

“WHO SEES WITH EQUAL EYE, . . .
ATOMS OR SYSTEMS INTO RUIN HURL'D”?¹

(Received 4 February, 1993)

Brian McLaughlin's paper is an admirably clear presentation of the current debate over the relation of classical and connectionist models of cognitive processes. What is particularly welcome is the clarity with which McLaughlin emphasizes that this debate must ultimately be resolved in the laboratory. It is presumptuous of philosophers to make predictions about how things will turn out, but since we get neither money, fame, nor much respect, the pleasures of presumption are all the more enjoyable. Even so, I intend to be minimally presumptuous here, for I have not come to praise connectionism so much as to question the strength of the attack McLaughlin orchestrates. I will first briefly review how the problem of systematicity has been characterized, and then spend most of my time articulating three problems that I believe raise some worries about whether there is a well-defined problem here at all. Last, I shall make a few remarks on the notion of implementation. There is a great deal here that I would like to dally over, but time is short.

I. SYSTEMATIC CAPACITIES

A. *The Classical View of Systematicity*

McLaughlin characterizes systematicity in terms of statistical correlations among [the possession of] cognitive capacities, and the examples he gives are fairly revealing:

- (1) [possession of] the capacity to see a square shape as above a triangular shape is highly statistically correlated with the [possession of] capacity to see a triangular shape as above a square shape.

primitive capacities to see lines and angles. Thus these two capacities do involve the same concepts, though they don't show it on the surface. Needless to say, this analytic strategy has proven quite useful on numerous occasions.

But there is probably a limit on the use of such explanations of correlation. The classicist seems committed to the existence of primitive capacities, and correlations between them will not be explicable by showing that they are complexes of a common set of more primitive capacities. If there are correlations between primitive capacities, they are, for the classicist, brute facts. There seems no analysis of red or green that would explain why the capacity to see the one is highly correlated with the capacity to see the other — more highly correlated, in particular, than with seeing blue.

At first blush, this harbors no problem for the classicist. McLaughlin recognizes that classicists cannot offer a decompositional analysis of the primitive capacities; that, he is more than willing to cede, is something that might well be in the domain of the connectionists. Are we then to assume that the primitive capacities are atomic, independent of each other? Can't there be statistical correlations in the presence of atomic capacities? If there are, and the connectionist can somehow explain them, why *must* the connectionist resort to classical forms of explanation in dealing with 'nonprimitive' capacities?²

Perhaps, of course, correlations among nonatomic, nonprimitive capacities can be explained only classically, but even if there is some argument to this effect, it presupposes that another hurdle had already been cleared. The classicist is committed to the notion that there are certain atomic, primitive cognitive capacities. Such a capacity would have to be one that involves no other capacities, or at least none at the same level of analysis, and just which are those? As I read the history of conceptual analysis and AI, one should be quite skeptical about the prospect of isolating any clear set of primitives.

Since the explanation of any capacity, ability, or disposition standardly involves showing how it can arise from simpler (but not necessarily simple) capacities, abilities, and dispositions, it's hard to imagine rock bottom short of quantum physics. Yet the classicist needs to hold, I think, that there is some relatively clear line to be drawn between atomic psychological capacities and the capacities in terms of which

they will in turn be explained. But this is one of the things a good connectionist doubts, and the best efforts of numerous brilliant workers has not produced noticeable progress on this score.

C. *Independence and Statistical Correlation*

In order to be unproblematic, talk of statistical correlations or conditional probabilities between capacities must assume that the capacities are independently identifiable. This does not mean, of course, that there can be no conceptually related ways of describing the capacities. It does mean that if the *only* ways we have of picking out the capacities involve conceptual connections among them, there will be less work for a statistical or ultimately causal account of their connection to perform. There may still be a causal story to be told, for presumably some causal account of the capacities underlies any true conceptual account, but it may be difficult to disentangle the causal and conceptual levels of description and explanation. If the identity and individuation conditions of the cognitive capacities in question are not sufficiently independent, the attempt to describe the phenomenon to be explained as the presence of statistical correlations between the possession of different capacities breaks down, and the explanation the classicist offers may be a conceptual account masquerading as a causal account. Since the capacities McLaughlin is concerned with are capacities to have intentional states, he intends to individuate them by individuating the intentional state involved.

There are two questions to be raised here: First, can we identify intentional states (and thereby cognitive capacities) independently of each other? If we individuate cognitive capacities derivatively from the semantic content of the intentional state involved, then the independent specifiability of the capacities will depend on how independently specifiable semantic contents are.³ That is, the degree to which one is a semantic holist will predict one's willingness to accept the independence of the cognitive capacities. The more holistic one's semantics, the less likely one is to think that cognitive capacities are independently specifiable states connections between which are to be explained by law-like statistical regularities. Fodor's classicism and his aversion to semantic holism are pieces of the same cloth. Semantic holists, on the other

hand, must wonder what all the excitement is about statistical correlations between cognitive capacities — it's a lot like scurrying around to find an empirical theory to explain the astronomically high correlation between husbands and wives.

If I can find a being that can respond differentially to squares but not to triangles, should I still say it has the capacity to recognize or see squares? What this differential response is evidence of will depend to at least some degree on what other patterns of response we find in the organism. Surely Clark's assertion that the systematicity of cognition is a conceptual truth belongs here — namely, as a point about the holism of semantics. For the holism of semantics, if true, is conceptually true, as far as I can tell. If we identify and individuate meanings by their places in complex systems, then any intentional state we identify and individuate by its semantic content will also have to have systematic relations to other semantically related intentional states. If not, we have no right to identify it the way we did to begin with. (Since this does not determine just how faithfully the relations among intentional states must mimic the semantic relations, on this reading, Clark's assertion does not conflict with Dennett's assertion that systematicity may not be as thoroughgoing as Fodor, Pylyshyn and McLaughlin think, for the holism of semantics does not place specific requirements on the rigor of the structure of the cognitive system.)

Indeed, my reflections on this debate have led me ever more strongly to the conclusion that it is at root still a debate about semantics — not in the sense that the participants are quibbling over mere words, but in the sense that the key to the whole puzzle is semantics, how representations represent. For the real attraction of classicism is not a theory of how there can be correlations among distinct capacities, but how organisms can be in states with the incredibly fine-grained structure of intentional states at all. That intentional states involve compositional structures of a finite collection of primitives is the only answer to this puzzle that we really command; it is the only answer we can currently employ with confidence in building our own imitation intentional systems. Systematicity as characterized by Fodor, Pylyshyn, and McLaughlin is difficult for the connectionist primarily because, having rejected the compositional structures of the classicist to begin with, we aren't yet sure how exactly a connection system can be in a fine-grained

intentional state in the first place, much less how it might connect to some other, different but equally fine-grained intentional state. This is the challenge of elaborating a fully worked out theory of distributed representation, one that, to my meagre knowledge, has not been satisfactorily met.

This leads directly into the second question raised by worrying explicitly about the identity and individuation conditions of intentional states: just how fine-grained a system must we posit? Almost everyone admits that language-using organisms will require a fine-grained system that enables very subtle distinctions between contents of intentional states. And currently only a classical system of symbols in a combinatorial syntax and semantics is sure to fit that bill. But is an equally fine-grained, and therefore equally rigorous system required to identify and individuate the intentional states of infralinguals? The classicists seem to assume that one grain fits all, but if we could make do with a far more coarse-grained typology of intentional states for infralinguals, it is not clear that the full classical architecture is the only approach that gives us enough distinctions of the right kind to do the job.

When Dennett, for instance, suggests that a monkey might think that a lion wants to eat him, but might be quite incapable of thinking that he wants to eat the lion, he is suggesting that we should not project the kind of systematic relations that are so easily available to language-users like us onto the mental states of infralinguals. The mental states of infralinguals will undoubtedly also exhibit systematic, that is, regular, relationships — but perhaps not relationships that require the full classical architecture to model.

D. *Systematicity and Generality*

For, until we have a clearer sense of how to individuate cognitive capacities, it is far from clear that at least some of what is claimed to be systematicity in cognition is not simply *generality*. That organisms capable of seeing red triangles as above blue squares can also see blue squares as above red triangles may be due to the fact that they have a *general* capacity to identify relative spatial location. Employing a systematic rule-governed set of representations may be one way of explaining generality — but is it the only way? Pattern recognition —

something connectionist systems are notably good at — can clearly be general without classical systematicity.

The capacity to throw a red ball and then a green cube is highly correlated with the capacity to throw a red cube and then a green ball — but throwing, I take it, need not have a combinatorial structure to explain this fact. Throwing is a general capacity applicable to most things we can heft. To think that there is a capacity-to-throw-balls and a different capacity-to-throw-cubes the co-presence of which in one organism needs explanation is at best odd. One can imagine strange circumstances in which those capacities are separable and an explanation of their compresence does require explanation, but not in the context of our world and evolutionary history, I believe. The throwing capacity developed in a world in which enough different shapes presented themselves that specialization to one shape alone would have had little value in comparison to a more generalized capacity.⁴ Perhaps the capacity to desire a green triangular object and a red square is highly correlated with the capacity to desire a red triangular object and a green square object because there is a general capacity to desire that developed in a world in which the objects of desire differed enough to guarantee a generalized ability. Perhaps instead of bringing to the world a representational system already equipped with the universality of logic, we have only certain general capacities with enough slack to handle the vagaries of the world in which they function. Until we know better how to individuate capacities, it seems possible that the capacities of infralinguals are interestingly general without necessarily being systematic in the sense in question here.

II. PROBLEMS IN IMPLEMENTATION

I don't have much space left, but while I am muddying the waters, I think it is important to point out that what counts as an implementation is not really all that clear. Fodor, Pylyshyn, and McLaughlin want to claim that it is most likely that connectionism will figure into the overall cognitive story as the sub-psychological structure in which a classical architecture is implemented. Under exactly what conditions is one architecture (call it A) implemented in a different architecture (call it B, for base architecture)? It seems a necessary but clearly not sufficient

condition that whatever structures in B are supposed to implement A preserve the input-output characteristics of A. Further, there must be some mapping from each of the possible states of A onto (sets of) states of B such that all relevant functional relations are also preserved, that is, the relevant causal relations among the states of A must map onto the causal relations among the (sets of) B-states. I'm not sure this is yet a fully sufficient set of conditions for implementation, but even so there are important questions that need clarification.

Making sense of a partial implementation doesn't seem a problem, but what is the right thing to say when the mapping from A to states of B is not clean; B only approximates the configurations of A? There are a number of options:

1. We could then say that B does not implement A.
2. We could say B is a faulty implementation of A.
3. We could say that B is a good, fair, poor implementation of A.

We don't usually spend a lot of time worrying about how to characterize a particular implementation relation, nor need we, for we can usually refer to the intentions of the system's designer to clarify whether it is appropriate to talk about an implementation at all, and to help us fix the standard for a good implementation. But, of course, with brains and cognitive systems no such information is available to help us. Absent a designer, just how applicable is the notion of implementation? Is the chemistry of nucleic acids an implementation of genetic theory? While I know what someone would mean if she made such a pronouncement, it seems dangerous to me, in that we usually use "implement" to describe processes driven from the top down. This danger threatens the classicist's claim as well. Nature did not develop brains in order to implement a classical architecture. Brains and cognition developed bottom-up; that a classical architecture is at least approximately true of them is a wondrous testimonial to the subtlety of evolution. We can, of course, explicitly abstract from such implicatures and develop a purely formal notion of implementation — but the issue in debate between the classicists and the connectionists is not a purely formal

one, for it concerns the relevant explanatory priority of two different models of mental mechanisms.⁵

If we had a great classical theory and a great connectionist theory of cognition, I'm not at all sure what would decide whether the latter implements the former. Of course, if we had these great theories, we probably wouldn't care much whether the relation is an implementation relation.

III. FINAL SALLY

Analogies are great fun, so I want to put into play one more analogy and see how it plays. For 2000 years every astronomer in the Western world operated under the assumption that any acceptable explanation of the movements of the heavenly bodies had to attribute them uniform circle motion, a constraint apparently first articulated by Plato in the *Timaeus*. There were lots of reasons and motivations for this constraint, and some absolutely beautiful systems were developed employing it. But the irregularities in the data kept forcing small betrayals of this fundamental commitment, ever more epicycles, then eccentricity, and finally the equant. The Copernican adjustment really didn't help all that much by itself, though it succeeded in getting rid of the major epicycles and the equant. The great simplification came when Kepler realized that the apparent irregularities in the data were perfectly regular if the planets moved in only approximately uniform (though still lawful) and only approximately circular motion. The circle is pretty, simple, and powerful, but it is the slightly messier ellipse that slides smoothly onto the messy data of planetary motion. The classical approach to cognition is pretty, simple, and powerful, but I sometimes wonder if the messier connectionist approach won't end up sliding more smoothly onto the messy data of cognition.

NOTES

¹ Pope, *Essay on Man*, Epistle I.

² The connectionist would not naturally accept the primitive/composed distinction the classicist uses to sort out the domain, for no cognitive capacity will be considered primitive by the connectionist. Of course, the connectionist will assume that many

cognitive capacities will turn out to be orchestrations of simpler cognitive capacities, but this doesn't entail any ultimate *atomic capacities*.

³ Note that as of now we have no other interesting way to individuate cognitive capacities than via intentional states, and no other interesting way to individuate intentional states other than via semantic contents. Insisting on *interesting* ways of individuating these things disqualifies totally unhelpful ways like "the intentional state John was in as he walked through the door before breakfast" or "the capacity Herbert used in getting out of the handcuffs."

⁴ It is true that one's capacity to throw a ball may be more highly developed than one's capacity to throw a cube, even though they are the same capacity. Through practice someone like Roger Clemens can do amazing things with a little sphere, where he might not be nearly as proficient with a cube. But is this because, as the classicist claims, the cube-throwing program that enables the cube-throwing capacity is not as good as the ball-throwing program that enables the (different) ball-throwing capacity, or is it simply that practice has smoothed the cooperative interactions of the general abilities to grip, propel, and release in the one case more than in the other?

⁵ McLaughlin discounts the idea that no parallel process can literally implement a serial process, but I have my doubts that this can be so easily overlooked. In any interestingly parallel process the causal relations among the relevant functional states will not map the causal relations in a serial process directly. But preserving the mappings of relations among the states is a very important part of being an implementation, even in a thin formal sense of the term.

Department of Philosophy
University of New Hampshire
Hamilton-Smith Hall
Durham, NH 03824
USA