Distributed Cognition and Extended Mind Theory

I. Introduction

It may seem natural to think of the mind as a stream of conscious experience occurring squarely behind the eyes, or perhaps as some single, persisting subject of these experiences, hovering in the center of the skull. The past hundred years of scientific thinking about the mind have challenged this view in a variety of ways. The most recent challenge, and the most striking to date, rests partly on the distributed nature of cognition – the fact that intelligent behavior emerges from the interaction of a variety of elements, some of which may be spatially removed from the locus of behavior. From such distributed models of cognition, many authors have inferred the extended-mind thesis, the claim that the mind itself spreads into the world beyond the boundary of the human organism.

Distributed cognitive models have had direct impact on, and to some extent have been inspired by, research in the social sciences. Studies of insect behavior, for instance, form a bridge between the interests of cognitive scientists and matters to do with group-level behavior: large numbers of social insects, each of which "mindlessly" follows such simple information-processing rules as "drop my ball of mud where the pheromonal concentration is the strongest," design elaborate nests. This illustrates both how intelligent-looking results can arise from a distributed – and one might think fairly unintelligent – process and also how their emergence might be social in nature: environmental conditions induce various sub-populations to play different roles in the life of the insect colony. Some robustly cognitive, human social processes also seem amenable to distributed theorizing: contemporary scientific results, in particle physics, for example, often involve the contribution of hundreds, or even thousands, of individuals and instruments; here, each individual exercises a rich set of her own cognitive resources while playing a role in a much larger, highly structured enterprise. An intermediate case might be the modeling of traffic patterns: individual humans can reason in flexible and complex ways about driving and routes of travel, but, constrained by the presence of other automobiles and surrounding infrastructure, drivers' contributions to traffic flow, and the resulting traffic patterns, have much in common with large-scale behavioral patterns of social insects.

The remainder of this entry consists of three sections. Section II describes distributed cognitive modeling and the extended-mind thesis in more detail. Section III reviews critical reactions to the extended-mind thesis. Finally, Section IV briefly discusses fruitful areas of ongoing research on distributed cognition and the extended mind.

II. From Distributed Cognition to the Extended Mind

In the early twentieth century, Sigmund Freud proposed that subconscious mental states sometimes drive human behavior, thereby taking a significant step away from the commonsense view of the mind bruited at the outset. Given that the Freudian subconscious deals primarily in emotions and desires, however, Freudian innovations might seem to have little to do with intelligence or rationality *per se*. In contrast, the cognitivist revolution and its computer model of the mind marked a wholesale departure from the Cartesian view, at least with regard to the role of introspection in the life of reason. The new cognitive science set out to model the mechanisms the operation of which accounts for central aspects of human intelligence – memory, reading, means-end reasoning, perception, speech-processing, etc. – with no particular emphasis on consciousness. The overriding goal was to model behavioral data or produce simulations of human-grade performance; it mattered not whether the details of such models jibe with the revelations of introspection. From the standpoint of the new cognitive science, all that matters is that a series of mathematically defined states, causally connected by precisely defined operations, produces intelligent behavior.

The demotion of consciousness together with the focus on mechanisms paved the way for distributed accounts of cognition. If one need not ground one's models in a conscious, first-person perspective, one is free to look for cognitive mechanisms wherever the causal contributors to human behavior might be. Think again of the historical context. The computer model of the mind enjoyed substantial empirical success, but much of this in artificially limited domains or in the solution of narrowly circumscribed problems. Special-purpose systems can diagnose patients given lists of symptoms, but such systems do not exhibit human-like intelligence, flexibly and smoothly navigating their way through the real world, in real time. In response, a new breed of cognitive scientists began searching for a richer – or at least different – set of causal contributors to intelligent behavior, influences that might help to produce human behavior without requiring that their contributions take the form of computer code. Much of this research focused on the entire, interactive system of brain, body, and world and, in doing so, cast the constraining influence of environmental structure in a new light. In addition to being a source of input, the environment appears to work together with bodily processes to guide intelligent behavior – both literally, as in navigation, and metaphorically, as when it limits a decision-maker's options. The child does not learn to walk via the maturation-based appearance of an innate, internally stored motor program, for example; rather, walking emerges as the joint activity of the child's musculoskeletal system, the solid floor, and gravitational fields, all interacting as physical contributors to a stable cycle of movement through space. The visual system does not explicitly encode the proposition that objects in the environment are likely to be relatively rigid. Instead, the visual system computes the layout of the environment in a way that works effectively only in environments in which most of the objects are rigid. The rigidity assumption is left implicit and, thus, when the visual system functions properly, it does so by pooling its own contributions together with those of an environment that contains rigid objects. Philosophers took notice of, and sometimes contributed to, this developing trend in cognitive science, arguing, for example, that from an evolutionary perspective, we should expect environment-dependent cognitive processing to be widespread.

Into this intellectual foment, Andy Clark and David Chalmers cast "The Extended Mind" and in doing so crystallized the vanguard's vision. This vision draws heavily on empirical work related to distributed cognition. For example, a central notion from the literature on distributed cognition is that of an epistemic action. Such actions do not bring the physical world any closer to a state in which the subject's goal has been achieved; instead, epistemic actions divulge information about out how the subject might best pursue her goal. The rearranging of Scrabble tiles on one's rack, for instance, does not change the state of the game board so that it is any closer to having a new, completed word on it; rearranging letter adds nothing to the board itself. Rather, this epistemic action provides information to the player's cognitive mechanisms so that she might more readily identify a word among her set of letters. Such actions take place partly in the environment, and thus we can see how elements of the problem-solving process itself might be outside of the organism. What matters with regard to cognitive-scientific explanation is the role these contributors play in the production of intelligent behavior, not whether the contributors happen to appear within the boundary of skin and skull. Given the plausible additional premise that at least part of one's mind is physically located where one's cognitive processes are located, empirical work on distributed cognitive processing supports the extended-mind thesis.

The preceding style of reasoning also applies to such pedestrian mental states as belief. Perhaps the contents of most of a typical individual's beliefs are neurally encoded. Nevertheless, such contents could be encoded externally – in a trusty notebook, for example – and, if these external encodings were to interact with the organism in the right way, they would constitute part of the material basis of human mental states. So long as (a) the subject carries the notebook with her – the information in it reliably available – (b) she trusts what's written in the notebook, and (c) she can fluidly deploy the information stored there, the notebook does not differ relevantly from internal resources, at least with regard to its role in the production of intelligent behavior. Thus, mental states recognized by our traditional, commonsense understanding of ourselves – beliefs, memories, perceptions – can reside, at least partly, beyond the boundary of the human organism. Moreover, new cognition-enhancing technological devices appear almost daily, and thus we can expect this phenomenon to become ever more pervasive.

III. Critical Reaction

Sometimes proponents of the extended-mind thesis press a merely prejudice-removing agenda; they hope to neutralize readers' internalist biases by showing how cognition and mind *could* comprise environmental states or processes. Illustrative examples (such as the

Scrabble example given above) play an indispensable role in this project, but these illustrations often lead double-lives, being meant also to suggest that extended cognition surrounds us in the actual world. Here is where the more substantial question lies: Do cognition and mind actually extend into the environment in deep and substantial ways, extensively enough to ground a revolution in cognitive science and a re-conception of ourselves as persons?

Advocates for the extended view typically claim that (a) it constitutes the best interpretation of existing cognitive-scientific results, (b) it offers the most promising framework for future research in cognitive science (even if the existing results don't conclusively support the extended view over its rivals), or (c) it captures the deep nature of the mind – where the plausibility of (c) derives from the plausibility of (a) or (b). It is no surprise, then, that critics often attack the cognitive-scientific basis of the extended-mind thesis.

In this vein, some critics have charged that an extension-friendly cognitive science is bound to lack scientific unity. External objects and states, and the processes in which they participate, differ significantly from the internal ones heretofore fruitfully studied in cognitive science. The behaviorally relevant causal profile of neurally encoded memories, for instance, differs enormously from the causal profile of, say, graphite scratches in a notebook. It may well be useful to study the complex ways in which humans interact with and exploit the environment during problem solving; but the external processes so studied will almost certainly differ in substantive ways from internal ones – at least if attention is directed to real-life cases involving paradigmatic cognitive capacities. As a result, there will be no reason to treat the entire run of states and processes, internal and external, as cognitive – as being of essentially the same kind. This thought is closely related to another: even when the organism enters into an ongoing causal relation with portions of the environment, that does not by itself extend cognition or mind into the environment. Reasoning to the contrary commits the so-called causal-constitution (or coupling-constitution) fallacy.

Thus it is demanded that philosophers of cognitive science produce a mark of the cognitive or at the very least some cluster of central traits that tend to qualify a state or process as genuinely cognitive. By adverting to such a mark, proponents of the extended view might explain what unifies internal and external processes during extended cognitive processing and also make clear the conditions under which an ongoing causal connection brings into existence a fully cognitive process.

Critics themselves have proposed various marks of the cognitive – or at least necessary conditions on something's being genuinely cognitive – to ground their arguments against the extended-mind hypothesis or to undercut arguments in support of the extended view. Mental representations, for example, have played a central theoretical role in cognitive science. Perhaps, then, in order that a process be genuinely cognitive, it must consist in the manipulation of representational states – that is, states that picture, refer to, or describe parts of the world. Even if some external units are representational, their representational status derives from our interpretation of them, which is thought to show that external processes are not cognitive in their own right.

A further critical approach contrasts the extended view with others that fall under the general rubric of 'situated cognition'. According to one of these, the embodied view, nonneural bodily states or processes constitute significant parts of human cognition.

According to the embedded view, human cognitive processing is highly interactive and exploits structure in the environment, even though the external materials don't become proper parts of cognition itself. The critic claims, though, that data thought to support the extended view should be taken to support the embodied or embedded view instead. Once we have established the utility of a persisting cognitive system located in the body, and have explained how it interacts with the passing parade of external stimuli and materials, it appears gratuitous to say that, when the integrated, persisting cognitive system – typically housed within the body – interacts with the external materials, a new, fully cognitive system comes into being.

For all that's been said, this critical approach might seem to be a wash: the relevant data cohere well enough with either an extended or embedded-embodied view. This would be to ignore the central theoretical role of the relatively integrated, persisting system. It is the core construct in cognitive-scientific modeling and, as such, seems to offer the most promising basis on which to distinguish merely causal contributors to intelligent behavior from those processes the location of which bears on the location of the mind. All forms of cognitive modeling – computationalist, connectionist, dynamicist, as well as the brute-biological – specify a set of integrated elements and operations governing their interaction; the functioning of this system explains intelligent behavior as the result of interaction between the system and whatever materials or stimuli it happens across. On this view, then, being a state of a part of the integrated, relatively persisting system serves as a necessary condition for something's being a genuinely cognitive state. This constitutes a criticism of the extended view only because, and to the extent that, for humans, the integrated, relatively persisting cognitive system appears within the organism.

The consideration of a final critical perspective will bring more clearly into focus the relation between cognition and mind. One might accept that cognition is distributed but think that our enduring and stable conception of the mind should play a winnowing role when we evaluate the implications of distributed models of cognition. It is arguably a central feature of our concept of the mind that it be the locus of control of intelligent behavior. Thus, even if we bracket the demand for a mark of the cognitive and even if the persisting cognitive system stands as merely one element of many in cognitive-scientific modeling, this system has privileged status: it serves as the locus of control of the various forms of behavior cognitive scientists hope to explain; therefore, where we find it, we find the mind. If the integrated, relatively persisting cognitive system appears within the human organism – as seems likely to be the case for most humans, most of the time – then so does the human mind.

IV. Future Directions

Debate about distributed cognition and its philosophical implications remains lively, as does the development and testing of distributed models. Among others, the following areas of research – of particular note in connection with the social sciences – flourish.

Some authors have attempted to characterize mathematically the relevant structural properties of integrated intelligent systems. According to one suggestion, a wide range of networks – including the brain and well-functioning computer networks – exhibit a distinctive trade-off between the specialization of their components and the sharing of information among those components.

Along similar lines, structural properties have been sought at the level of social groups. For example, memory researchers have found that people who work together to solve problems often take on specialized roles within the group (say, with regard to what sort of information they're responsible for remembering); the distribution of such roles facilitates some efforts but may hinder others. This sort of investigation might help us to decide whether group minds appear among the extended minds. It also has the potential to contribute substantially to research in the social sciences – from the analysis of voting behavior to the study of the dynamics of online communities.

Lastly, it's worth emphasizing the contribution dynamical-systems-based approaches have made to the discussion of distributed cognition and the extended-mind thesis. Computer-based and neural-network-based metaphors have dominated research in cognitive science. Nevertheless, a significant minority of cognitive scientists instead view the mind as something more like a physical system in motion, one best described using differential equations that express dependence relations among various continuously interacting quantities. On this view, cognition occurs when the organism enters, even for a short time, into an interdependent relation with physical systems beyond the boundary of the organism – with each component in the system constraining the other in a relationship of ongoing mutual determination. Proponents of such dynamical modeling often claim that cognition and mind permeate the entire system: the organism and the external materials to which it's coupled. It remains to be seen whether the dynamical-systems approach can be used to model paradigmatically cognitive activities such as long-term planning and theoretical reasoning. Nevertheless, exploration of dynamical processes should help us better to understand at least some cognition-related interaction between thinkers and their environments.

Robert D. Rupert

See also Collective Agents; Collective Intentionality; Embodied Cognition; Group Mind; Situated Cognition

Further Readings

Adams, F., and Aizawa, K. (2008). *The bounds of cognition*. Malden, MA: Blackwell.

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*, *1,* 7–19.

Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: MIT Press.

Hurley, S. L. (1998). *Consciousness in action*. Cambridge, MA: Harvard University Press.

Noë, A. (2004). *Action in perception*. Cambridge, MA: MIT Press.

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

Rupert, R. D. (2009). *Cognitive systems and the extended mind*. Oxford: Oxford University Press.

Wheeler, M. (2005) *Reconstructing the cognitive world: The next step*. Cambridge, MA: MIT Press.

Wilson, R. A. (2004). *Boundaries of the mind: The individual in the fragile sciences*. Cambridge: Cambridge University Press.