©2010 The Institute of Mind and Behavior, Inc.
The Journal of Mind and Behavior
Winter 2010, Volume 31, Number 1
Pages 37–48
ISSN 0271–0137

The Boundaries Still Stand: A Reply to Fisher

Kenneth Aizawa

Centenary College of Louisiana

In his recent critical notice of *The Bounds of Cognition* in this journal, Justin Fisher advances a set of concerns that favor the hypothesis that, under certain circumstances, cognitive processes span the brain, body, and world. One is that it is too much to require that representations in cognitive process must have non-derived content. A second is that it is possible that extended objects bear non-derived content. A third is that extended cognition might advocate the extension of certain general categories of cognition. A fourth is that *Bounds* misapplies Andy Clark and David Chalmers' so-called "parity principle." The purpose of this rejoinder is to show how Fisher's concerns can be, or have already been, addressed.

Keywords: extended cognition, bounds of cognition, mark of the cognitive

One of the great virtues of Fisher's critical notice of Frederick Adams and Kenneth Aizawa's *The Bounds of Cognition* is that it identifies some prominent lines of response to the book. I am grateful to the *Journal of Mind and Behavior* for giving me the opportunity to reply to Fisher's notice, since this I believe this will advance the debate over the existence and evidence for extended cognition.[1]

A familiar claim in the extended cognition literature is that much of the history of psychology has been marked by prejudice. This is the prejudice — a remnant of Descartes' enduring legacy — that cognitive processes occur only in the brain. Cognitive psychologists simply assume that the mind is realized by the brain. We find one or another version of this charge in Clark and Chalmers (1998), Haugeland (1998), Rowlands (1999, 2003), and elsewhere. Rather than supposing that cognitive processes occur only within the brain, the advocates

Requests for reprints should be sent to Kenneth Aizawa, Charles T. Beaird Professor of Philosophy, Department of Philosophy, Centenary College of Louisiana, Shreveport, Louisiana 71104. Email: ken.aizawa@gmail.com

[1]The following discussion is based upon replies given at an "Authors meet Critics" session with Frederick Adams, Justin Fisher, Lawrence Shapiro, and myself, at the 2009 meetings of the Southern Society for Philosophy and Psychology in Savannah, Georgia.

of extended cognition propose that there are good grounds for thinking that cognitive processes span the brain, body, and environment. The extended cognition movement should, therefore, be seen as a liberating revolution.

Much of *The Bounds of Cognition*, however, is an attempt to rebut this charge of prejudice. It is an attempt to vindicate what is, on all sides, taken to be the long-standing intracranialist practice in psychology. Our method of doing this begins by articulating a rather weak form of cognitivism according to which there are non-derived mental representations that are manipulated in a relatively limited number of ways. The next step is to point out that, if cognitive processes are as cognitivists claim and if such processes are found only within the brain, then we have some non-question begging reason that vindicates what has been claimed to be mere prejudice.

Viewed in this light, some of Fisher's anxieties about *The Bounds of Cognition* should dissipate. For one thing, it should not be hard to see why so little of the book is dedicated to "playing offense," namely, giving a positive account of cognitivism.[2] It is because cognitivism is a relatively familiar view. Why spend time on the familiar? A general rule of thumb is that a book should treat of what is new. In our case, what is new would be "playing defense," illuminating and challenging the arguments that are given in support of the new hypothesis of extended cognition. For a second thing, one can see why we do not worry too much about the issue degenerating into a terminological dispute. If the entire extended cognition movement dissolves into a mere terminological debate, then cognitive psychology could proceed undisturbed by worries about extended cognition. That is the cheap response, but it is basically correct. We are concerned with the legitimacy of the current enterprise of scientific cognitive psychology. It does not really matter to us whether the processes studied in this field are called "cognitive" or not. What does matter is that there will continue to be a legitimate scientific enterprise that studies processes that are typically found in the brain and not typically found spanning the brain, body, and environment, and that these processes will have many of the features formerly studied by cognitive psychologists. If past practices are vindicated, it does not much matter if those practices are labeled "cognitive psychology" or not. Indeed, we could adopt one of Fisher's suggestions and say that "cognitive-processing-conservatively-construed" does not extend.[3] Then, good old-fashioned intracranialist cognitive psychology could go on as it largely has for the past few decades.

Fisher's comments are not, of course, exhausted by anxieties. He has specific objections. He gives reason to doubt that non-derived representations are necessary for cognition. He is shocked that we have not argued that external objects

[2]Cf., Fisher (2008), p. 349, for comments on the relative amounts of effort dedicated to "playing offense" versus "playing defense."

[3]Cf., Fisher (2008), p. 356.

cannot bear non-derived content. He objects to what he takes to be our parochial understanding of cognition. And, finally, he objects to our (mis)application of the so-called "parity principle." So, let me address each of these arguments in turn.

The Non-Derived Content Condition Is Implausible

In Chapter 3 of *The Bounds of Cognition*, we argued that a necessary condition for cognitive processing is that it involves non-derived representations, representations that mean what they do independently of other representational or intentional capacities. Fisher, however, challenges this by providing four putative counterexamples.[4] These are four types of cognitive representations that appear to derive their content from other representations. He proposes that:

(A) The content of thoughts and experiences derive from the content of perceptual states.

(B) The content of some perceptual states are derived, in part, from the content of other perceptual states.

(C) The contents of long-term memory derive from the contents of perceptual states and short-term memory states.

(D) The contents of imaginational states derive from the contents of associated perceptual states.

These are some excellent examples and they provide what I find to be the most compelling published version of an attack on the invocation of non-derived content [others include Dennett (1990), Clark (2005), and Menary (2006)]. A fully satisfying rejoinder to these examples might well require an entire paper in its own right, but one can begin to see one's way out of most of these putative counterexamples by, first, recalling that what we have referred to as the derived/non-derived distinction is also sometimes referred to as the conventional/non-conventional distinction. That is, the meanings of derived representations, such as those of written words, a white flag, and a red flashing light, are sometimes said to be created by conventions. Correlatively, non-derived representations have their content other than by conventions, i.e., non-conventionally. So, the short and simple reply to Fisher's examples is to say that we should not be confused by the word "derived." The word "derived" has several meanings, only one of which is the relatively narrow one intended in this context of non-derived representations.[5] We can avoid the misleading meanings of "derived" if we note that, in the examples that Fisher cites, it is not as though one portion of the brain sets up conventions in virtue of which other portions of the brain get content. It is not, for example, the case that the perceptual systems get together

[4]Cf., Fisher (2008), p. 348.

[5]This point is made in a reply to Dennett (1990), in Chapter 3 of *The Bounds of Cognition*.

and in any sense establish a convention that certain states of other parts of the brain will have such and such a meaning. That is the cheap reply.

One can get a sense of one of the misleading senses of "derivation," one not discussed in *The Bounds of Cognition*, by considering the following example. Suppose one wants to create a Turing machine that computes the successor function, $S(x) = x + 1$. To do this, one might introduce a convention according to which the number n is represented by a string of n "1"s on the Turing machine tape. Here we have assigned content to strings of "1"s. These strings now have what we mean by derived content. Next, suppose we give this Turing machine the following program

$$\begin{array}{l} S_0\,1\,R\,S_0 \\ S_0\,0\,1\,S_1 \end{array}$$

and we start it with its read-write head scanning the leftmost "1" of the string "1111." What this Turing machine will do is scroll right across the set of "1"s and add one more. The program has performed a computation and might be said to have derived a representation that means five. This is, perhaps, a legitimate use of the term "derived" and "derived representation," but it does not accord with the use in play with the derived/non-derived, conventional/non-conventional distinction. The convention according to which a number n is represented by a string of n "1"s is a derivation of content from the minds of those who established the convention, but the execution of the Turing machine program to produce the string "11111" is not a derivation of content in the intended sense.

This example seems to provide a promising reply to Fisher's examples (A), (C), and (D). In these examples, it looks like certain representations are set up, then there are cognitive derivations that simply re-deployment antecedently meaningful representations. So, for example, in understanding cases such as (A), it could be that certain mental representations acquire their content in a non-derived way by means of some causal theory of mental content, but then later certain mental computations in thought deploy these antecedently meaningful expressions.[6] Similar stories might be told *mutatis mutandis* for the contents of long term memory and imagination. By contrast, case (B) seems to require more work. Fisher maintains that "many perceptual states are intrinsically relational, and represent various aspects of one's environment in a way that displays their relations to other represented aspects" (Fisher, 2008, p. 348). So, take the concept, say, of X being on Y. Perhaps to acquire the concept of X being on Y one has to invoke mental representations of things in the X slot and mental represen-

[6]See Stich and Warfield (1994), Rupert (2008), or Adams and Aizawa (2010) for some proposals regarding how such causal theories of mental content might work.

tations of things in the Y slot. Thus, one needs to have the contents of the things in the slots to have the content of the concept of something being on something else. So, we have a representation of the relation of being on that does not mean what it does independently of other representational or intentional capacities. This is an important case and more work would need to be done to reply to it, but it is perhaps premature to abandon the idea of non-derived representations. Perhaps mere additional articulation of the theory of non-derived representations is needed.

Is it Possible that External Objects have Non-Derived Content?

Fisher is shocked that we "don't even consider the question of whether external devices like an Alzheimer's notebook might satisfy the various accounts they gesture towards" (p. 349). In the first place, it seems that Fisher simply missed a spot. In Chapter 3, we wrote, "In advancing the hypothesis that cognition involves non-derived representations, we do not presuppose that machines or other objects outside the head *cannot* bear non-derived content" (p. 39). We, then, go on to invoke this point in a reply to Dennett's critique of non-derived content. In the second place, one might recall that our view is that it is possible for cognition to extend. It is only a matter of contingent empirical fact that cognition typically does not. So, we are committed to the view that it is possible that external devices might satisfy some conditions on non-derived meaning. Thus, we do not think that this condition by itself blocks absolutely all possible instances of extended cognition. It, at most, rules out some cases. Whether or not particular external objects bear non-derived content must be assessed on a case by case basis. In fact, we reiterate these two points in our discussion of Andy Clark's (2005), critique of non-derived content. We wrote that,

> We do not propose that non-biological structures are incapable of supporting non-derived content. Our view is that, as a matter of contingent empirical fact, cognitive processing typically occurs within brains, even though it is possible for it to extend. To keep this view consistent, we have to maintain that it is possible for non-derived content to occur outside the brain and body. (Adams and Aizawa, 2008, pp. 47–48)

In the third place, it seems to us reasonable to say that, were one's theory of non-derived content to yield the conclusion that ordinary notebook inscriptions have non-derived content, then that would be very close to a *reductio* of the theory. Ordinary notebook inscriptions are a paradigm of derived content bearers. In the fourth place, as Fisher notes, it is indeed correct that Cummins' picture theory of representation leads to many, many things bearing content in virtue of isomorphisms. This, however, is one of the problems of Cummins' theory. It promiscuously assigns content all over the place. Moreover, isomorphism is

a symmetric relation, where representation is an asymmetric relation. These problems are discussed in Aizawa's book on *The Systematicity Arguments*, to which Adams and I referred in our footnote 16 on p. 37. In the fifth, and final, place, we do not intend to endorse any of the specific theories of naturalized content we mention in the book. In fact, we have published criticisms of both Fodor's asymmetric dependency theory and Cummins' picture theory. Our point in mentioning them is not to endorse them, but merely to draw the reader's attention to this relatively familiar project in cognitive science, namely, the naturalized semantics project. We do not endorse an account of our own, since such an endorsement is not essential to our current needs.

The short of it, therefore, is that we *do* consider the question of whether external devices like an Alzheimer's notebook might bear non-derived content. We also consider whether some of theories of naturalized content, such as those of Fodor, Cummins, and others, might lead to the conclusion that a notebook might have symbols with non-derived content. The view is that notebooks can, in principle, bear non-derived content, but in typical cases they do not. To date, such a view has not been challenged by the advocates of extended cognition. But, the debate is young.

What about a Broader Notion of Cognitive Processing?

Fisher's response to our proposal regarding cognitive mechanisms is relatively familiar. (Other instances of this strategy may be found in Clark [2008], and Sprevak [2009].) It sets up an opposition between Adams and Aizawa's narrowly construed theory of cognition and an extended cognition's broadly construed theory of cognition. From here, Fisher develops two lines of criticism. First, we should prefer the general theory of cognition. Second, once we see these two positions, we can see once again that the whole debate between extended cognition and its opponents threatens to degenerate into a mere terminological dispute over the word "cognitive." The matter of a terminological dispute was treated above, so now we can focus on this hypothesis of extended general cognition.

We discussed this possibility in section 4.1 of our book, among other places. In fact, we discussed it in our original paper, "The Bounds of Cognition."[7] The gist of our reply was that, while ordinary language often preserves broad, generic categories for things that are studied in science, the advance of scientific understanding often leads to a diminished significance to this usage. One example we used was heat. There are, of course, many things that are hot and heat is generated in many different ways, but there is no single discipline of "heat science." Instead, heat is treated in different ways by different sciences. The heat created by friction falls to physical theory, the heat created by exother-

[7]See Adams and Aizawa (2001).

mic chemical reactions falls to chemistry, and the heat of microbial action in decomposition falls to biology. Instead of using heat as an example of a general theory, Fisher proposes flight.[8] Yet, like some general category of heat, some general functional characterization of flight, such as any scheme for locomotion through the atmosphere, is of limited scientific interest. How much use would this definition be to someone designing a helicopter, a space rocket, or a fighter jet? How would this help the ornithologist studying bird flight or the entomologist studying insect flight? As the study of flight advances, deeper principles and mechanisms in specific contexts become more important, where the superficial concept of locomotion through the atmosphere is diminished. Memory is another case in point. Surely there is a generic, ordinary language conception of memory, but in the scientific study of memory, one from time to time finds something like the view that there is no such thing as memory. Instead, there is only long-term memory, short-term memory, procedural memory, declarative memory, or other specific types of memory, as there might be. The rough idea is that simple generic notions suitable for everyday use are of diminished importance with the advance of science. We have proposed that insofar as the hypothesis of extended cognition ties its fortunes to ordinary, common sense notions, its enduring relevance to science is diminished.

The Misapplication of the "Parity Principle"

Fisher's final objection, that we have misapplied the "parity principle" to the famous Inga-Otto thought experiment, comes as a surprise.[9] It seems that if there is a misapplication of the "parity principle," the blame for that lies with Clark and Chalmers. But, that does not really matter. Even Fisher's preferred way of applying the principle merely forces us to reorganize our presentation of material. So, let us work through this. Recall that Clark and Chalmers' principle maintains that,

> If as we confront some task, a part of the world functions as a process which, were it *done in the head*, we would have no hesitation in recognizing as part of the cognitive process, then that part of the world *is* (so we claim) part of the cognitive process. (1998, p. 8)

Commenting on this, Fisher adds that

---

[8]In truth, Fisher seems to send mixed messages about the wisdom of pursuing general functional characterizations, such as that flight is any scheme for locomotion through the atmosphere: "It's doubtful that there is a single correct functional characterization for flight" (p. 351f). Fisher's reasons for rejecting general functional characterizations are, perhaps, not entirely different than those given in Adams and Aizawa (2001, 2008).

[9]Cf., Fisher (2008), p. 353–354.

> To apply the Parity Principle, [A and A] would need to consider some creature who employs an intracranial process that is otherwise just like Otto's use of his notebook, and they would need to ask whether that intracranial process would count as cognitive. (pp. 353–354)

Agreed. We need a creature that meets two conditions: (a) It has an intracranial process that is just like Otto's use of his notebook and (b) That intracranial process counts as cognitive. Next, Fisher claims that, "[A and A] choose a surprisingly poor candidate for this role: the ordinary human Inga" (p. 354). But wait a second! We did not choose to compare Otto and Inga; Clark and Chalmers did. And, it is plausible to suppose that, Fisher's assumptions notwithstanding, Clark and Chalmers take Inga to be a creature that meets conditions (a) and (b). Here is why. *Ex hypothesi*, Inga is a normal human being, so that it is pretty plausible that there are intracranial processes in Inga that would count as cognitive. So, Inga meets (b). Moreover, Clark and Chalmers write a number of things (some of which are cited in *The Bounds of Cognition*) indicating that *they*, at least, take Inga and Otto to be invoking processes that are in relevant respects the same:

> For in relevant respects the cases are entirely analogous: the notebook plays for Otto the same role that memory plays for Inga. The information in the notebook functions just like the information constituting an ordinary non-occurrent belief; it just happens that this information lies beyond the skin. (Clark and Chalmers, 1998, p. 13)

Then, later,

> In both cases the information is reliably there when needed, available to consciousness and available to guide action, in just the way we expect belief to be.
>
> Certainly, insofar as beliefs and desires are characterized by their explanatory roles, Otto's and Inga's cases seem to be on a par: the essential causal dynamics of the two cases mirror each other precisely. We are happy to explain Inga's action in terms of her occurrent desire to go to the museum and her standing belief that the museum is on 53rd street, and we should be happy to explain Otto's actions in the same way. (ibid.)

Then still later,

> To provide substantial resistance, an opponent has to show that Otto's and Inga's cases differ in some important and relevant respect. But in what respect are the cases different? (ibid.)

So, it seems reasonable to us, that Clark and Chalmers think that Inga meets condition (a) as well as condition (b). But, Adams and Aizawa think, along with Fisher, apparently, that Inga and Otto are really quite different. This is what we are driving at in pages 137–140 of *Bounds* by reference to such things as primacy effects, recency effects, and the generation effect. But, rather than

simply siding with us, Fisher chides us for making an inappropriate comparison between Inga and Otto. Instead, we should have considered a comparison between Otto and a creature we will call "Marty the Martian."[10] So, how does this work?

Marty is like Otto in the sense that Marty "remembers" things by using a tiny pencil inside his head to inscribe the information he wants to remember. Now, Fisher thinks, one has the real test case. Now, Fisher thinks, we have a creature that meets conditions (a) and (b). Marty has an intracranial process that is just like Otto's use of his notebook and this intracranial process counts as cognitive.

Ok. But, wait another second. We can grant that condition (9a) is satisfied, but what is the case for saying that the pencil writing is cognitive in Marty's case?

> My own inclination is to say that the imagined aliens' memory *would* count as a cognitive process. To be sure, the aliens' memory would work quite differently from ordinary human memory, and would have strengths and limitations quite different from those of ordinary human memory. But these seem to be irrelevant differences of fine grained implementation. What's being implemented strikes me as a clearly cognitive capacity — the capacity to remember things. (Fisher, 2008, p. 19)

We talked about a Martian thought experiment like this on pages 48–49 of *The Bounds of Cognition*. Given such examples, we can either treat this as a counterexample to our theory of cognition or reject Marty as a cognitive agent. Given this choice, we ~~think~~ can rationally stand by our theory. What examples like Marty show, we think, is that it may well be possible to accomplish many tasks, such as getting to the Museum of Modern Art, without really thinking about it. This is where we would, again, invoke the differences between Marty and normal humans. Thus, humans have primacy and recency effects in memory, the generation effect, the laws of learning and forgetting, and so forth. These are the psychology textbooks effects that we have described before.

Fisher anticipates this last response and dismisses the facts to which we allude as "irrelevant differences of fine-grained implementation." (This is a common reply to the differences we point out between Inga and Otto.) Yet, it is unclear to us what basis he has for this judgment. The facts to which we allude are taken from cognitive psychology and since our overarching project is to vindicate the familiar intracranialist cognitive psychology, why would this be irrelevant? If we want the psychologically relevant facts, why should we not look to psychology itself? So, the upshot here seems to be that the Inga/Otto thought experiment and the Marty/Otto thought experiment raise exactly the same issues in only slightly different ways. One way or another, one has to

[10]Cf., Fisher ~~(2008)~~, p. 354. (2008, p 354).

come to grips with the idea that cognitive psychology recognizes certain differences between how memory typically works and how notebooks typically work.[11]

To conclude, one should concede that Fisher's comments point out a number of respects in which *The Bounds of Cognition* might have developed its views more clearly and perspicuously. Nevertheless, the flaws that Fisher thinks he finds are generally not serious. To a significant degree, the book presented enough exposition and argumentation to both illuminate and challenge the hypothesis of extended cognition.

au: correct year? 1994

au: city correct? (see next page) Malden

References

Adams, F., and Aizawa, K. (1995). Fodorian semantics. In T. Warfield and S. Stich (Eds.), *Mental representation: A reader* (pp. 223–242). Cambridge, Massachusetts: Blackwell.

Adams, F., and Aizawa, K. (2001). The bounds of cognition. *Philosophical Psychology, 14*, 43–64.

Adams, F., and Aizawa, K. (2008). *The Bounds of Cognition*. Boston: Blackwell Publishers.

Adams, F., and Aizawa, K. (2010). Causal theories of mental content. In E.N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 2010 Edition). Available at: http://plato.stanford.edu/entries/content-causal/

Aizawa, K. (2003). *The systematicity arguments*. Boston: Kluwer Academic Publishers.

~~Aizawa, K. and Adams, F. (2005). Defending non-derived content. *Philosophical Psychology, 18*, 661–669.~~

Clark, A. (2005). Intrinsic content, active memory, and the extended mind. *Analysis, 65*, 1–11.

Clark, A. (2008). *Supersizing the mind: Embodiment, action and cognitive extension*. Oxford: Oxford University Press.

Clark, A., and Chalmers, D. (1998). The extended mind. *Analysis, 58*, 7–19.

Dennett, D. (1990). The myth of original intentionality. In K.A. Mohyeldin Said, W.H. Newton-Smith, R. Viale, and K.V. Wilkes (Eds.), *Modeling the mind* (pp. 43–62). Oxford: Oxford University Press.

Fisher, J. (2008). Critical notice of *The Bounds of Cognition*. *Journal of Mind and Behavior, 29*, 345–358.

Haugeland, J. (1998). Mind embodied and embedded. In J. Haugeland (Ed.), *Having thought* (pp. 207–237). Cambridge, Massachusetts: Harvard University Press.

Menary, R. (2006). Attacking the bounds of cognition. *Philosophical Psychology, 19*, 329–344.

Rowlands, M. (1999). *The body in mind*. Cambridge: Cambridge University Press.

au: is this cited in text?

au: are these references in text?

au: is this correct as we have it (not editors' initials go before last names)

(2008, p. 355)

[11] Fisher ~~(2008), p. 355,~~ also charges us with misreading the three versions of Clark and Chalmers' Tetris case. After quoting Clark and Chalmers at length in *Bounds*, we wrote "We take the point here to be that there is no principled difference among these cases, hence that there is a cognitive equivalence between the three cases, so that we have an instance of extended cognition in case (2)." That still looks to be a pretty plausible interpretation of what Clark and Chalmers meant when they wrote,

> How much *cognition* is present in these cases? We suggest that all three cases are similar. Case (3) with the neural implant seems clearly to be on a par with case (1). And case (2) with the rotation button displays the same sort of computational structure as case (3), distributed across agent and computer instead of internalized within the agent. If the rotation in case (3) is cognitive, by what right do we count case (2) as fundamentally different? (~~Clark and Chalmers~~ (1998, p. 7)

TR: this should be further reduced.

Of course, Adams and Aizawa, appear to agree with Fisher that the cases are not the same. But, that does not show that Clark and Chalmers think that the cases are not the same. Clark and Chalmers evidently think the three cases are, if not the same, then similar and "on a par," so that such differences as one finds are not cognitively relevant.

au: in this cited in text?

Rowlands, M. (2003). *Externalism: Putting mind and world back together again*. Montreal: McGill–Queen's University Press.
Rupert, R. (2008). Causal theories of mental content. *Philosophy Compass*, 3, 353–380.
Sprevak, M. (2009). Extended cognition and functionalism. *The Journal of Philosophy*, 106, 305–327.
Stich, S., and Warfield, T. *Mental representation: A reader*. Malden, Massachusetts: Blackwell Publishers.

au: year? 1994

au: correct city? Yes

au: Stich and Warfield or Warfield and Stich?