Perception .?. Action

"The mind is a strange machine which can combine the materials offered to it in the most astonishing ways." — B. Russell

How do action and perception link up? According to the standard view—what Susan Hurley calls "the classical sandwich" [10]—information is processed in three stages of our cognitive architecture: an initial perceptual stage, a central stage and a final motor stage. Perceptual and motor processing are essentially peripheral whereas intermediate processing—referred to as cognition—is described as central and holistic.

In this essay, I will first elaborate on the standard view before I turn to consider an alternative skill based approach to action and perception proposed by Alva Noë [13]. According to his enactive approach, perception is in itself an explanatory activity. My project will be to show where potential incompatibilities with the standard view arise. However, taking a closer look at Noë's approach will reveal that his representation avoiding conception of cognitive processing[1] does not do what it promises. It is either inconsistent or eventually collapses into a version of the standard view; for even if perceptual and motor processing are identical, the standard view can largely be preserved. Only our mental architecture then constitutes a loop rather than a sandwich.

1 The Standard View

A cognitive system produces some sort of output (usually behavior) in response to some input (e.g. visual stimulation). The question of how such a system manages to get from input to output arises at Marr's *algorithmic level* [11], i.e. the answer requires moving beyond a general task specification and considering how the system encodes and transforms the information available to it. A common

[1] A word of terminological clarification is advisable. In this essay, I am going to use the terms "cognition" or "cognitive stage" to refer to central processing or the central stage, respectively. Whenever else the word "cognitive" occurs, I shall refer to the entire process, system or architecture, including perceptual and motor stages.

assumption is that the route from perception to action can be described in three stages: perception, cognition and action.

An agent's sensory periphery provides input for perceptual processing—e.g., the retinal stimulation elicited by, say, a ball in front of the agent—whose output feeds into central cognition. Central processing, in turn, provides input to the motor processing stage which finally yields observable output, e.g. a reaching movement.

To understand this three-stage view, it is important to distinguish the peripheral processing mode of perceptual and motor stage from what is found in between. In his "Modularity of Mind" [5], Fodor famously puts forward the distinction between *modular* and non-modular, *central*, processing—on which the standard view is based to a considerable extent—; where the sooner is found in the perceptual and motor stages while the latter is associated with cognition.[2] The most important characteristic of modules is, according to Fodor, *informational encapsulation*. Modules are special-purpose inference making computational systems whose access to background information is limited to a proprietary database; a module's operations only have access to information in the corresponding database while at least some other information available to at least one other cognitive process is not available to this very module. Further, though not essentially, modular cognitive systems are domain specific, innately specified, hardwired, i.e. associated with neuroanatomical mechanisms, operate quickly, autonomous, and bottom-up (cf. ([4], [5]). They operate on lower levels, on what is directly presented to them (in the case of vision for instance luminance contrast, color, orientation, ...) and need not, if at all, draw on large amounts of background information or knowledge available to the agent; such peripheral modules typically operate subconsciously (subpersonally). Sensory input feeds into modules building up a representation of the environment on which higher (personal) level processes subsequently operate. Fodor himself puts it clearly:

> [W]hereas perceptual processes are typically modularized—hence encapsulated, hence stupid in one of the ways reflexes are—the really "smart", really "higher" cognitive processes (thinking for example) are not modular, and, in particular, not encapsulated. ([5], p. 518)

Talking about "smart" cognitive processes, Fodor refers, of course, to central cognition. It is here that—once the agent arrives at a conscious perception of the

[2] It is important to stress that Fodor is *not* a proponent of massive modularity, i.e. he does not claim that the mind entirely consists of modules. "Modular" to him means only "to some interesting extent" ([4], p. 37).

represented environment—background knowledge and reasoning come into play. Processing the output of peripheral modules, the central stage allows the fixation of beliefs, make decisions and form intentions. Unlike peripheral and motor stage, the cognitive stage operates in the realm of propositional attitudes. Fodor characterizes central processing as *Quinean* and *isotrophic*, viz. the propositional attitude system as a whole has certain epistemic properties and a holistic structure in which everything is potentially relevant to everything else. Eventually, the results of central processing serve as input to motor processing modules which are responsible for the execution of intended behavior—e.g. reaching for the ball.

Taken together, this leaves us with a conception of our mental architecture according to which information processing is essentially sequential. A central element is sandwiched between two modular systems or "peripheral buffer zones" ([10], p. 21) connecting up world and agent. The standard view is based on the following assumptions:

(1) Cognitive processing is three-fold involving two modular stages and a central one sandwiched in between.

(2) There is a unidirectional causal flow from perception—via cognition—to action.

(3) Perception provides the starting point for cognition and action in that it computationally builds up the representations manipulated during central processing and eventually fed into the output modules.

(4) Action and perception are clearly dissociated.[3]

2 Problems Ahead

The standard view together with its four basic assumptions has just been sketched. Possible alternatives are, for example, provided by *skill theories* of perception. Related to Heidegger's notion of "Zuhanden-sein" (readiness-to-hand) [9], such theories claim that sensation and perception consist in a set of action capacities the perceiver possesses; the action possibilities provided by the percept constitute perceptual content. As a well known example for skill theories might serve Gibson's idea of direct active perception via *affordances*. Affordances are the possibilities for action a perceived object offers. Agents have to learn them before they can use them.

[3]Though, Fodor's modularity thesis as such does not explicitly require a strict distinction between action and perception. The sandwich view, however, clearly does.

As Myin and O'Regan point out, skill based approaches as compared to traditional theories of perception stand out in at least two respects

> First, by emphasising that perception concerns the activity of an organism in an environment, the traditional focus on the 'inner' as the locus of importance is abandoned. This implies—and this is the second respect—that perception is not, as in many traditional approaches, seen as the establishment of inner representations of the outside world, but rather as active engagement with this outer world [...].
> ([12], p. 32-33)

I will now turn to consider one such skill based approach in detail: the enactive sensorimotor approach. As will turn out, as seems as if this view challenges the assumptions on our cognitive architecture made by the standard view.

According to Alva Noë's enactive approach [13], all perception is skillful explanatory activity—mere stimulation, i.e. sensory input, is insufficient for us to perceive and become aware of what lies in front of us; instead, practical knowledge relating our action to changes in stimulation is required to make sense of the input. This knowledge O'Regan and Noë refer to as *sensorimotor contingencies* (or sensorimotor dependencies) [14].[4]

At the heart of sensorimotor theory lies the claim that sensory input chan-ges as a function of motor actions. To get the flavor, consider a simple straight line. What a straight line *is*, is defined by what happens as we look at it from different positions. If we turn our heads, for instance, the stimulus' orientation changes. As we move our eyes along the line, no changes in sensory stimulation occur. The ability to perceive a straight line depends on our knowledge of how the sensory stimulation changes as a function of manipulating the percept.

Empirical support for the thesis that vision is based on the ability to actively modify incoming sensory stimulations comes from cases of sensory substitution. A relatively well-studied case is *tactile-vision sensory substitution* (TVSS) devices. A blind subject is equipped with a video camera, e.g. mounted on spectacle frames.

[4]Strictly speaking, there are two classes of sensorimotor contingencies (SMC): modality related SMC which are related to sensation (these concern how the senses are affected by given stimuli) and attribute related SMC which are related to perception (they concern the results of categorization of objects and events). While the sooner depend on features of the perceptual apparatus (for vision: how an image is projected on the retina, which kinds of receptive fields the neurons have, the distribution of rods and cones across the retina, etc.), the latter depend on features of objects in the environment (their position in space, color, illumination, shape and texture, etc). In either case, active exploration determines changes in the perceived properties of the percept.

The camera transmits the optical data to a small computer that translates incoming signals into electrotactile stimuli. These are in turn—using a small field of electrodes—delivered to the subject's dorsum of the tongue, fingertip or skin on the back [1]. The subject learns to *see* objects using this device. Even relatively simple devices allow her to identify, e.g., the ball in front of her and estimate her distances to it. Interestingly, TVSS turned out to be of almost no use to subjects, as long as they were unable to actively manipulate the camera, i.e. modify the incoming signal.

The original sensorimotor thesis makes a claim about visual modalities in particular. In "Action in Perception", however, Noë claims that the approach can be extended to all sensory modalities. Though

> at a low level of characterization, the modalities are radically different, implicating distinct forms of sensorimotor patterns [it] is possible to abstract from these differences. ([13], p. 110)

Thus, the sensorimotor dependencies in play might vary from modality to modality—those for audition involve the position of our ears relative to the source of sound, those for vision involve the position of our fovea relative to the percept, etc.—while they can still "represent common spatial qualities" ([13], p. 110). Both, seeing through TVSS and seeing through retinal stimulation allow to identify objects' shapes, colors and positions; seeing as well as hearing indicates that the object of interest lies in a certain distance.

With the enactive approach comes the claim that action, thought and perception cannot be clearly dissociated. We perceive the world by way of actively manipulating it. To make sense of sensory stimulation, we have to appeal to our acquired sensorimotor knowledge, knowledge that Noë describes as non-representational and *conceptual* (in so far as it allows us to make sense of the incoming stimulation). The resulting perceptual experiences present us with possible ways the world could be, something Noë calls *judgeable contents* or *thoughts*. Answering the question of whether or not things are as experienced makes us then come to a certain belief. Thus, perception ultimately consists in actively entertaining these judgeable contents and is therefore "a way of thinking about the world" ([13], p. 189). Note Noë's emphasis that what we *entertain* are not usually representations. Typically, he claims, what we judge and manipulate are *virtual representations*, which really are just *presentations*. A bit of clarification is required. To use an example similar to Noë's own, imagine yourself surfing through the web searching for Christmas presents. What you see on the screen in front of you is a presentation, not a re-presentation of the data you are receiving from the shopping server.

Though Noë is not entirely clear on it, I take the distinction he is trying to make to go somewhere along the lines Benny Shannon proposed [16]. Shannon introduces the notion of *presentation* for those things that are currently within an agent's grasp, those things an agent's *current now* encompasses. As Shannon argues, presentations are to be contrasted with *re-presentations*. Where representations bear characteristics similar to those of presentations but do not concern what lies actually in front of the agent. Re-presentations come into play, for example, when we remember past situations and events.

His core idea seems to resemble Noë's in so far as both propose that we can immediately use things actually present in our current environment instead of having to represent them. Representations are expensive to acquire for an organism and where they are not necessary there is no good reason to appeal to them. Given the environment and our embeddedness in it, we can take advantage of the immediate presence of what is there and need not build it up internally before we can act upon it. Using our bodily skills, we can exploit the direct links between us and the environmental stimuli.

Going back to the shopping example from above, this is to say it is unnecessary to download the complete shop database and re-present the shop internally on your computer. What is currently presented to you is sufficient to retrieve the information you are looking for. Also, the actual page offers various opportunities to actively manipulate the data on the screen (e.g. by scrolling up and down) and thereby explore it further. Still, these presentations have a striking disadvantage: once the actual source is gone, they disappear with it. If the internet connection breaks down there is no way for you to access the shop data anymore.[5] In the shopping case as well as in vision, we never have the entire world (or shop) present at once and are dependent upon it to persist.

However, Noë does not want to abandon representations *completely* from cognitive processing. Rather, he holds that "the role of representation in perceptual theory needs to be reconsidered" ([13], p. 22). What Noë rejects is a unified internal world-model being generated by vision (i.e. exactly what Marr [11] described visual processing to serve for). Noë claims the world to be merely presented and thereby serving as its own representation; or, as Rodney Brooks puts it, "[t]he world is its own best model" ([3], p. 583).

The immediate question coming to mind then is *where* representations come

[5]Of course, nowadays computers have a cache that allows to reconstruct several previously viewed pages even when no internet connection present. The content of a cache might amount to a proper representation but, for the current example, this backup function is not what we are interested in.

into play. Noë himself does not elaborate on this subject. The only claim he clearly and repeatedly makes is that it is not the perceptual system that serves to build up internal representations. But still, already in the very beginning of his book, he admits that

> No doubt perception depends on what takes place in the brain, and very likely there are internal representations in the brain (e.g. content-bearing internal states). What perception is, however, is not a process in the brain but a kind of skillful activity on the part of the animal as a whole. ([13], p. 2).

Interestingly, as the attentive reader has possibly noticed, in an earlier quote Noë claims that different modalities—once we abstract away from their differences—can "represent common spatial qualities" ([13], p. 110). But how can they, if the exercise of sensorimotor skills does *not* give rise to internal representations? I will come back to this question.

Finally, now that Noë's skill based sensorimotor approach has been outlined, let me return to the design of our cognitive architecture. With respect to the assumptions introduced in section 1, the following tensions arise as we compare sensorimotor theory to the classical sandwich view:

(1) Cognitive processing is not clearly separable into three stages. Rather, action, perception and thought are all abundant in the bodily exercise of sensorimotor skills. Our cognitive architecture is made up of a hierarchy of sensorimotor contingencies.

(2) The causal flow through three stages (perception, cognition, action) is replaced by an active interplay of sensorimotor skills and environment; we might describe it as a feedback loop between the agent and her surroundings. Note that Noë describes sensorimotor contingencies as conceptual ([13], p. 183); whereas Fodor's (perceptual) modules are impenetrable to concepts.

(3) Since there can be no perception without action, the starting point of any cognitive activity must be the agent's exercising some motor skill. Once the agent has acquired modality-specific sensorimotor knowledge, she can make sense of the incoming stimulation and consciously perceive. Internal representations are not required for this process. What is being manipulated is not something internally stored or built up but rather the environment itself. Further, it is not computational processes defined over representations that constitute cognition. Practical knowledge—relating the execution

of explanatory activity to changes in stimulation—applied to the perceiver's environment shall take its place.

(4) Perception *is* in itself action. There is no proper way to dissociate the two capacities.

It thus seems as if the sensorimotor approach is radically incompatible with the standard view and calls—since its claim that perception is in fact skilled bodily exercise of a certain capacity is empirically quite plausible[6]—for a complete revision of what our cognitive architecture look like. In the next section I will argue that this is not the case. For, despite its prima facie revolutionary appearance, Noë's approach either turns out to be inconsistent or largely collapses into the standard view.

3 Harmless, Though

Historically, perception and action have been studied largely separately. However, nowadays few would deny that both are closely related, if not interdependent or even incorporated into one another. The suggestion that action an perception go hand in hand—that they might even be meditated through identical neurophysiological mechanisms—is supported, apart from the merely behavioral evidence presented earlier, by the discovery of *mirror neurons* [15]. Mirror neurons are neurons found in motor areas that are associated with—i.e. they show characteristic activation patterns during—perception, typically seeing perform conspecifics perform an action, as well as the execution of that very action. It apperars therefore that O'Regan and Noë's sensorimotor contingency theory pushes into the right direction. But is their approach really as drastically incompatible with the sandwich view as outlined above?

Before I consider assumptions (1) to (4) again, let me first return to the issue on representations left open in section 2. O'Regan and Noë's claim is that sensorimotor contingencies do not serve to built up an internal representation. But still, they do not abandon representations entirely from cognitive processing. Further, Noë describes different modalities to be representing, at an abstract level, common spatial qualities. But how could they, if an object is perceived by exercising sensorimotor skills that do *not* built up internal representations? It seems that (at a crucial point) the sensorimotor approach is inconsistent. On the one hand,

[6] I have presented the example of TVSS earlier. Further support comes from change blindness, perceptual completion (filling in the blind spot) and inversion glasses.

it does not allow representations to be built up during visual perception. On the other, however, it uses them.

In response to O'Regan and Noë's sensorimotor theory, Vittorio Gallese and Christian Keysers take up the issue of mirror neurons and quite rightly argue that

> the merit of focusing on the relevance of sensorimotor contingencies for perception [...] is not to falsify the importance of representations for perception but to help us understand the nature of these representations. ([6], p. 984)

And further:

> Ironically, in a way mirror neurons instantiate both the very expertise of sensorimotor contingencies so central to [O'Regan and Noë's] theory, and the representation of the world, the importance of which [O'Regan and Noë] argue against. ([6], p. 984)

On their (Gallese and Keysers') view, representations in the brain might be motor consequences. Sensorimotor dependencies could then offer a novel *way of understanding how representations work*. But their establishment does by no means justify the denial of internal representations of the world being built up by some sort of explanatory activity. According to Gallese and Keysers, action control and representations are intrinsically related; and the medium through which both are mediated is the mirror neuron system.

I do not know whether the mirror neuron system is the be-all and end-all in this discussion. But it makes it at least conceivable that there are ways to combine the plausible gist of sensorimotor contingency theory—that perception is essentially a skillful explanatory activity—with a representation based approach to perception and cognitive processing. And, furthermore, mirror neuron activity could serve to show that Noë's *presentations*, are in fact some sort of re-presentations since already the presentation of an action is sufficient to elicit the corresponding motor representation in the observing agent. The claim that presentations might turn out to be representations gains even more force when considering neural activity elicited in response to visual stimulation as compared to visual imagery. In MT (a brain area in the visual cortex associated with motion detection) for example, activation patterns during vision and visual imagery are strikingly similar [8]. This suggests that similar mechanisms are involved in coding *present* and *re-presented* stimuli. Thus, Noë's argument from presentations being cheaper for the organism plainly fails.[7]

[7]However, this is not to deny that sensorimotor activity is a way of manipulating the

All this on the table, what do we make of our inconsistencies with the standard approach? I shall consider (1) to (4) in reverse order.

(4) The last assumption of the standard view (a clear dissociation of perception and action) is the only one that ultimately needs to be rejected fit sensorimotor contingencies into the picture. Noë's view identifies perception with skilled bodily exercise of sensorimotor contingencies; perceiving thus is a way of acting. Motor output is the means by which new input to the system is generated, perceptual processing *is* motor processing. This being the case, action and perception cannot be dissociated. The identification of perception and action leads the classical sandwich to turn into a loop connecting peripheral processing and a central stage in both directions.

(3) Once we realize that the route form perception and action to cognition and back runs in a loop, it is easy to see how motor activity can be the starting point for perception and cognition. A simple (maybe accidental) first movement suffices to give rise to the very first discoveries of sensorimotor dependencies. Once the loop has been entered, there is no need for any unidirectional causal flow originating from a certain input pattern; on the contrary, there is constant feedback between perceptual motor processes in the periphery and cognition (central processing by the agent).

With this, two major questions arise. First, if sensorimotor contingencies constitute peripheral processing, and if our cognitive architecture consists in a hierarchy of sensorimotor contingencies (as Noë suggests), what remains for central cognition? And second, can the exercise of sensorimotor skills indeed be described as a peripheral, i.e. modular process? The first question will be answered further below; to approach the latter, reviewing Fodor's criteria for modularity is advisable.

Initially, and most importantly, modules are informationally encapsulated. I take this requirement to be fulfilled by sensorimotor contingencies, since they are described as applied merely to the currently present environment. Though described as conceptual,[8] sensorimotor dependencies can only assess limited information (the knowledge associated with the respective modality). Domain specificity comes, at least at a low level of description, with modality specificity. As is obvious from the presentation above, quick bottom-up, subpersonal operations are characteristic to sensorimotor contingencies, too. Like Fodorean modules, sensorimotor contingencies operate on lower levels, on what is directly presented

environment.

[8] I will return to the question of conceptuality in my considerations about requirement (2).

to them.[9] The only background information they draw on is the acquired seonsorimotor knowledge. So far, sensorimotor dependencies might thus be counted as modular. However, they clearly violate the criterion of being innate since they have to be acquired. Possibly, we could regard the *capacity* to acquire sensorimotor dependencies as innate and circumvent the problem in saying that sensorimotor dependencies are generally innate, but the agent still has to learn how to use them practically. This then, could also allow them to be hardwired, and hence fulfill another criterion. Still, the idea that some ability has to be acquired to be capable of modular processing is certainly not what Fodor had in mind. But since innateness (and being hardwired as well) is not too crucial a criterion for modularity, and since—once acquired—there need not be any *functional* difference between innate and learned mechanisms, let us regard this difference as a minor one.[10]

A little more problematic is the fact that Fodorean modules are described as inference making *computational* systems building up representations of the environment for higher level processes to operate on as opposed to the environment itself being manipulated using practical sensorimotor knowledge (i.e. the tension described in the second part of (3)). I shall return to this problem after considering the remaining assumptions. For the moment just note that—for the main part—sensorimotor contingencies could, roughly, be described as Fodorean modules.

(2) With regard to the second assumption, the above considerations imply that—since our cognitive architecture now has a circular rather than a linear structure— there being a feedback loop between agent and environment is perfectly unproblematic. And it is conceivable to assume that sensorimotor dependencies are the means by which input and output are mediated between the agent and her environment. The second part of (2)—that Fodorean modules are not penetrable by concepts whereas sensorimotor contingencies are described as *being* conceptual—is not as easily to handle. Being conceptual, on Noë's account, is something like *can be made sense of*. Something is conceptual in so far as we *understand* and *judge* it. But this cries out to be found at the personal, rather than the subpersonal, unconscious level; whereas Noë describes sensorimotor dependencies to be predominantly subpersonal. In my opinion, the distinction between

[9] Input is provided directly by the agent's environment in both cases. Still, there is a difference in so far as the input is actively manipulated on the encative view whereas the standard view rather assumes some kind of passive delivery to sensors. However, the active exploration and intake of the environment does not conflict with modularity. It even comes automatically when assuming perception is a way of acting.

[10] Anyway, the worst case would be having to offer a characterization of modules slightly different from the Fodorean version to make sensorimotor contingencies modular.

personal and subpersonal level is not—as Noë states—irrelevant ([13], p.30). If sensorimotor contingencies operate subconsciously, I cannot see why it should be important that the agent can make sense of or judge the incoming stimulation, i.e. why sensorimotor dependencies need to be conceptual. If they are conceptual, on the other hand, i.e. they give rise to understanding, I am inclined the think that whatever is going on should be found at the personal level. The only way out of this dilemma is to distinguish personal from subpersonal level sensorimotor contingencies.

We are thus really dealing with two different kinds of sensorimotor dependencies,[11] one of which operates subpersonally, the other within the agent's conscious grasp. The former type does not need to be conceptual and could thereby fulfill Fodor's requirement of non-conceptuality on modules.[12] The latter type, on the contrary, might come into play elsewhere in our cognitive architecture. This binary divide between subpersonal and personal level provides us with an answer the question of what—given that subpersonal level sensorimotor contingencies constitute peripheral processing—constitutes central cognition: conceptual, personal level sensorimotor dependencies (henceforth "central" sensorimotor contingencies).[13] But can they fulfill the requirements for central processing? In order to do so, they must be *Quinean*, *isotrophic* and operate in the realm of propositional attitudes. Since Noë describes that the way we come to a belief is by judging the perceptual content provided (by some perceptual activity) to us, I take the third requirement to be unproblematic. That the propositional attitude system as a whole has certain epistemic properties and that it builds up a holistic structure in which everything is relevant to everything else, does not stand in contradiction to Noë's description either. (In fact, he is barely concerned with beliefs and desires.) Finally, according to the standard view, the results of central processing have to somehow feed back into motor processing modules, i.e. they must somehow be made available to perceptual motor activity. This could simply be accounted for by certain lower level sensorimotor contingencies connecting up with central ones; similar to the way central processing is linked to peripheral modules on the standard view.

[11] The distinction made here is entirely different from that between modality related SMC and attribute related SCM I mentioned in section 2.

[12] Admittedly, Fodor's criterion is stricter than just being non-conceptual. He requires it to be impossible for modules to be penetrated by concepts. However, I think we can understand low-level sensorimotor contingencies to meet even this demand.

[13] Assuming central sensorimotor contingencies are connected to multiple lower level ones makes it easy to see how they can overcome informational encapsulation and domain specificity.

(1) All this, of course, has consequences for the remaining assumption. Cognitive processing can now be seen as being two stage, not three stage. Noë seems to go even further in equating action, perception and thought. But, as my discussion above indicates, even if our minds are shaped by a complicated hierarchy of practical skills" ([13], p.31), it seems that there have to be at least two different kinds of them. Therefore, I suggest that the two stages—perceptual and motor processing on the one hand and cognition on the other—are realized by two sorts of sensorimotor contingencies: cognition by central ones (which might be positioned further up in the hierarchy) and peripheral processing by low-level ones.

Finally, let me return to the second part of requirement (3). The tension left to consider was that sensorimotor contingencies are described as practical knowledge applied to the environment whereas the standard view postulates computational modules building up inner representations of the external world that are subsequently manipulated. Going back to our considerations about representations in the brain and the mirror neuron systems at the beginning of this section, it is possible to interpret the exercise sensorimotor contingencies as building up some sort of representations: motor consequences. At the same time, this does not mean that the environment is not being manipulated. In fact, the manipulation could be represented with the motor consequences. Once the representations are established, there is no trouble in seeing how computational inference rules can be defined over them. Practical manipulation of the environment would then, ultimately, go along with rule based computations applied to representations.

In conclusion, the seemingly radical incompatibility with the standard view ends up requiring no more than a slight—empirically plausibly—modification; namely that our conception of a sandwich-like cognitive architecture should be replaced by a loop.

4 Upshot

I started out with the question of how action and perception link up with one another. The "classical sandwich" suggests information being processed linearly: perceptual modules process input; their output feeds into central (non-modular) processing; the results provided by this second stage eventually serve as input to modular motor processing yielding behavioral output. In section 2 I introduced the skill based sensorimotor theory of perception put forward by O'Regan and, especially, Noë. According to this approach, perception is in itself an explanatory activity. Taking a closer look at sensorimotor contingencies, some inconsistencies

and conceptual problems have been pointed out. My project in the last section has been to show that the radical challenges posed upon the standard view by the enactive approach do not hold what they promise. In fact—under an appropriate interpretation—the empirically plausible sensorimotor approach broadly collapses into a slightly modified version of the representation based standard view; where this modification consists in assuming that our mental architecture does take the form of a loop rather than a sandwich.

References

[1] Bach-y-Rita, P. and Kercel S.W. (2003). Sensory substitution and the human-machine interface. *Trends in Cognitive Science*, 7 (12), 541-547

[2] Bermúdez, J.L. (2005). *Philosophy of Psychology: A Contemporary Introduction*. New York and London: Routledge.

[3] Brooks, R.A. (1991). Intelligence without reason. *Proceedings of the 12th International Joint Conference on Artificial Intelligence*. Sydney, Australia. 569-595.

[4] Fodor, J.A. (1983). *The Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, MA: MIT Press.

[5] Fodor, J.A. (2006). Précis of The Modularity of Mind. In: J.L. Bermúdez, ed. *Philosophy of Psychology: Contemporary Readings*. New York and London: Routledge.

[6] Gallese, V. and Keysers, C. (2001). Mirror neurons: a sensorimotor representation system. *Behavioral and Brain Sciecnces*, 24, 983-984.

[7] Gibson, J.J. (1986), *The Ecological Approach to Visual Perception*. Hillsdale, NJ: Lawrence Erlbaum.

[8] Goebel, R. et al. (1998). The constructive nature of vision: direct evidence from functional magnetic resonance imaging studies of apparent motion and motion imagery. *European Journal of Neuroscience*, 10, 1563-1573.

[9] Heidegger, M. (1957). *Sein und Zeit*. Tübingen: Niemeyer.

[10] Hurley, S.L. (1998). *Consciousness in Action*. Cambridge, MA: Harvard University Press.

[11] Marr, D. (2206). Excerpts from Vision. In: J.L. Bermúdez, ed. *Philosophy of Psychology: Contemporary Readings*. New York and London: Routledge.

[12] Myin, E. and O'Regan, J.K. (2002). Perceptual consciousness, access to modality and skill theories. *Journal of Consciousness Studies*, 9 (1), 27-46.

[13] Noë, A. (2004). *Action in Perception*. Cambridge, MA: MIT Press.

[14] O'Regan, J.K. and Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciecnces*, 24, 939-973.

[15] Rizzolatti, G. et al. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, 3, 131-141.

[16] Shannon, B. (1993). *The Representational and the Presentational*. London: Harvester Wheatsheaf. Part II.