centering organism.

This possibility does not depend on identifying what is to be explained in terms of external states. A different possible reformulation of Externalism would emphasize that external states are needed to explain behavior when that behavior is itself identified in terms of external states, not merely in terms of bodily movement. Peacocke's 'constitutive Externalism' goes in this direction: "...what is distinctively explained by a set of externalist states are relational facts about events or objects, relational facts which go beyond mere bodily movements.....If the person had not seen the person as being in the garden, he would not have pointed towards the garden. It is counterfactuals like that which support the explanatory importance of externalist states" (Peacocke 1992-93, p. 207).

By contrast, the suggestion here is that dynamic feedback reconfigures the explanatory position to be occupied by mental states. This point does not depend on already identifying what is to be explained in world-involving terms. It may even be part of what implicitly motivates such world-involving identifications of behavior. External states may be needed to explain even the particular pattern of events at the body surface, so long as a dynamic view is maintained. We can agree with Peacocke that external states are needed to explain externally individuated actions. But we should disagree with his claim that "... brain states and efferent connections will be enough to explain any given bodily movement" (Peacocke 1992-93, p. 207). Even if we do not individuate what is to be explained Externalistically, dynamic feedback falsifies this claim. Perhaps this claim could be defended by assuming temporal atomism; but that is a high price to pay (see Consciousness in Action, essay 1).

The way dynamic feedback involves external states in explanation raises no special worries about magical mediating mechanisms, by which environments affect the causal powers of mental states without affecting nervous systems (of the kind that have exercised Fodor[14]). These worries depend on assumptions of duplicability, which are just what are undermined by dynamic feedback considerations in relevant cases. The states of dynamic singularities spread across nervous systems and environments, and their causal powers don't need further mediating mechanisms.

10. The reappearing self.

A. The considerations that make duplication problematic in our cases illustrate the truism that the distinction between intrinsic and relational states is relative to a boundary: relational states of persons are after all intrinsic states of something bigger, of systems that include persons: persons-in-environments. Not only are no magical mechanisms implicated, but if anything the suspicion of magic faces the other way. Why assume there is something in principle special about causal relations within some boundary, something that fits internal physical states alone to have the causal powers of mental states? The causal relations between nervous systems and environments are intricate and continuous. There is nothing specially oomphy about causal relations inside the skin, or inside the head, nothing specially capable of pushing and shoving. So there is nothing causally mysterious or inhospitable to materialism or naturalism or realism about relational states of persons. And there is no magical causal boundary around persons. Viewed subpersonally, they are in principle transparent to causality.[15]

If the challenge to the General duplicability assumption succeeds, where would this leave our sense of what and where selves are? It would suggest that we can afford to dispense with the idea of a sharp boundary around subpersonal states and vehicles of content. At the subpersonal level, when we look at dynamic causal flows and patterns, distinct selves show up as distinct singularities, foci of systematic and complex feedback relationships, which cluster in and around and move with the active biological organism but which also loop out into the environment.

I argue that an adequate account of what it is to have a unified perspective may need both normative and subpersonal components (see Consciousness in Action, essays 1, 3, 5). The notion of dynamic singularity may be the subpersonal functional idea of unity needed to complement a normative approach to the unity of consciousness and to explain the perspectival interdependence of perception and action. Nevertheless, at the personal level, what it is to be a subject and agent is open to view, just where it seems to be, in the normatively constrained lives of creatures whose perception/action system embeds them in and deals directly with the world. Within a two-level view of the interdependence of perception and action, the role of norms of coherence and rationality in individuating units of consciousness at the personal level[16] frees us up to recognize the permeability of causal boundaries at the subpersonal level. The role of dynamic causal singularities in complementing normative considerations within a two-level account of unity does not depend on sharp causal boundaries or on Internalism about vehicles.

Appeal to dynamic feedback systems (recurrence, reentry, comparators, etc.) has been a theme of various recent empirical approaches to both consciousness and cognition (see work by Edelman, Gray, Thelen and Smith, Elman et al, etc). Perhaps we can now begin to see why, from a philosophical point of view, this convergence on the importance of dynamic feedback in understanding experience is no accident.

This essay is especially indebted to discussions with and comments from Alan Allport, Ned Block, Andy Clark, Martin Davies, Dan Dennett, Jeffrey Gray, Richard Gregory, Bernard Kobes, Nicholas Rawlins, Rowland Stout, and Michael Wheeler.

[1]. Note that Taylor recognizes a phenomenally distinct possible outcome in which complete behavioral adaptation is achieved, but the world is not seen the right way round through goggles. In such a case "...putting on and removing the spectacles would result in an abrupt transition from one stable perceptual field to another with the opposite left-right

ordering, and without aftereffects." In fact he relies on the qualitative difference between these two cases in making his predictions. See p. 205. So Taylor's view is not behaviorist in a sense that would deny such a difference. See in particular his remarks on consciousness at the end of his final, philosophical chapter.

In some respects his position there must count as philosophically far ahead of its time. It could be regarded even now as fairly radically Externalist, given its combination of emphasis on the "respectability of the once-despised phenomena of consciousness" (p. 363) with the view that "the brain cannot be understood except by considering the behavior of the whole system of which it is a part, that is, an organism living in an environment to which it must learn to respond in various appropriate ways if it is to survive" (p. 356). He expresses misgivings, not about the attempt to understand consciousness in physical terms but rather about such attempts confining their attention to events taking place within the nervous system (p. 361).

[2]. Taylor's account of the experiment differs from Kohler's in claiming that head movement and not eye movement was critical; see Taylor 1962, pp. 255, 259. For present purposes the difference is immaterial; the text follows Kohler's version. For related color adaptation but using various pitches to adapt color perception rather than goggles and eye movement, see Taylor 1962, ch. 10. See also Rock 1966, p. 203n.

[3]. These two cases are representative of a group of cases, discussed further in essay 9.

[4]. Parallel remarks can be made about the content of intentional actions; the parallel is developed in essays 9, 10.

The distinction between instrumental and noninstrumental dependence can also be used to express what is controversial about Externalism about content. Everyone agrees that content can vary with external states that have effects on internal physical states. This is merely instrumental dependence, in that external states are merely a means to changes in internal states. What is controversial is whether content can vary with external states when internal states do not vary, as in Inverted Earth cases--that is, whether content can be noninstrumentally dependent on external states.

[5]. The case, on its face, is naturally described in terms of the dependence of perceptual content on intentions or actions. However, the account of them as violating the Input-output picture describes them in terms of the dependence of perceptual content on motor output. Is this move from dependence on intentions or actions to dependence on output warranted? Doesn't it assume the very picture that is being criticized?

No. In these cases perceptual content varies even though sensory input appears to be constant. If perceptual content varies with the content of intentions or actions, it is then natural to assume that these in turn depend on output, just because input is constant in these cases. None of this entails that basic intentional content must in general be a noninstrumental function of output only, as the Input-output picture would have it. The output-dependence assumption becomes appropriate just when input is constant. This is compatible with holding that in the general case both intentional and perceptual content may depend noninstrumentally on relations between input and output.

For consideration of further objections concerning the distinction between input and output, see essay 9, sect. 6.

[1] The model for noninstrumental dependence can be thought of in terms of an analogy to Twin Earth arguments for externalism, where the content of mental states depends on external/distal states or events noninstrumentally, that is, not just by means of their influence on internal physical states. In such arguments, the content of mental states varies with external states while internal physical states do not vary. But such external

states may still be causally related to the mental states in question; indeed, they are typically assumed to be so. A question externalists typically address is whether content is responsible to external causes or to internal physical states. And externalists hold that external causes can have a deep or constitutive relation to the content of mental states. (See and cf. e.g. Block 1990; Burge 1979, 1986a, 1986b; Davies 1991b, 1992, 1993, etc.)

Now substitute output states or events for external states, and input states or events for internal physical states, and consider perceptual content. The analogous question now becomes: can the content of perceptual states depend on output states noninstrumentally, not just by means of their influence on input states? That is, can perceptual content vary with output states while input states do not vary? If so, such output states may still be causally related to the perceptual states in question. Causal and constitutive relations are no more incompatible here than they are in the case of externalism.

[6]. Though, as Taylor notes, there may well be a correlation between sensory color information and motor signals associated with verbal behavior. Verbal behavior, for all its additional structure and complexity, is still motor behavior. And if color perception can depend noninstrumentally on motor factors, then we may well wonder, without making any concessions to philosophical behaviorism or to verificationism, whether color perception could depend noninstrumentally on linguistic factors. Could we design an experiment that would look for motor-dependent color adaptation involving verbal motor responses rather than nonverbal motor responses? (Cf. Singer et al 1979 on judgmental vs. motor activity in relation to adaptation, p. 81; Varela et al 1991, ch. 8.)

[7]. See for example Goodale and Milner 1992; Jeannerod 1997, pp. 16ff; see also Hurley 1998, essay 5, sect. 4.B for further exposition; and see Thelen and Smith 1995, ch. 10, for an application of the 'what'/'where' distinction within a dynamical systems approach to development and cognition.

[8]. This characterization of the debate includes those who disagree about the 'local supervenience' of phenomenal character or 'qualitative content', as well as those who disagree about representational content. So 'content' is being used very broadly, to indicate what experiences are like as well as what they are about, even if you can't fit what they are like into a 'that...' clause. Radical Externalists apply Content Externalism even to such qualitative contents--or deny that it can come apart from representational content. Nothing turns on using 'content' this broadly; you can substitute 'character' in the phenomenal case. The aim is merely to place issues about Externalism for phenomenal character with issues about Externalism for representational content.

[0]0. Other ways of drawing the distinction between cases in which duplication is problematic and cases where it is not might involve the primary quality/secondary quality distinction, or the distinction between inversions that involve several modalities of information as opposed to inversions that involve only one modality. The systematic interdependency between perception and action in 'where' system cases itself involves a kind of multimodality, if we regard motor information as another modality of information. But it is not clear that a case involving no systematic interdependency of perception and action but more than one sensory modality would create the same difficulties for duplication.

[1]1. Trouble for the Input-output picture stems from relations of input to motor signal or efferent copy feedback, whereas the argument against Duplication also highlights relationships of other sorts of input to proprioceptive, afferent feedback. But complex

dynamic feedback systems typically involve both, as well as afferent feedback loops that go outside the body. The arguments of this essay invoke on all three types of feedback.

[2]2. Or indeed multimodal considerations.

[3]3. The need for output adjustments, and for input-output harmony, in setting up Twin Earth cases, is recognized clearly by Davies, though he underestimates the difficulty of meeting it (e.g. in his circle/ellipse case, which inspired my El Greco cases; cf. sect. 6 below). But to deny Twin Earth arguments can apply to certain types of case without considering output side adjustments or attempting to establish input-output harmony is to fail to recognize this requirement to begin with, and hence to fail to give Twin Earth arguments a fair run; see McGinn 1989, pp. 63ff.

[4]4. For a more technical account, see for example the Introduction to Port and van Gelder 1995 on complex dynamic systems and coupled subsystems in which the parameters of one system are the state variables of the other and vice versa.

Kelso has issued a warning against confusing the causal complexity dynamic systems involve with feedback circuits in the limited and specific sense appealed to by control systems (1995, p. 9; see also Clark 1997, p. 106; and see essay 10 below on control systems and feedback-regulated signals). But feedback is appealed to here is a more general sense: as circular causal flow, including flow from outputs back to inputs (when these can be distinguished and identified). In principle this may or may not take the simple form of a control system. But complex feedback in the form of 'dynamic singularity' centered on an organism is emphasized here. So there is no conflict between the present view and Kelso's comments. Indeed, in essay 10, sect. 6.C, it is argued that the way basic intentions emerge from complex subpersonal feedback relationships cannot be understood in terms of a control system with a preset reference value. Like Kelso, we don't have to answer the question of who sets the reference value. Kelso explains how self-organizing systems avoid regresses and homunculi by demystifying emergence.

[5]5. "...[W]e suppose that environmental differences have the consequence that the same nerve firings and muscle contractions as in the actual situation result in a quite different bodily trajectory"; see Davies 1992, p. 37ff, and his defense of the legitimacy of 'counternomic' suppositions against the 'wildly science fictional' charge.

[6]6. But vision itself may adapt when full freedom of movement is exploited, as in the Papert case.

[7]7. Thanks here to Rowland Stout for detecting an error in an earlier version of this argument.

[1]. Consider loosening the constraints on action further. Instead of just being allowed to point at different objects, suppose Dom is allowed to lie down. He is looking and pointing at the top of the 1 to 1 blue sphere, but the line from his eyes to the red spot is now at quite a shallow angle. If Twin Dom lies down, his eyes are at the same absolute distance from the ground as Dom's eyes are, no longer twice as high. The relevant angle from Twin Dom's eyes to red spot will be shallower that it was when he was standing, but not as shallow as the new angle for Dom. Lying down changes relative height of eyes to red spot in different ways for Dom and Twin Dom. The distortion of Twin Dom's retina cannot itself perform the needed position-sensitive transformation, since in the rigid scenario it does not itself change with position. Similar remarks apply if instead of lying down himself, Dom turns his three spheres around 90 degrees, so that their red spot face left instead of up, and Twin Dom also tips his eggs over onto their sides.

[2]. Just for good measure, consider another auditory issue that arises for duplication in the plastic case. It concerns the switch between two normal mechanisms of auditory

localization, reflectance and relative arrival time of sounds at the two ears. Very roughly: Relative arrival time cannot be used to localize a sound if its source is equidistant from the two ears, but reflectance can be. Reflectance can distinguish between a sound directed in from of your nose and one directly above your head, even though they are equidistant from your ears. Relative arrival time is used when it is available. As someone moves, whether a sound is equidistant from his two ears will change: a sound that is directly in front of someone's face will no longer be equidistant when he turns his face to the left, though other movements will preserve equidistance. So whether or not relative arrival time information is used to localize sound is normally a function of Dom's movements in relation to objects. His normal movements in effect turn auditory localization mechanisms on and off. Twin Dom's normal movements, of moving his head and body and of manipulating objects, will also do this. But will they do it in a way compatible with central duplication, given the plastic vertical distortion of objects and body? Will equidistance relations and hence rates of switching between reflectance and relative arrival time processes differ between Dom and Twin Dom, for given pairs of counterpart movements and objects? Will a given movement for Dom preserve equidistance while its counterpart movement for Twin Dom does not, so that relative arrival time processes kick in for Twin Dom but not for Dom? If so, can a unimodal auditory transformation be applied to Twin Dom's auditory inputs that will cancel out this difference just when it arises, leaving central states unaltered? Or would the needed transformation depend in a complex way on the relationships between, among other things, motor information about head tilt and arm angles, various modalities of information about object and body orientations, and so on? Such selective transformation of information as a function of relationships between multiple other streams of information is itself a central operation, and would itself constitute a central difference between Dom and Twin Dom.

If the antiduplicationist hypothesis holds for cases involving reflectance vs. relative arrival time, it cannot be avoided by the safe zone strategy either. When Dom is functioning normally, his actions naturally do trigger one or the other mechanism of auditory localization. There is nothing artificial or interfering about this effect of his actions. It is part of the normal functioning the duplicationist aims to capture, so it cannot be wholly suppressed. But selective suppression just when central duplication is threatened itself defeats central duplication.

I am indebted to Nicholas Rawlins for raising these issues.

[3]. Here I am indebted to Josefa Toribio.

[4]. Consider Hutchins' comments, in discussing the architecture of cognition:

"The proper unit of analysis is...not bounded by the skin or the skull. It includes the socio-material environment of the person, and the boundaries of the system may shift during the course of activity. Temporal boundaries are important too. ...arbitrary boundaries on the temporal extent of the unit of analysis also risk cutting pathways in ways that leave things inexplicable." (1995, p. 292)

See also and compare: Kelso 1995 on hysteresis and temporal context-dependence in the cortex, e.g. at p. 262; Thelen and Smith 1994, p. 175.

[5]. See essays 4, 9, 10, on how the difference between otherwise similar active and passive movements contributes to eliminating perceptual ambiguities. How could helmets keep track of this difference without invading central processes in one way or another?

[6]. See for example Churchland 1995 on the transformation of multidimensional vectors and the importance of feedback or recurrence.

[7]. Are these empirical issues or not? Neural networks and dynamic systems can involve mathematical structures of such complexity that the only way to discern their dynamic properties may be to let the model run on a computer and see what happens (see Elman et al 1996). This consequence of complexity begins to bring the empirical/a priori distinction under scrutiny. If the relationships of organisms and nervous systems to environments involves this kind of dynamic complexity, applications of the empirical/a priori distinction in philosophy of mind may also be brought under scrutiny. Consider, for example, the role of this distinction in certain arguments about Externalism and first-person authority (e.g. McKinsey 1991).

[8]. Compare Davies 1993, p. 247.

[9]. I am grateful to Martin Davies for prompting me to clarify this point.

[0]0. See and compare various comments on boundaries and causal spread. van Gelder and Port (in Port and van Gelder 1995, p. 34), write that: "...a central element of the dynamical perspective...is that cognitive processes span the nervous system, the body, and the environment; hence cognition cannot be thought of as wholly contained within the nervous system". Cf. Randall Beer in Port and van Gelder 1995, "Computational and Dynamical Languages for Autonomous Agents", p. 130; Thelen and Smith, pp. viii, 17, 91; Hutchins 1995, pp. 292, 316, 364; etc.

11. Assumptions of duplicability, whether general or specific, are strictly assumptions that internal states or processes can be duplicated, for some significant boundary, in a different environment. The supposition that internal plus environmental states can in principle be duplicated is of course granted.

[2]2. Consider Searle 1992, pp. 69, 107.

[3]3. See for example Randall Beer, "Computational and Dynamical Languages for Autonomous Agents", in Port and van Gelder 1995, p. 131.

A distinction needs to be drawn between the role of external states in explanations that derives from dynamic feedback considerations, and the incidental role of external states that just happen to be along the route of causal input or output. States of Mars, for example, need not be especially significant to the reformulation of Externalism if information from the blue sphere on Earth merely happened to be causally routed via Mars before reaching the Earthling's nervous system.

[4]4. Here's a typical expression of this view, which Fodor uses to argue against Externalism.

> ...for one thing to affect the causal powers of another, there must be a mediating law or mechanism. It's a mystery what this could be in the Twin...cases; ....since ...the only mechanisms that can mediate environmental effects on the causal powers of mental states are neurological. (Fodor 1993, p. 41)

Notice that Fodor's point is directed against an appeal to external states in a Twin case, so it depends on assuming duplicability.

[5]5. A purely subpersonal point of view would feature something analogous to the kind of x-ray vision Richard Dawkins describes in The Extended Phenotype:

> We look at life and begin by seeing a collection of interacting individual organisms. We know that they contain smaller units, and we know that they are, in turn, parts of larger composite units, but we fix our gaze on the whole organisms. Then suddenly the image flips. The individual bodies are still

> there;...but they seem to have gone transparent. We see through them to the replicating fragments of DNA within, and we see the wider world as an arena in which these genetic fragments play out their tournaments of manipulative skill. (1982, Pp. 4-5)

But the softening of subpersonal causal boundaries is compatible with the reclaiming of personal boundaries.

[6]6. See essays 3, 4; see also Hurley 1989, ch. 8.