

The Coupling-Constitution Fallacy

Ken Aizawa
Department of Philosophy
Centenary College of Louisiana
Shreveport, LA 71134

Fred Adams
Department of Philosophy
University of Delaware
Newark, DE 19716

In the face of both common sense and contemporary science, an increasing number of philosophers and psychologists have found themselves attracted to the view that, in ordinary tool use, we have instances in which cognitive processes span the cranial boundary and extend into extracranial space. As Daniel Dennett puts it, “*minds are composed of tools for thinking* that we not only obtain from the wider (social) world, but largely leave in the world, instead of cluttering up our brains with them” (Dennett, 2000, p. 21, italics in original). According to this kind of environmentalized cognition, when a student takes notes in class, the student literally commits information to memory. When someone uses pencil and paper to compute large sums, cognitive processes extend into the pencil and paper themselves. Fans of environmentalized cognition differ among themselves about just what types of tool use bring about the extension of the cognitive into the extracranial world. So, where some may wish to maintain that all brain-tool couplings will let the cognitive enter the artifactual, others restrict the view to only certain species of brain-tool couplings.¹ Be these refinements as they may, all fans of environmentalized cognition have enthusiastically embraced what they recognize to be a radical departure from

¹ Haugeland, (1998), for example, is fairly explicit about the kinds of couplings they have in mind.

orthodoxy.

In this talk today, we don't propose a critique of environmentalized cognition *per se*. Nor do we here propose a defense of common sense or scientific orthodoxy. Rather, we wish to draw attention to what seems to us to be the most common fallacy that advocates of environmentalized cognition commit. We call this the "coupling-constitution" fallacy. What is common to these arguments is a move from the observation that process X is in some way causally connected to a cognitive process Y to the conclusion that X is part of the cognitive process Y. We can see this coupling-constitution fallacy in other contexts. Consider the bi-metallic strip in an ordinary thermostat. The expansion and contraction of this strip is coupled to the ambient temperature of a room and the air conditioning apparatus for that room. Nevertheless, this gives us no reason to say that the expansion and contraction of the strip extends beyond the limits of the strip and into the room or air conditioner. The Watt governor provides another example. The combustion of fuel in the governed engine is tightly coupled to the rotation of the weighted arms, yet the process of combustion does not extend beyond the bounds of the engine.² This is the generic form of a coupling argument, but we find a range of specific variation in the literature. What we wish to do here is to draw attention to the numerous versions of this fallacy.

1. The Simple Fallacy. In what we call the simple coupling argument, all that is invoked is a causal connection or looping between the cognizing organism and its environment. Raymond Gibbs, (2001), provides a brief illustration of the simple coupling argument by appeal to what is involved in windsurfing. Gibbs is concerned with intentions, rather than cognition, but

² For other examples, see XXX.

the fallacy remains the same:

The windsurfer continually affects and is affected by the set of the rig, so the behavioral intention to successfully windsurf emerges as a result of the interaction between the person and environment. Focusing on the agent alone, or on how the agent responds to the environment, fails to capture the complex nuances of windsurfing behavior. Just as it is important to understand the significance of paper and pencil when one does long division, where the cognition of doing long division is in part “offloaded” into the environment, the intentionality in windsurfing is best understood as a distributed cognitive behavior involving a person, a device, and the environment (Gibbs, 2001, pp. 117-118).

Gibbs’s discussion here is ambiguous. In the first place, there is an epistemic interpretation of what Gibbs is after. This is the idea that we cannot (should not) study intentions (cognition) without keeping an eye on the interaction between the organism and the environment. For the present, we are not concerned with this thesis. Suffice it to say, however, that we think that the environmentalist’s charge that cognitive science does not attend to environmental interactions is overblown. The second interpretation of the foregoing passage finds an ontological issue in it. It concerns the bounds of cognition, or how the processes involved in windsurfing might be divided into the cognitive and the non-cognitive. Gibbs at least comes close to the ontological issue when he claims that the intentionality in windsurfing is best understood as a distributed cognitive behavior involving a person, a device, and the environment. Unfortunately, he gives no reason to think this is so. In describing the windsurfer case, Gibbs simply assumes that, in virtue of a causal coupling, the windsurfer and her environment should be analyzed as a single cognitive/intentional whole. So, this seems to us to be a pretty clear illustration of the coupling-constitution fallacy.

2. *The “Credit Assignment” version of the fallacy.* Andy Clark, (2001), runs a similar kind of argument, although in this case he introduces a tangential question about what is

responsible for a particular product. The example involves writing an academic paper.³

Confronted, at last, with the shiny finished product the good materialist may find herself congratulating her brain on its good work. But this is misleading. It is misleading not simply because (as usual) most of the ideas were not our own anyway, but because the structure, form and flow of the final product often depends heavily on the complex ways the brain cooperates with, and leans upon, various special features of the media and technologies with which it continually interacts. We tend to think of our biological brains as the point source of the whole final content. But if we look a little more closely what we may find often is that the biological brain participated in some potent and iterated loops through the cognitive technological environment. We began, perhaps, by looking over some old notes, then turned to some original sources. As we read, our brain generated a few fragmentary, on-the-spot responses which were duly stored as marks on the page, or in the margins. This cycle repeats, pausing to loop back to the original plans and sketches, amending them in the same fragmentary, on-the-spot fashion. This whole process of critiquing, rearranging, streamlining and linking is deeply informed by quite specific properties of the external media, which allow the sequence of simple reactions to become organized and grow (hopefully) into something like an argument. The brain's role is crucial and special. But it is not the whole story. In fact, the true (fast and frugal!) power and beauty of the brain's role is that it acts as a mediating factor in a variety of complex and iterated processes which continually loop between brain, body and technological environment. And it is this larger system which solves the problem. ... The intelligent process just *is* the spatially and temporally extended one which zig-zags between brain, body, and world (Clark, 2001, p. 132).

Here we can agree with everything up until that last sentence. Here we find a familiar pattern in environmentalized cognition, a long description of the causal connections between the brain and environment followed by the move to the view that these causal loops constitute part of the brain's cognitive processing. Yet, from the fact that cognitive brain processes are *coupled* to environmental processes we cannot simply infer that the environmental processes *constitute* part of the cognitive process. To think otherwise is to commit the coupling-constitution fallacy. Now, as an aside, anyone can admit that the brain and its tools are jointly responsible for the production of the journal article and that the brain and its tools jointly constitute a single causal

³ Actually, the example first appears in Clark, (1997), but its use to support environmentalized cognition is less marked there.

process.⁴ Still, joint responsibility for a product or outcome does not suffice to show that the brain and its tools constitute a single *cognitive* process. After all, it is the interaction between the spinning bowling ball and the surface of the alley that between them lead to all the pins falling. Still, there is no “extended bowling ball” meshing with the alley.

3. *The “System Version” of the Fallacy.* Clark & Chalmers, (1998), run a version of the coupling argument inserting the idea that humans and their tools form a cognitive system.⁵ They write,

In these cases [of tool use], the human organism is linked with an external entity in a two-way interaction, creating a *coupled system* that can be seen as a cognitive system in its own right. All the components in the system play an active causal role, and they jointly govern behavior in the same sort of way that cognition usually does. If we remove the external components the system’s behavioral competence will drop, just as it would if we removed part of its brain. Our thesis is that this sort of coupled process counts equally well as a cognitive process, whether or not it is wholly in the head (Clark & Chalmers, 1988, pp. 8-9).

Clark, (2003), develops this same idea by claiming that humans are hybrid artifact-organism systems, or cyborgs.⁶

We can grant for the sake of argument that the combination of a human being with pencil

⁴ This jointly responsible idea figures more prominently in the version presented in Clark, (1997). Haugeland, (1998), runs the same “collaborative” line about navigating to San Jose. By driving the interstate, one relies on the structure of the interstate and on one’s cognitive abilities in dealing with roads. Thus, the road and the brain collaborate in navigating to San Jose and they constitute a single causal process. Still, that does not make the road and brain a single *cognitive* process. Establishing the latter stronger claim is what the environmentalized cognition is after.

⁵ In fact, one can pick up a reference to a system in the passage from Clark, (2001), cited above. Much of Haugeland, (1998), can be viewed as a more elaborate case of the system version of the coupling-constitution fallacy.

⁶ Doesn’t the cyborg example play into the intracranialist’s hand? After all, cyborgs are hybrids of organism and artifact, rather than simply organisms. So, shouldn’t humans with tools be hybrids of cognizers and artifacts, rather than simply cognizers?

and paper constitutes a system, that a person with a laptop computer constitutes a system, that a person with a notebook constitutes a system, and so forth. We can also concede that humans and their tools constitute cognitive systems. Yet, this does nothing to establish the idea that cognition is environmentalized. It does not follow from the fact that one has an “X system” that every component of the system does X. Obviously there are systems that consist of many types of components and involve a multiplicity of processes. An air conditioning system, for example, can involve a thermostat, a compressor, an evaporation coil, a fan, and so forth. Perhaps we can say that the process of cooling of the air as it passes over the evaporation coils is the process of “conditioning the air,” but surely the liquefaction of Freon and the electrical processes within the thermostat and the opening and closing of the circuit in the thermostat are not air conditioning. Surely nothing forces us to lump all of these processes under a single descriptor, “air conditioning.” Another example is a personal computer, a computing system. Suppose, for the sake of argument, that we don’t limit the notion of computing to what the CPU does. Suppose that we understand computing broadly so as to cover any sort of information processing. Thus, we might count the process of reading a floppy disk, reading a compact disk, and turning the computer on as kinds of information processing, hence as kinds of computing. Even on this very broad understanding of computing, it is still not the case that every process in this computing system is a computing process. There is the production of heat by the CPU, the circulation of air caused by the fan, the transmission of electrons in the computer’s cathode ray tube, and the discharge of the computer’s internal battery. Think of a sound system. Not every component produces sounds. The speakers do, but amplifiers, lasers in CD players, volume controls, and tone controls do not. Again, not every component of an X system does X.

4. *Gibbs's Interpersonal Coupling Fallacy*. Gibbs (2001) claims that "intentions are, in many cases, emergent products of interactions between individuals, and between individuals and the environment, and that therefore they exist in a distributed manner across individuals" (Gibbs, 2001, p. 106). Clearly, Gibbs is an environmentalist about at least some intentions and, as we have seen, is prone to commit the simple coupling-constitution fallacy. In addition, however, he advances some more complicated versions of the fallacy.

One of Gibbs's arguments is based on a dialogue he observed in a bar. The dialogue begins after John spills a beer.

John: I wonder if there is a towel behind the bar.

Nicole (goes over to the bar and grabs a towel): Here you go.

John: Oh thanks! I wasn't actually asking you to get a towel for me. I just was thinking aloud about whether there might be a towel that I could get from the bartender. But thanks.

Gibbs begins his analysis of this dialogue by saying that, "John intends his utterance with a particular meaning, but changes his mind and accepts Nicole's interpretation of what he said."

We think that Gibbs's treatment of this case is flawed in many ways, so it will take a while to work through these problems before we can ultimately relate it to the other coupling arguments.

So, first off, we think that Gibbs simply misunderstands John's comment. John is not changing his mind about anything. He is not adopting Nicole's interpretation of what he said; in fact, he is explicitly rejecting it. John says, "I wasn't actually asking you to get a towel for me," which is an explicit rejection of what he thinks Nicole thinks he intends. When he says, "But thanks," he means that, even though he didn't intend for Nicole to get him a towel, he is thankful that she did it anyway. It looks as though John's initial intention remains constant throughout the whole episode.

Not to rest our argument too much on what Gibbs might take to be our idiosyncratic understanding of the foregoing dialogue, we might try to develop an imaginary scenario in which John does change his initial intention. How would the scenario have to be different in order for John to have really changed his mind about his original intention? Let's say that at t_0 he had the intention merely to wonder out loud and so he proceeds to utter, "I wonder if there is a towel behind the bar." Nicole then goes and gets the towel and says "Here you go." Now at t_1 , let John say "Thanks. I'm glad you discerned what I intended." Now at least Nicole's actions have provoked a kind of conflict between the intention John had at t_0 and the intention he implies (at t_1) he had at t_0 . This, however, still does not seem to be an instance of actions at t_1 changing John's intentions at t_0 . Indeed, the mechanics of this exchange are that of a comic scene with Inspector Clouseau. Clouseau clearly intends one thing, has something unexpected arise, but then tries to play off the surprise as what he intended all along. What reason is there to think that John changed the intention he had at t_0 , rather than that he changed *his interpretation of the intention he had at t_0* ? It could be that John suffers from a failure of memory or self-deception. It must surely be admitted that self-deception or failures of memory can lead to distorted interpretations or assessments of the intentions one had in the past. So, why not in these types of cases? Gibbs provides no reason to prefer the view that John changed his intentions at t_0 to the view that John merely changed his assessment of his intentions at t_0 . Worse, Gibbs appears to be completely insensitive to this distinction. Nowhere is this more evident than when he writes,

The fact that John altered *what he believed to be* his original intention shows that Nicole's interpretation of his intention actually shaped John's *own conception* of what that intention may be (Gibbs, 2001, p. 110, emphasis added).

What Gibbs says here can be conceded by the orthodox view of the locus of cognition. What

Gibbs is hoping for, but has provided no argument for, is much stronger, namely, that John's intentions at t_0 were changed.

But, suppose we set aside the infelicity of Gibbs's original example wherein John says, "I wasn't actually asking you to get a towel for me." Further suppose that at t_1 John really is able to alter the intention he had at t_0 . Further, let us suppose that there are no problems with backwards causation, that there is nothing wrong with events at t_1 causally influencing events at t_0 . (We think we're being especially generous here.) Still, Gibbs must come to grips with the fundamental flaw in coupling arguments, namely, the fact that events at one time causally influence cognitive events at another time does not make it the case that those first events constitute part of a single cognitive process that includes the cognitive events. More concretely, the fact that Nicole's and John's actions made some cognitive difference to John's intention at t_0 is not enough to establish that Nicole's and John's actions are part of the same cognitive process or state as John's intention at t_0 .

Further evidence that Gibbs is guilty of confusing constitution relations and causal relations in the analysis of this case is supported by his claims following another sample dialogue. He notes that "speakers intentions also clearly shift as a result of conversation and may at times not be viewed as solely a product of an individual speakers's mind" (Gibbs, 2001, p. 111). It is surely common ground that intentions shift or change over the course of a conversation. I ask you to pass the salt. That, against a backdrop of other factors, might cause you to form the intention to pass the salt. And, of course, in such a case, there is a perfectly good sense in which your intention is not solely a product of your mind, namely, your intention is not caused exclusively by events within your own mind. Yet, such an admission does nothing to

challenge the standard view of the locus of cognition. For all that has been conceded, we can still maintain that your intention to pass the salt is entirely *constituted* by events and processes within your cranium. So, even under quite generous concessions, Gibbs has not produced an argument for environmentalized cognition or environmentalized intentions.

5. *Conclusion.* In this paper, we have drawn attention to the sundry forms of coupling-constitution fallacy. We think that one cannot simply move from the fact that an environmental process is coupled to a cognitive process to the conclusion that it constitutes part of that cognitive process. It is, therefore, evident that, at the very least, other conditions must be in place in order to license a move from coupled to constituted. Although we don't have time to defend the view here, we are on record as maintaining that this will involve the operation of distinct types of covering laws. Simply stated, cognitive processes are covered by cognitive laws, such as those one finds in cognitive psychology textbooks. Such laws for memory include those implicated in primacy and recency effects, the superior memory of common nouns for names, etc.⁷ Such laws for psychophysics include Weber's law and Fechner's law. We speculate that once one begins to develop appropriate principles for moving from coupling to constitution, one will thereby be committed to the view that cognition is typically a purely intracranial process.

References

- Clark, A. (2003). *Natural-Born Cyborgs*. Oxford: Oxford University Press.
- Clark, A. (2001). Reasons, Robots, and the Extended Mind. *Mind & Language*, 16, 121-145.
- Clark, A., & Chalmers, D. (1998), The extended mind, *Analysis*, 58, 7-19.

⁷ One might view part of Robert Rupert's, (2004), discussion as one way of fleshing out this idea.

Dennett, D. (2000). Making tools for thinking. In Sperber, D., (Ed.) *Metarepresentations: A multidisciplinary perspective*. Oxford: Oxford University Press. (pp. 17-29).

Gibbs, R. J. (2001). Intentions as emergent products of social interactions. In Malle, Bertram, F., Moses, Louis, J., and Baldwin, D. A., (Eds.). *Intentions and Intentionality*. Cambridge, MA: MIT Press. (pp. 105-122).

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

Rupert, R. (2004). Challenges to the hypothesis of extended cognition. *Journal of Philosophy*, 101, 389-428.