Mog – Copenhagen talk: Three Worries about C&T's paper

This was supposed to be "two responses to the Cosmelli and Thompson paper" but it's now more "three worries that I have about it". I should state at the beginning that I really like the spirit of the paper, but that I think that in order for it to convince orthodox philosophers of mind and cognitive science there are a few clarifications that need to be made.

1) Causal Constitution Problems

In the orthodox account of consciousness there is a problem as to what is constitutive of consciousness and what is merely causally required. Those things that are merely causally required are considered to be background or enabling conditions. On this view, generally the body, and indeed the bulk of the brain are seen in this way – they are causally necessary in that without them the neural correlates of consciousness would not be able to be present. Their function is just that: background and enabling conditions of consciousness. The brain in a vat thought experiment sits nicely with this account of consciousness as it is easy to see that if we substitute these background and enabling conditions – as long as we do it so well as to allow all the inputs and outputs to carry on just as before – we would still have just the same consciousness as in a non envatted brain; the core realiser of consciousness, the part that is responsible for distinct subjective feels, is somewhere in the brain. In short there is a neural correlate of consciousness, and this neural correlate is said to be constitutive of consciousness.

C&T point out that when we look at the nervous system as a non-linear dynamical system rather than through the glasses of the traditional "input/output" models, it becomes difficult to see how there can be a core realiser of consciousness in the brain. Consciousness on this account is a property of a whole system, and the nervous system is just a subsystem of this whole system – the living organism. As they note, the causal/constitution distinction can't be made on this type of account.

"In dense nonlinear systems where all state variables interact with each other, any change in an individual variable becomes inseparable from the state of the entire system. In such cases, **the distinction between regular causes (regularities in the system's behavior) and singular causes (unique nonrepeatable events that change the system's behavior) becomes meaningless (Wagner 1999), and there is arguably no core realizer for a given property or behavior less than the system itself**." (p.8)

Quick Aside (There is a worry here that the claim that dynamical systems accounts do not lend themselves to distinctions between causal and constitutive elements is debatable as some dynamical systems accounts involve notions of "controllers" which appear to fulfill the same function/are mappable onto as constitutive elements but I am going to leave this to the side as I'm not well versed enough in dynamical systems to present a strong case either way on this.

Despite this (re. quote) and the insistence in the paper that they are not interested in the metaphysical questions of causality and constitution, throughout the paper they continue to talk in causal/constituent terms, for example:

"**If certain brain processes simply could not be realized in the absence of the body, and these brain processes included those crucial for consciousness, then we would have reason to believe that the body is not merely causally enabling for consciousness, but also constitutive.** The argument of this paper is that the brain-in-vat thought experiment, when spelled out with the requisite detail, suggests precisely this result." (p. 9, my emphasis).

On first reading this may look as if they are leaving themselves open to objections about the "causal/constitution error error"; that the only way of distinguishing causal processes from constitutive ones is by already assuming an internal/external distinction (whether we regard external as external to the brain or nervous system or body..). Such distinctions, as Hurley 2006 tells us, therefore tell us nothing about the nature of consciousness. Instead, we should seek to give explanatory accounts of consciousness rather than constitutive or causal ones (though a good explanatory account may give us information about constitution) (Hurley, 2006).

Because the expressions of C&T's account like the one above are in direct contradiction to what they state elsewhere in the paper, I suggest that the spirit of the paper does not in fact warrant any accusations of causal/constitution error error and that we should read these passages as suggesting that,

if we were to view the problem through the orthodox lens which grants that a constitutive claim about consciousness can be made then, by the orthodox arguments, and given the physiological evidence, **if it would make sense to claim that a part of the brain is constitutive of consciousness then we ought to conclude that the body is also constitutive (rather than merely causal) of consciousness**. If this is how the paper should be read, we must ask what this really means… Does it make sense to make such a claim – doe we retain any useful meaning of constitution in such a case?

Although we are looking at the systems as dynamical systems it is difficult to avoid this terminology and still say something meaningful about consciousness. We can give an explanatory account: to explain consciousness we must appeal to the brain/body system – there is no more basic system that we can explain consciousness in… - so rather than calling the brain/body system the minimal constituent base for consciousness perhaps C&T should call it the "minimal explanatory base"?

It still seems to me as if some important information is not being expressed by calling a system the minimal explanatory base for consciousness, that would be captured by calling something the minimal constituent base, but I am unsure as to whether this is because although my heart has made the leap to the non-linear dynamic approach the rest of me lags with my history in the orthodox tradition…

Summary:
I am not clear as to what kind of account of consciousness they are seeking to give. They explicitly reject the orthodox account of consciousness and the metaphysical

notions of causality and constitution which are enmeshed in this account in favour a non-linear dynamical systems account in which causality and constitution have no real meaning. However, they continue to talk in terms of causality and constitution throughout the paper. This could be for three possible reasons:

a) They are trying to show that even within the orthodox framework the philosopher's brain in a vat makes no sense.
b) The orthodox terminology is so entrenched in us that it is difficult to get an explanation across in purely dynamical terms.
c) Purely dynamical explanations seem to miss something vital out which we can grasp nearer using orthodox terminology.

2) The Sensorimotor Vat

C&T make a good case that, ignoring the metaphysical question of whether a particular structure instantiates consciousness at a particular moment in time, there is good evidence to suggest that the brain (or the parts of the brain that are considered to be NCC's) are not merely a taker of inputs and a giver of outputs but rather are a part of a non-linear dynamical system. If this is the case it makes no sense to isolate one part of such a tightly coupled system and focus on that assuming that the rest of the system is just a background condition/enabling condition for that part as the orthodox folk do. The philosophers' brain in vat, they conclude, will not do. If we want to talk of brains in vats kept up and running and conscious then we must include explanations of the systems that they are part of. Once we do this, they claim, we are effectively no longer talking of a brain in a vat, but rather a biologically autonomous, sensorimotor agent: "a body in a world".

I follow C&T about half way. I can see that to keep the brain alive and running, what is required is to rebuild a certain amount of the system (at least functionally) of which the brain is a part. So whatever technology we use we are completing the loops required for organismic regulation – the end result being to all intents and purposes a "body".

This (organismic regulatory system) however, according to C&T is only a subsystem of the biologically autonomous sensorimotor agent that they claim is what we must now substitute for the philosophers' brain in a vat. This is where I think the difficulties are encountered. C&T claim that consciousness is a system-level property of that level of system – the autonomous sensorimotor system. But the evidence given for this is developmental evidence. That being an autonomous sensorimotor agent is required for consciousness to develop is a weaker claim than the one I am assuming that C&T want to make, and an explanatory account of consciousness does not require a developmental account.

There could be two different interpretations of what the brain in a vat thought experiment is being used for here: To see what is required for a biv to have exactly the same conscious experience as me – this gets back to the philosophical thought experiment – we can start asking questions like will it know that it is a biv etc… I don't want to dispute that if we were to create a biv that had experiences just like me

it would need (in effect, i.e. functionally or in some type of simulated way) to be an autonomous, sensorimotor system. However I don't believe that this is what the experiment is being used for here. If it were, orthodox defenders of the philosophers biv would object that the original thought experiment allows for all this – you are supposed to set up the vat with any amazing technology to enable the brain to not notice that it is no longer not in a body in a world. They are not tied to the technology being serial. They would point out that even though there is no longer a "real body" or a "real world" exactly the same experiential life can be had by the biv, thus still pumping the intuition that, as long as everything else is functionally adequate it is the brain itself that instantiates consciousness and thus we need not lose the causal/constitutive distinction.

Alternatively, the experiment could be used to see what the minimal constitutive base for conscious experience is. Despite C&T clearly stating that they are not concerned with the metaphysical questions concerning bivs I get the feeling from the paper this is actually what they are pushing for. For them the minimal constitutive base for consciousness is nothing less than the autonomous system. An autonomous system however does not need to be a sensorimotor system, and the biv that C&T want to leave us with is fully sensorimotor as well as autonomous.

I find the evidence that sensorimotor regulation is developmentally required for consciousness convincing, and I would suggest that it could be argued that even non-developmentally if our biv is to have exactly the same experiences as us it will require to put things in memory and so forth for which reinforcement is necessary, for which sensorimotor regulation (whether real or simulated) over time is very likely necessary. However, although it is an interesting claim that for extended human experience to occur in a system it must be sensorimotor as well as autonomous, this is a claim about more than the minimal constitutive base for experience.

The non-sensorimotor but still organismically regulating system would seem to be the minimal constituent base for conscious experience. As sensorimotor regulation may be required for memory and so forth these may just be flashes of experience, but experiences none the less. Given the evidence in this paper there doesn't appear to be anything that could be gotten from a sensorimotor autonomous system, rather than a merely autonomous system that would justify the attribution of minimal conscious experience to the former but not the latter. They could tell some sort of Noe type story about requiring sensorimotor capacities to be able to make sense of sensations and thus to have experiences, but I don't know if they want to go down that road….

Summary:

Autonomous systems do not need to be sensorimotor. The evidence that sensorimotor regulation is linked to consciousness is convincing developmentally, but not constitutively. Although C&T state explicitly that they are not interested in the constitutive question, (1) as my first worry explained it is not clear that this is the case all the way through, and (2) the developmental evidence that they consider seems to lead them to the conclusion that rather than a "brain in a vat" being the null hypothesis of the constitutive basis for consciousness, it should be a "body in a world", i.e. an autonomous, sensorimotor agent. However it seems to me that, if we ignore the developmental stories about the importance of sensorimotor regulation, the

evidence just shows that the minimal constitutive base for consciousness is a brain in a sub-body, i.e. an autonomous but non-sensorimotor system.

3) The importance of Autonomy

If the body ought to be seen as part of the constitutive base for consciousness because it is so tightly coupled to the brain that it constitutes a system, why ought we not to see oxygen as part of the constituent base too, after all oxygen is just as tightly coupled.

Thompson's response to this worry is to state that **dense coupling does not necessarily entail autonomy**, and that nothing less than an autonomous system is required as the constitutive base for consciousness (personal communication).

If dense coupling does not entail autonomy then what does?

An autonomous system is: Self-organising, Self-producing, Self-controlling

And more specifically:

> "In an autonomous system, the constituent processes (i) recursively depend on each other for their generation and their realization as a network, (ii) constitute the system as a unity in whatever domain they exist, and (iii) determine a domain of possible interactions with the environment." (Varela, 1979, p. 55). (p.65 Mind in life)

The brain/body/oxygen system would therefore not constitute an autonomous system because processes within that system do not fulfill these criteria. However:

> "What counts as the system in any given case, and hence whether it is autonomous or heteronomous, is context- dependent and interest-relative." (p. 72, Mind in Life)

However, an ecosystem can be seen to be an autonomous system, yet it is not clear to me how it fulfills the criteria: it may be self-organising and self-producing but it does not seem to be self-controlling in any normal sense of the term, nor does it determine a domain of possible interactions with the environment as far as I can see…

If viewing a system as autonomous is just an explanatory tool, and the system could just as well be seen as heteronomous or as a sub-system of another system, then it is unclear to me that the constituent processes really do constitute the system as a unity and thus I am unclear as to how the appeal to autonomy saves Cosmelli and Thompson from the threat of the explanatory spread of consciousness out of the living system and into the environment.