How to Situate Cognition: Letting Nature Take its Course

Article · January 2005

DOI: 10.1017/CBO9780511816826.004

2 authors:

Robert A. Wilson
University of Alberta

Andy Clark
The University of Edinburgh

To appear in M. Aydede and P. Robbins (eds.),
The Cambridge Handbook of Situated Cognition

How to Situate Cognition: Letting Nature Take its Course*

Robert A. Wilson and Andy Clark

1. The Situation in Cognition

The situated cognition movement in the cognitive sciences, like those sciences themselves, is a loose-knit family of approaches to understanding the mind and cognition. While it has both philosophical and psychological antecedents in thought stretching back over the last century (see Clancey, this volume, Gallagher, this volume), it has developed primarily since the late 1970s as an alternative to, or a modification of, the then predominant paradigms for exploring the mind within the cognitive sciences. For this reason it has been common to characterize situated cognition in terms of what it is not, a cluster of "anti-isms". Situated cognition has thus been described as opposed to Platonism, Cartesianism, individualism, representationalism, and even computationalism about the mind. A cluster of positive characterizations of the situated cognition movement has also been given, both in terms of adjectival descriptions based on the approach of one or more figure of influence (e.g., Gibsonian, Heideggerian), and in terms of a medley of related but under-explicated notions, such as embodiment, embeddedness, and externalism. Importantly, researchers who self-identify with the “situated movement” in the cognitive sciences have not been paralyzed by a relative lack of attention to the conceptual articulation of their paradigm. They have instead gotten on with the task of showing how the situated perspective leads to interesting and novel approaches to understanding particular cognitive abilities.

In this chapter we do some conceptual stocktaking. We propose a way of thinking of situated cognition that captures at least one important historical strand to the situated cognition movement but, more importantly, which also provides the field with some normative direction. In a nutshell, the basic idea is that we should think of situated cognition as a form of cognitive extension, or, rather, as a variety of forms that such cognitive extension can take. Our aim is not to specify the essence of situated cognition (a misplaced goal, here as in many places in philosophical theorizing), nor is it to do justice to all of the work that has, at various times, been referred to under the heading "situated cognition" (or one of its nom de plumes). Rather, it is to provide a way of conceptualizing situated cognition that helps both to focus and reorient the study of cognition as a situated phenomenon.

The chief innovation of the chapter thus lies in a preliminary articulation of the variety of forms that cognitive extension can take. This articulation involves two main dimensions, one tracking the nature of the augmentative resource, the other tracking the durability and reliability of the augmentation. The resulting matrix captures many forms of cognitive extension that seem almost trivial and can be taken for granted by most parties to the debate over the nature of cognition. But it also accommodates a variety of forms that cognitive extension can take that challenge us to reassess what cognition is, who we (as cognizers) are, and what the future holds for the study of cognition. In addition, once we provide such a nuanced framework for thinking about the varieties of cognitive extension, we are better able to respond to several important recent objections to the conception of situated cognition as cognitive extension.

To set the scene, we begin with some potted history, in the form of a brief tour through some of the anti-isms and positive characterizations of situated cognition. One caveat, however, before we get going. An important strand to the situated cognition movement that we are unable to discuss here directly concerns the ways in which cognition is embodied. There is a range of discussion of this idea elsewhere in the volume, as well as in our own previous work (Clark 1997, 2003; Wilson 2004), but it remains a separate project to relate the embodiment of cognition to the idea of cognitive extension that we are concentrating on here. Suffice it to say that we think that taking seriously the embodiment of cognition will reinforce the perspective that we are developing, primarily because many forms of embodied cognition, properly understood, will turn out to involve just the kinds of cognitive extension that we articulate here.

2. Situated Cognition: A Potted Recent History

Cognition, it is widely agreed, both stems from and generates the activities of physical individuals located in particular kinds of environments. But despite this general consensus, the views of cognition and the mind that constituted a kind of status quo in the field by the end of the 1970s took a particular view of the relationship between cognition (or the mind), individuals, and environments, a view that was typically understood to have fairly direct implications for how to study cognition and the mind. According to this view, the cognitive sciences were to embrace what Jerry Fodor (following Hilary Putnam) called "methodological solipsism", and were, in effect, to bracket off the world beyond the individual in characterizing and individuating cognitive states and structures. We will follow common practice and refer to this as the individualistic conception of the mind and cognition (Wilson 1995, 2003, 2004a, 2004b).

On this conception, cognition takes place inside the head, wedged between perception (on the input side) and action (on the output side), constituting what Susan Hurley (1998) has called a kind of cognitive sandwich. It involves the computational processing of mental representations, where these are language-like both in their constituents—concepts as words—and in their structural composition—mental representations as generated by, and decomposable in terms of, an underlying mental syntax. Such processing relies purely on features of the symbols themselves and the rules by which they are manipulated, rather than on broader features of the individuals who instantiate them and the environments in which they operate. On the version of this view most influential amongst philosophers of mind, much of this cognitive architecture is universal across the species and innately specified, and the task of developmental and cognitive psychologists is to uncover these innate structures, and to understand how it is that they eventually gave rise to the diversity that we appear to see in everyday cognitive activity.

Even though cognition takes place inside the head, on this view, cognition is not simply to be given a neural description. For cognition is taken to be, in an important sense, substrate-neutral: it is not what you are made of but how you are functionally organized that matters for cognition. This view, functionalism, has sometimes (e.g. Clark 1989, p. 21) been presented as reflecting the old blues adage that "it ain't the meat it's the motion" that matters—in this case, for mind. In truth, however, within the individualistic paradigm for the study of cognition that we are sketching there wasn't so much attention to motion as to functional structure and organization. That's why researchers constructing computer programs (and sometimes program-driven robots) could view themselves as contributing not simply to the study of cognition through simulation, but to the enterprise of creating genuinely cognitive beasts, with "artificial intelligence". And that's at least one reason why the neurosciences were often viewed, at least until the 1980s, as a kind of junior partner in the venture to understand and explain cognition. Thus, while the neurosciences could tell us much about the <u>realization</u> of cognitive processes and structures in certain kinds of cognizers, their findings could not usurp the investigations of psychologists, linguists, and computer scientists, who were exploring the functional structure of the mind.

Fodor's <u>The Language of Thought</u> (1975) and <u>The Modularity of Mind</u> (1983) are particularly striking expressions of this kind of view in philosophy, with Fodor taking much of his cue from Noam Chomsky's work in linguistics. But this view of cognition was manifest across the cognitive sciences more generally, including in psychology and artificial intelligence. Exemplars include Zenon Pylyshyn's <u>Computation and Cognition</u> (1984) and the physical symbol system hypothesis (see Newell and Simon 1976, Newell 1980) according to which cognition is the algorithmic manipulation of symbol structures. Perhaps the best-known, large-scale project in artificial intelligence that steadfastly adopts a classic, individualistic perspective is the CYC project, an ongoing, 20-year project to create a general-purpose, common sense reasoner (see www.cyc.com).

Philosophical doubts about individualism were first raised in the now classic arguments of Hilary Putnam (1975) and Tyler Burge (1979), both turning on the question of whether individualistic views of the mind and cognition could adequately account for meaning or mental content. These original challenges to individualism were cast in terms of whether psychological states, particularly intentional states, should be individuated or taxonomized in accord with the constraint of individualism. For this reason one of us has elsewhere (Wilson 2000, 2004a) called the resulting forms of externalism <u>taxonomic externalism</u>; it has also been called "traditional externalism" and "philosophical externalism". The general idea was to continue to view intentional or content-laden mental states as some kind of internal state of the individual, but to argue that, nonetheless, because of their content, they did not supervene on, or were not metaphysically determined by, what fell within the physical boundary of that individual. Hence the debates over whether folk psychology was individualistic (Burge

1979, Fodor 1982, Loar 1988), whether the notion of content used in David Marr's celebrated computational theory of vision was internalist or externalist (Burge 1986a, Egan 1992, Segal 1989, Shapiro 1997), and on the issue of the relationship between individualism and the normativity of the mental (Burge 1986b, Ebbs, 1998, Pettit 1993).

More radical forms of externalism about the mind—what one of us (Wilson 2000, 2004a) has called locational externalism, Mark Rowlands (1999) calls environmentalism and, following Andy Clark and David Chalmers (1998), many call the extended mind thesis—hold that the mind and the cognitive processes that constitute it extend beyond the boundary of the skin of the individual agent. The extended mind thesis identifies cognitive systems themselves as reaching beyond individuals into their physical and social environments. The thesis challenges individualism directly by implying that an individualistic psychology could only, at best, tell part of the story about cognitive processing: the inside story. An early gesture at such a view was wide computationalism (Wilson 1994), the view that the computational systems that make up the mind can extend into and include as a part of themselves, aspects of an organism's environment, a view that we will discuss further below.

Locational externalism, environmentalism, and the extended mind thesis are radical forms of externalism in at least two ways. First, they do not rest on claims and intuitions about whether the content of a pair of states of two individuals in different environments (or one individual in two such environments over time) is the same or different, or on questions about how particular intentional states are taxonomized, or on ones concerning the role of the physical or social environments in individuating such states. Instead, they appeal to the nature of psychological processing, to the arbitrariness of the head (or the skin) for bounding cognitive systems, and to the structure of real-life, on-line cognitive activity in the world. Thus, if the extended mind thesis is true, it is true in virtue of something implementationally deep about cognition, rather than some debatable view of mental content. Second, the extended mind thesis is not simply a view of how we "talk about" or view cognition and the mind—about the epistemology of the mind, one might say—but about what cognition and the mind are—about the ontology of the mind. We thus arrive at the basic idea of this chapter: that work in situated cognition is best viewed as an ongoing series of investigations of the space of cognitive extensions.

3. Extensions in Biology, Computation, and Cognition

Cognitive extensions, like house extensions, come in a surprising variety of forms. Some are truly, massively, staggeringly transformative, while others are content to project a previously existing theme. Some are seamless and elegant, while others expose a barely connected hodge–podge of warring materials. Some are visibly permanent, built to persist as long as the main structure, while others are more temporary, apt for rapid dissolution or replacement. Some appear homely, while others can seem alien (some find them monstrous). And the very idea of cognitive extension, just like its bricks-and-mortar cousin, can be apt to provoke objection and outcry. In this section and the next we present what seem to us to be the main varieties of possible cognitive extension. Distinguishing these varieties is, we believe, the crucial first step towards allaying the fears, revealing the attractions, and generally restoring peace to the neighborhood.

We should go further, however, and not simply think of cognitive and house extensions as alike but as instances of the very same kind of activity: the activity of modifying one's environment from the situation in which one find's oneself in ways that meliorate that situation. Thinking is a kind of building, a kind of intellectual niche construction that appropriates and integrates material resources around one into pre-existing cognitive structures. In cognition, agents modify or augment the capacities that those pre-existing structures enable. Part of the point of suggesting that we move beyond the "minds are like houses" metaphor and view cognition itself as a kind of building activity is thus to introduce a deeper analogy for understanding the idea of cognitive extension, one anchored in recent thinking about genetics, development, and evolution. Several recent paradigms in the biological sciences—in particular, niche construction theory, developmental systems theory, and the idea of an extended physiology—advocate a move beyond the boundary of the organism and an accompanying reconceptualization of some of the processes at the core of biology. The idea of cognitive extension can be very fruitfully approached by means of these examples (see also Wilson 2004b, 2005).

Niche construction theory has been articulated by John Odling-Smee and his colleagues over the past twenty years, and receives its most elaborate, synthetic exposition in Odling-Smee, Laland and Feldman's <u>Niche Construction: The Neglected Process in Evolution</u> (2003). Niche construction, they say "occurs when an organism modifies the feature-factor relationship between itself and its environment by actively changing one or more of the factors in its environment" (2003:41). Niche construction (see also discussion in Clark in press-a) is a widespread process in the natural world, encompassing not only the construction of nests and burrows in mammals, birds, and insects, but the manipulation of existing social structures, the off-loading of physiological functions to environmental resources, and the appropriation and adaptation of environmental resources as biological and cultural tools. But while niche construction has long been recognized as an important process in certain contexts, Odling-Smee et al. place it center stage in the biological sciences, viewing it as introducing a missing dynamic aspect to the theory of evolution that has radical implications for the biological sciences. Niche construction, on this view, provides the key to integrating ecology and genetics via an extended notion of inheritance, ecological inheritance. And, they have argued, the pervasive character of niche construction and its contrast with natural selection make it "nothing less than a second selective process in evolution" (2003:178).

Developmental systems theory derives from Susan Oyama's The Ontogeny of Information (1985), and has its flagship presentation in Oyama, Griffiths and Gray's collection Cycles of Contingency (2001). Developmental systems theory (DST) is rooted in skepticism about the emphasis placed on the role of genes in accounts of development and evolution that arose in the wake of Richard Dawkins' The Selfish Gene and the development of sociobiology in the mid-1970s. Proponents of DST argue that genes are but one (albeit important) developmental resource, and that the basic unit of development is the developmental system that (typically) contains genes as components. Genetic systems thus do not exhaust the kinds of developmental systems there are, with other developmental systems, such as the chromatin marking system, being epigenetic, while others, such as those for the transmission of behavior and culture, stretch not simply beyond the boundary of the nucleus of the cell but beyond the organism into the environment (see Jablonka and Lamb 2005). These wide or extended developmental systems have previously (and problematically) been conceptualized quite independently from individualistic developmental systems, separated by the divide between "culture" and "biology", or that between learning and inheritance, or environment and genes.

Finally, the ecological physiologist J. Scott Turner has recently presented a range of examples—coral reefs (consisting of polyps, calcite deposits, and protozoan flagellates), termite mounds (complete with their own interspecies communities), mole cricket burrows (see also discussion in Clark 2005)—in which, as Turner says of the first of these examples, an "important component of the physiological process takes place outside the animal" (2000, p.24), in this case, the polyp. Turner takes his inspiration from Richard Dawkins's idea of the extended phenotype, arguing both that physiological processes, such as metabolism, energy transfer, and homeostasis, extend beyond the boundary of the organism, and that this provides grounds for thinking of the physiology of organisms themselves as extending beyond their own outer membranes, whether it be skin, shell, or soft tissue.

These recent emphases on niche construction, developmental systems, and extended physiology in the biological sciences suggest blind spots in past biological thought and fruitful directions for future research. And the new suggestions all share a premise central to our idea of situated cognition as cognitive extension: that the individual organism is an arbitrary stopping point for the scientific study of at least a range of relevant processes in the corresponding domain. In the domain of cognition, no one is an island.

Time then to get a little closer to cognition itself. Central to much cognitive science is the assumption that cognition is computational, an assumption manifest in general appeals to the metaphor of the mind as computer, in the construction of computational models for particular problem-solving processes, and in the claims of so-called "strong AI" that appropriately programmed machines are (or would be, if we could only construct them) cognizers. At the start of this chapter we mentioned computationalism in the list of "isms" that have sometimes been thought to be challenged by the idea of situated cognition. We think, however, that any such inference is mistaken. It is mistaken because computation itself can be an extended process in just the sense in which we are suggesting that cognition can be an extended process.

This is the view that one of us has called wide computationalism (Wilson 1994, 1995: ch.3), and the basic idea behind it is simple. Traditionally, the sorts of computation that govern cognition have been thought to begin and end at the skull. Computationalism has thus been viewed as entailing an individualistic view of cognition. But why think that the skull constitutes a magic boundary beyond which cognitively relevant computation ends and mere causation begins? We are creatures embedded in informationally rich and complex environments. The computations that occur inside the head are an important but non-exhaustive part of the corresponding computational systems. A wide or extended computational perspective opens up the possibility of exploring computational units that include the brain together with aspects of the brain's beyond-the-head environment. Wide computational systems thus literally extend beyond the confines of the skull into the world.

The first point we want to make is that the idea of extended computation is entirely non-controversial in non-cognitive contexts. Indeed, it is presumed in a large variety of contexts in which computational techniques are brought to bear on understanding causal mechanisms and processes. Computational processes occur within discrete entities—whether they be biological cells, computer chips, or larger entities comprised of these units—but they can also occur between such units. Typically, such extended computational processes constitute a larger computational system, but that should not obscure the fact that, with respect to those units, genuine computational processes physically extend beyond the boundaries of those units. There is nothing ersatz, for example, about the computations that flow across a grid system or a LAN.

The second point is that there are at least some cognitive contexts in which the idea of extended computation should be just as uncontroversial. As an example, and to illustrate how extended computationalism modifies the traditional view of cognition as computation, consider Figures 1 and 2, which depict two ways in which one might multiple two three-digit numbers.

Traditional Computationalism

Multiplying with only internal symbols

Computational system ends at the skull; computation must be entirely in the head.

1. Code external world.

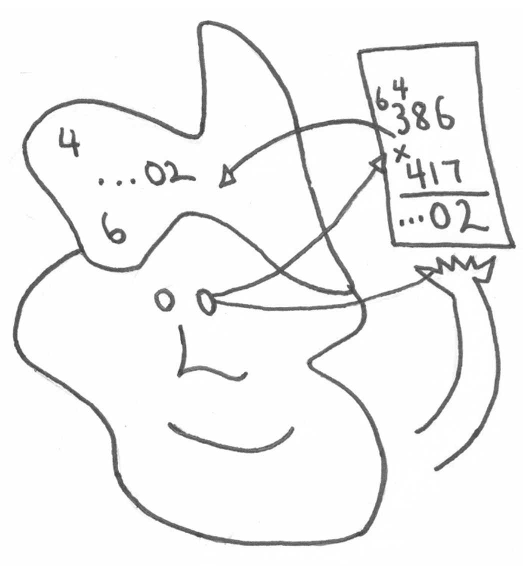
2. Model computations between internal representations only.

3. Explain behavior, based on outputs from Step 2.

Figure 1: Traditional Computationalism

On the traditional computationalist view, the first step in the process of multiplication is to code some form of input (e.g., visual or auditory input) from the world into internal symbols. The computations involved in multiplication then take place entirely between these internal symbols; computation ends at the skull. Contrast this with an extended computational view of multiplication, which involves some kind of internal symbol manipulation or computation, but which also involves the active manipulation of symbols—such as those on a piece of paper—that are not in the head, as depicted in Figure 2:

Multiplying with
internal & external symbols



Extended Computationalism

Computational system can extend beyond the skin into the world; computation may not be entirely in the head.

1. Identify representational or informational forms--whether in the head or not--that constitute the relevant computational system.

2. Model computations between these representations.

3. Behavior itself may be part of the wide computational system.

Figure 2: Extended Computationalism

On the extended computationalist view, multiplication begins as a causal process between external and internal symbols, but that initial relationship is then incorporated as part of the computational process itself. The resulting computational system itself is not restricted to what is inside the head, but includes both internal and external symbols. This allows one to understand the entire activity of multiplication as it is typically performed as a dynamic series of perception-action cycles that are computational in nature.

It is this conception of extended computation, we claim, that is invoked in much of the recent work in situated robotics (e.g., Brooks 1991a, 1991b, 1999), animate vision (e.g., Ballard et al. 1997), and cognitive anthropology (e.g., Hutchins 1995). It is not a distinct, novel, or ersatz kind of computation but simply an extension of the standard view of computationalism, just as the three examples that we began this section with—that of niche construction, developmental systems, and extended physiology—are extensions of quite standard corresponding views in the biological sciences.

Wide computationalism or extended computation thus provides the basis for one direct argument for extended cognition. Suppose that we grant the assumption of computationalism that has structured much of the work in cognitive science. If the kind of computation that at least parts of cognition involve are extended, then those parts of or aspects to cognition will also be extended. This is to reject the traditional individualistic understanding of computationalism, of course, but we have suggested that doing so does not require a novel or non-standard view of computation itself. Rather, it is to sift the wheat of computation from the chaff of individualism. Thus, far from being incompatible with computationalism, situated cognition as extended cognition follows rather directly from it.

But whether or not one accepts computationalism, and the claims we have made about it here, there remain a variety of forms that extended cognition can take. We turn now to examine two major dimensions that structure much of this variety.

4. Articulating the Idea of Cognitive Extension

The first dimension concerns the nature of the non-neural resources that are incorporated into extended cognitive behaviors, dispositions, and activities. Such resources may be natural, technological, or socio-cultural in nature, and each of these determine distinct kinds of cognitive systems. The second dimension concerns the durability and reliability of the larger (extended) system. The system may be a temporary and one-off construct, a temporary but repeatable one, or something more permanent. Let's take the two dimensions in turn.

Natural extended cognitive systems are those cognitive systems containing natural resources from the cognizer's environment that have been functionally integrated into the cognitive repertoire of that cognizer. A natural resource for a creature is any feature of its natural world that it draws on for some aspect of its continued functioning. Oxygen is a natural resource for respiration, and fruit (for humans and other primates) is a natural resource for nutrition. But some organisms also require cognitive resources, and many of these form part of the natural world of those organisms. Natural resources, including cognitive resources, can simply be used by organisms, but sometimes this use does not merely fuel a pre-existing system—as in the above-mentioned cases of the respiratory and digestive systems—but augment the system itself and the capacities that it possesses.

To take a non-cognitive example, empty shells are a resource that hermit crabs use for protection. Since this use both physically and functionally changes the relevant capacities of hermit crabs, it is plausible and relatively uncontroversial to see the crab together with its appropriated shell as a unified entity, one with capacities and abilities due in part to the shell that a shell-less crab alone does not possess.

As in this example, one should expect to find natural extended systems for cognitive functions just when there has been sufficient world-mind constancy for organisms to reliably exploit that constancy to lighten their internal cognitive load. Such exploitation manifests what one of us (Clark 1989) has dubbed the "007 Principle", know only as much as you need to know to get the job done. As Clark said in articulating this principle, in general, "evolved creatures will neither store nor process information in costly ways when they can use the structure of the environment and their operations upon it as a convenient stand-in for the information-processing operations concerned" (1989: 64). Perception is one prima facie likely domain in which this is true. One of us has previously argued (Wilson 1994, 1995: ch.3) that it is here that some of the best candidates for wide computational systems can be found, and (Wilson 2004a: ch.7) that Dana Ballard's animate vision paradigm also exemplifies the 007 Principle and posits natural extended cognitive systems for perception.

A second kind of extended cognitive system appropriates not natural but technological resources. In contrast with natural resources, technological resources are artificial in the sense of being made or developed by human agents. But like natural resources, they encompass a diverse range of resources, from those that are used in a novel, one-off manner, such as a book that you use on the spur of the moment to jam open a window, to those that represent permanent features of our everyday life, such as prosthetic limbs. In the cognitive domain, technological resources include dedicated cognitive artifacts, such as instruments for the measurement and recording of data, and those that extend our sensory abilities; make-do procedures, such as scribbling post-it notes as reminders and guiding one's immediate behavior by reference to improvised lists; and devices with more general functions, such as cell phones and other telecommunicative equipment that can be used for cognitive augmentation (see Clark 2003). Like natural extended cognitive resources, such technological resources can serve merely as inputs to a skin-bound cognitive system, but there is a range of cases where they do more than this and become functionally integrated into a larger cognitive system.

An electronics engineer usually has a pretty clear sense of what is mere input to a system and what is an integrated addition that alters the system itself. The distinction here is intuitive enough, even if it is one that has been surprisingly hard to pin down (see Haugeland 1998 for a classic attempt, and Clark in press-b for a slightly different treatment). Much obviously turns on degree and complexity of integration. But much also turns on how one conceives the goal or purpose of the system at hand. The purpose of the radio being to receive, decode and play contents borne by radio signals, anything locally added that helps it do so (e.g. a better transistor at some key point, a signal amplifier, etc.) looks like an augmentation of the system rather than a mere input. This is so whether the additional transistor or amplifier falls inside the pre-existing boundaries of the radio or lies beyond them.

In most cases where we are tempted to speak of cognitive augmentation, the same rule of thumb seems to apply: we find cognitive augmentation where new resources help accomplish a recognizable cognitive task in an intuitively appropriate manner, e.g., by enabling the faster or more reliable processing of information required by some goal or project. The new resource need not bear a compositional similarity to the rest of the system. Rather, it needs to display functional integrity together with the rest of some cognitive system that serves the kinds of purposes that that cognitive system has served: to perceive, to decide, to remember, to behave. More radical cases of cognitive augmentation may be possible—such as those in which the resulting system serves radically different purposes—but as with well-crafted devices generally, we think that cognitive extension tends to be step-wise, building on the solid achievements of systems that have already earned their keep in some particular domain.

A third kind of extended cognitive system is also worth distinguishing, though it might be thought subsumable under either the natural or the technological (or both). These are <u>socio-cultural</u> systems, which are formed when there is stable reliance by an individual in her cognitive activity on social structures, other individuals, and their cultural products. These structures and products serve as resources for a range of cognitive activities. Perhaps the most striking examples of socio-cultural cognitive resources involve writing systems, broadly construed, which have constituted a relatively durable, public cognitive resource crucial to education, training, regimentation, commerce, and military conquest in the Western world for millennia. But there are many others, including those that derive from practices of distributing cognitive labor between individuals, from the parental transmission of information to children, and from the establishment of ritual, musical, and ceremonial orderings.

For many individuals, such socio-cultural resources are like natural cognitive resources in that they can be taken for granted as part of the normal conditions in which their cognitive abilities develop and they learn particular skills and facts. They are a kind of cognitive oxygen, simply given as part of the natural world in which at least much cognition takes place. On the other hand, socio-cultural resources and the cognitive systems that they partially constitute are not biologically—genetically, physically, or evolutionarily—givens, since they have been created and modified by the activities of past generations of people. Socio-cultural resources are distinguished by their origin, but they are worth distinguishing in reflecting on extended cognition because, we claim, they constitute a crucial part of some of those cognitive abilities and activities that distinguish human cognition from its nearest neighbors. There may be animal cultures, but it is only in <u>Homo sapiens</u> that we find diverse cultures of cognition, social structures and products that, whatever their own origins, now significantly augment the cognitive capacities of individuals who are embedded in them.

As we saw with extended cognitive systems that are either natural or technological, socio-cultural extended cognitive systems exist when the appropriate type of resource is not simply used by an agent but becomes functionally integrated into the cognitive functioning of that agent (see also Clark in press-b). We think that the kinds of socio-cultural resources that we have mentioned often meet this additional criterion. Writing systems, for example, are not simply used by agents with given cognitive abilities but significantly augment the cognitive abilities that those agents possess, such as the capacity for short- and long-term memory, the ability to keep track of the relationship between abstract propositions as is often required in reasoning, and the expression of self-knowledge (cf. Goody 1977). Mathematical notation does not simply feed existing mathematical abilities (though it does that, to be sure), but builds on those abilities to produce an agent with significantly greater mathematical capacities. The difference between the ability to multiply using Arabic numerals, versus that using

Roman numerals, serves as a reminder of how much specific forms of writing can contribute to particular abilities here.

So much for the first dimension to extended cognition, whether the cognitive resources it incorporates are natural, technological, or socio-cultural. The second dimension concerns the durability and reliability of the extended cognitive system that results from the functional integration of such resources. Extended cognition, we have so far proposed, occurs when internal and external resources become fluently tuned and integrated so as to enable the larger system—the biological agent plus specific items of cognitive scaffolding—to engage in new forms of intelligent problem solving. Sometimes these larger problem-solving ensembles are transient creations, geared towards a specific purpose (doing the accounts, writing a play, locating a star in the night sky), and combine core neural resources with temporary add-ons such as pen, paper, diagrams, instruments etc.. At other times, they involve more stable and permanent relationships between biological agents and extended cognitive resources. We first consider the more transient varieties.

Consider, by way of a staging analogy, the idea of a task specific device (TSD) discussed by Bingham (1988). The notion of a TSD was introduced as a theoretical tool to help tackle the problem of understanding the organization of human action. In brief, a TSD is a temporary but highly integrated assembly created to accomplish some kind of goal. In the motor arena, a TSD is a soft-assembled (i.e. temporary and easily dissoluble) whole that meshes the dynamics that are inherent in the human action system and the so-called “incidental dynamics” contributed by various extra-organismic factors and forces. TSD's, that is to say, are “assembled over properties of both the organism and the environment" (Bingham 1988, p.250). In each specific case, the biological action-system will need to recruit some complex, non-linear combination of contributions from its four chief subsystems—the link-segment system, the musculoskeletal system, the circulatory system and the nervous system—and do so in a way expressly tailored to accommodate and exploit the incidental task dynamics introduced by, for example, a handle on a paint pot, a bouncing ball, or a windsurf rig out on the open sea. These examples span the three main kinds of incidental task dynamics identified by Bingham, viz. those tasks that simply introduce inertial and dissipative properties or mechanical constraints (as when we carry the paint pot by the handle); those that involve absorbing, storing and/or returning energy (as when bouncing a ball); and those that involve coupling with systems that have their own independent energy sources (the windsurf rig powered by the wind and waves of the open sea).

Why study such TSDs? One reason, the most obvious one, is that it is these very ensembles that are locally at work in many of the most distinctive cases of human action. We alone on the planet seem capable of creating and exploiting such a wide variety of action-amplifiers, ranging from hammers and screwdrivers, to archery bows and bagpipes, to planes, trains and automobiles. But a second reason, far less obvious, is that working backwards from the analysis of these complex wholes may itself contribute important insights concerning the contributions and functioning of the biological human action system itself. Although a natural first thought would be to try to understand each of the four main biological subsystems in isolation, then perhaps to look at their coupled interactions, then finally to add in the incidental dynamics, it turns out that this simple step-wise approach may be doomed to failure. The reason is that the potential behaviors of the whole biological action system are determined by staggeringly complex non-linear interactions between the four main subsystems and the incidental dynamics. The good news, though, is that in a typical TSD the degrees of freedom of this large and unwieldy system are dramatically reduced. The whole point, in fact, of soft assembling a task-specific device is to reduce the initially high dimensional available dynamics to a much lower dimensional structure and thus to establish an effectively controllable resource (see eg., Fowler and Turvey 1978, Salzman and Kelso (1987). As a result:

> The challenge is to work backwards from a description of the reduced dynamics to an understanding of the manner in which subsystem dynamics couple and co-constrain one another to produce the observed dynamical system. Because information about both task-specific dynamics and the individuated resource dynamics is required, the strategy unites the efforts of behavioral scientists and physiologists in an integrated and coherent effort. (Bingham 1988, p. 237)

We have described this strategy in a little detail because many of the key ideas apply directly, it seems to us, to the case of those extended systems that involve temporary, transient forms of cognitive augmentation. Let us label these 'Transient Extended Cognitive Systems' (TECSs).

TECSs are soft-assembled, i.e., temporary and easily dissoluble, wholes that mesh the problem-solving contributions of the human brain and central nervous system with those of the (rest of the) body and various elements of local cognitive scaffolding. To further probe the structure of the space of possible TECSs, we might distinguish cases according to the durability and reliability of the relationship between agent and resource.

Thus we might want to distinguish temporary, one-off relationships from those that, though transient, are regularly repeated. To solve a new brainteaser, an agent may generate a brand-new, one-off kind of TECS. While a practiced crossword puzzle solver, though perhaps confronted with a new puzzle (and as usual armed with pen and paper) will rapidly generate a well-understood, often repeated, form of TECS. An intermediate case, for many, might be when working on the popular and strangely satisfying Sudoku puzzles that are suddenly cropping up in newspapers all over the world. We could repeat this kind of exercise in filling out the details of many other examples of transient extended cognitive systems. A seasoned journalist, armed with a word-processor and a bunch of notes, may rapidly cycle through a whole range of TECSs, some one-off, others repeated; ditto for the mountaineer equipped with compass, map and altimeter, and so on.

Now there is no doubt that, in each specific case involving a TECS, the biological brain is (currently) an especially active player, recruiting some complex, non-linear combination of contributions from various types of on-board neural circuit and resource, and doing so in a way expressly tailored to accommodate and exploit the additional representational and computational potentials introduced by, for example, the compass, the pen and paper, or the word-processing package. These examples are, incidentally, the rough cognitive equivalents of the three main kinds of incidental task dynamics identified by Bingham—viz. those that simply introduce useful information or constraints (the compass), those that support offloading and returning of information (the pen and paper), and those that introduce new active sources of information processing and representation-transformation (the word-processing software). But despite this crucial role for the biological brain, there is much to be gained from the complementary study of the TECSs in their own right.

As before, the most obvious, and highly motivating, reason to do so is that it is these very ensembles that are locally at work in many of the most distinctive cases of human reasoning and problem solving. Here too, we alone on the planet seem capable of creating and exploiting such a wide variety of cognition-amplifiers, ranging from maps and compasses, to pen and paper, to software packages and digital music laboratories. But once again, a second and perhaps less obvious motivation is that working backwards from the analysis of these complex wholes may itself contribute important insights concerning the contributions and functioning of the biological brain itself. For here too the various internal neural contributions interact in a complex, non-linear fashion, and here too we may hope to gain valuable leverage on this forbidding internal complexity by analyzing cases in which some of the many degrees of freedom are deliberately (and profitably, relative to some specific goal) reduced by the use of external props and aids. For example, work on diagrammatic reasoning is beginning to track the various ways in which different kinds of diagram impose constraints on reasoning and action that (when the diagram is effective) echo those of some target domain (Stenning and Oberlander 1995; see also Zhang and Norman 1994).

The more general project known as ”external cognition” explicitly aims to track and understand the complex and often unobvious relationship between internal and external forms of representation. Its guiding idea is that "cognition is a function of both the internal and the external" (Scaife and Rogers 1996), and its practice involves seeking to understand the different properties of the internal and external structures and the way they fit together in specific problem-solving contexts. Sustained attention to the properties of, and varieties of, TECSs may thus yield a good deal of indirect information concerning what the biological sub-systems are and are not good at, and the forms of representation and computation upon which they most likely depend. One key advantage is, of course, that in the case of the external props and aids themselves, we are able (as Hutchins 1995 nicely notes) directly to inspect the various forms of representation, and directly to observe many of the key information–processing moves and representational transformations.

Much more contentious-sounding to some than the notion of a TECS is the notion (Clark and Chalmers 1998) of an extended mind. Yet in terms of our two dimensional taxonomy, the notion of an extended mind is nothing more than the notion of a cognitive extension, of any one of our three kinds, that scores rather more highly on the second dimension of durability and reliability. The extended mind idea thus simply takes the kinds of observation that already motivate interest in TECSs, and asks what would happen if some such organization were made more permanently available. Thus Clark and Chalmers (1998) imagined an agent, Otto, so thoroughly fluent in the use of a relatively permanent cognitive augmentation—a notebook, containing addresses and other such information—that the resource was deployed without conscious thought or intention, its operation and contents typically trusted, and the information it made available poised to impact conscious reason and deliberate action in very broadly the same way as might the same information were it stored in biological memory. In such a case, they argued, we should treat the non-biological augmentation as part of the material supervenience base for some of Otto's long-term, non-occurrent, dispositional beliefs (for example, about the location of an art gallery). The notebook and the physical traces therein should be treated, that is to say, as the physical vehicles of some of Otto's own non-conscious mental states.

Or consider, very briefly, another example that one of us has discussed previously in presenting the idea of extended cognition (Wilson 2004a: ch.8). Kanzi is a human-raised bonobo (pygmy chimpanzee) who has been thoroughly embedded from an early age in human-centered environments located in the research laboratories and grounds of Sue Savage-Rumbaugh (see Savage-Rumbaugh and Lewin 1994, Savage-Rumbaugh, Shanker, and Taylor 1998). As part of that enculturation process, Kanzi has learned how to use a 256 symbol, portable keyboard to communicate with people around him. Kanzi's actual developmental trajectory has taken him from using a technological resource designed by human agents, initially in temporary interactions then in repeated cycles of interaction, to becoming a distinctive agent whose persisting cognitive system has come to functionally integrate a much richer set of socio-cultural resources into its purview. The system that Kanzi plus his keyboard constitutes forms a cognitive system with memory and other cognitive capacities that seem qualitatively distinctive from that of other, unaugmented, bonobos: capacities that are somewhere between those of humans and other apes. It is not simply that Kanzi's enriched learning environment has restructured his neural wiring (although it has almost certainly done that too), but that his cognitive restructuring has proceeded through a potent cognitive extension involving these stable symbolic structures in his environment. Otto and Keyboard-Kanzi are thus both cases where a relatively enduring augmentation, suggesting deep functional integration, plausibly results in a cognitively reconfigured agent, an agent with an extended mind.

There has been much recent debate over such radical-sounding claims, and we do not plan to repeat very much of it here (see Clark, this volume, for some discussion). Instead, we simply note that the step from TECSs to the extended mind is not really as large as it may initially appear, especially once one recognizes the many grades along the continuum from the fleeting to the permanent, and further articulates the trichotomy between one-off, repeated, and permanent relationships uniting individuals and cognitive resources, as introduced above. As we see it, an extended mind is what you get, given the more basic acceptance of the possibility of temporary, soft-assembled extended cognitive systems, if and when certain additional coupling conditions are met (for much further discussion, see Clark in press-c). Such coupling conditions are meant to ensure that the capacities of the hybrid system—the biological organism plus augmentation—are plausibly seen as the capacities of a specific individual (e.g. Otto). For we properly expect our individual agents to be mobile, more-or-less reliable, bundles of stored knowledge and computational, emotional, and inferential capacities, and so we need to be persuaded that the new capacities enabled by the addition of the notebook are likewise sufficiently robust and enduring as to contribute to the persisting cognitive profile we identify as Otto-the-agent. The bulk of Clark and Chalmers (1998) was an attempt to isolate and defend a specific account of the conditions under which we would be justified in identifying such an extended mind. These amounted, in the end, to a set of conditions that (i) established the reliable presence of the new capacities as and when needed, and then (ii) made sure that the mode of deployment of the resource (automatic, trusted) made it more like a proper part of the agent and less like a perceptually consulted independent oracle.

These conditions turned out to be fairly stringent, and it is unlikely that any actual notebook currently carried by a human agent will meet the demands. In the context of near-future technologies, however, it may be that reliable, more permanent, forms of personal cognitive augmentation will become relatively commonplace. Two interlocking key developments (see Clark 2003) likely to support such a transition are, first, the increasing use of portable (perhaps, though not necessarily, implanted) electronics, and second, the spread of ubiquitous and pervasive computing, infusing much of the routinely available material world with accessible information and added computational potential.

The point of choosing the simple, technologically unsophisticated, notebook, however, was both to dramatize the <u>importance</u> of the reliability and coupling conditions, and to highlight the relative <u>unimportance</u> of the intrinsic nature of the resource itself. For a simple notebook, plainly, is quite unlike biological memory in its representational format, computational activity, and (for whatever this is worth) material structure. Indeed, as critics seldom tire of pointing out, such differences obtain between most of our external props and aids and the inner biological engine whose cognitive capacities they augment in significant ways. But such disparities, far from being a problem, are (we will now argue) the source of much of the power and interest of cognitive extension.

5. Are Some Resources Intrinsically Non-Cognitive?

We said at the outset that our articulation of the very idea of cognitive extension would help to reveal shortcomings with some of the most prominent objections to the view of situated cognition as extended cognition.

In this section and the next we concentrate on two recent critiques that we regard as posing some of the most challenging objections to this view of situated cognition, those of Fred Adams and Ken Aizawa (2001, in press, this volume) and Robert Rupert (2004, in press). In both cases, we aim to respond to their critiques in part by pointing to the diverse forms that extended cognition can take, and in part (and perhaps more interestingly) by uncovering some deeper assumptions on which their critiques turn: assumptions that we think are thrown into doubt by the more complex framework we here invoke.

Adams and Aizawa have argued for what they term a "contingent intracranialism" according to which, as a matter of empirical fact, earthbound cognitive processes are currently (at least as far as we know) restricted to the heads—better, the neural circuitry and central nervous system—of biological organisms. They thus reject the extended mind thesis as described above, and assert instead that nothing that is properly speaking cognitive goes on outside the bounds of skin and skull. As one of us (Clark this volume, in press-c) has addressed these issues at some length, we shall not repeat the exercise here, except to draw attention to what seems to be a crucial underlying belief or dogma without which the argument would not go through. We shall call this the <u>Dogma of Intrinsic Unsuitability</u>. It goes like this:

> Dogma of Intrinsic Unsuitability
>
> Certain kinds of encoding or processing are intrinsically unsuitable to act as parts of the material/computational substrate of any genuinely cognitive state or process.

In Adams and Aizawa (2001) the Dogma emerges as the claim that certain neural states, and no extra-neural ones, exhibit "intrinsic intentionality", conjoined with the assertion that no proper part of a truly cognitive process can trade solely in representations lacking such intrinsic content, such as the conventionally couched encodings in Otto's notebook (see Wilson in press on intentionality and extended cognition). The upshot of this is that the notebook (or Kanzi's keyboard) is deemed unsuitable as an element in any putatively cognitive process (see e.g. Adams and Aizawa 2001, p. 53). The Dogma is also at work in their later suggestion that cognitive psychology, in discovering pervasive features of human biological systems of memory and perception, is discovering features that may be the signatures of the kinds of causal process essentially required to support cognition. As a result, the absence of these signatures in the case of certain augmentations and add-ons is presented as a reason to doubt that the augmentations and add-ons contribute to cognitive processing properly understood (see e.g. Adams and Aizawa 2001, pp.52, 61).

The biggest flaw in both of these arguments concerns the relations between putatively essential properties of a whole and essential properties of the parts. For even if we grant that every cognitive system needs to trade in intrinsic contents, it does not follow that each part of every such system need do so (see Clark in press-c). Something can be part of the supervenience base of something's having property P, without itself having property P. To take a simple parallel suggested by David Chalmers, someone can be a leader by virtue of their relation to other people, without those other people being leaders or having "intrinsic leadership".

The Dogma of Intrinsic Unsuitability is just that: a dogma. Moreover, it is one that is, we suggest, ultimately in some tension with a robust faith in one of the central tenets of computationally-inspired cognitive science, viz, the idea that pretty much anything, including any kind of processing or encoding, <u>if it is properly located in some ongoing web of computational activity</u>, can contribute to the nature and unfolding of the cognitively relevant computations, and hence emerge as a proper part of the material substrate of some target cognitive process or activity. Call this the <u>Tenet of Computational Promiscuity</u>. Given that we have defended the idea that cognition can be extended because the computations it involves are extended, this tenet leads directly to the view of situated cognition as extended cognition that we have been defending. When Computational Promiscuity meets Intrinsic Unsuitability, something surely has to give. We think what has to give is pretty clearly the notion of Intrinsic Unsuitability.

The pressure on the Dogma of Intrinsic Unsuitability here, however, does not stem solely from accepting the idea of extended computation. This is because the Tenet of Computational Promiscuity is an instance of the broader functionalist insight that causal or functional networks of certain kinds are what is crucial to cognitive capacities, rather than anything about the particular stuff in which those networks are realized. To be sure, we should require that functional networks provide more than a shallow behavioral mimicry of indisputably cognitive creatures. But the kind of view that we have developed here, which begins with such indisputably cognitive creatures and then argues that their cognitive systems are, in fact, extended, avoids this kind of problem at the outset.

Computational Promiscuity is, it is important to notice, fully compatible with the thought that certain kinds of computational structure may be <u>necessary</u> for fluid, real-world intelligence. It is even compatible with the thought that such necessary computational structures (if such there be) may all be located (at present at least) inside the heads of biological agents. It asserts only that <u>once any such necessary structuring conditions have been met</u>, there is then no limit on the kinds of additional resource that may then be co-opted as proper parts of an extended cognitive process.

Once any such core systems are in place, many other kinds of representational and computational resource may come to act either temporarily or permanently as proper parts of more complex, hybrid, distributed, cognitive wholes. In such cases, it is often the very fact that these additional elements trade in modes of representation and processing that are <u>different</u> from those of the cognitive core that often makes the hybrid organization worthwhile. Tracing and understanding such deep complementarity is, we claim, the single most important task confronting the study of situated cognition. The example of Kanzi mentioned earlier is a case in point. There can be little doubt that, were it not for a wealth of pattern-recognizing know-how, Kanzi would not have been able to learn to use and deploy the symbol board. Yet there can also be little doubt that keyboard–Kanzi—the larger cognitive whole that results from the fluent coupling between bio-Kanzi and the new resource—is a fundamentally different kind of cognizing entity from her unaugmented cousins. Keyboard-Kanzi is not simply a cognitive core with an add-on symbol board. She is a new kind of hybrid thinking device.

6. Is cognition extended or only embedded?

Perhaps, though, we can carry out this project without buying into the idea of literal cognitive extensions. Perhaps it is enough, indeed, to speak merely of complementary processes, some cognitive, some not. Robert Rupert (2004, in press) offers just such a challenge for the kinds of approach we have been describing. The challenge comes in two parts. The first concerns what Rupert sees as the severe costs of seriously adopting an extended cognitive systems perspective. The second concerns what Rupert depicts as the lack of added value provided by the adoption of an extended perspective. To offset the severe costs, he argues, the added value would have to be very great indeed. But in fact, the combined effect, he fears, is just the opposite. The large costs are offset by no correspondingly large gains, and so the project of studying 'extended cognition' is one that both philosophy and cognitive science would be wise to reject.

Rupert distinguishes two projects that he sees as competing proposals for understanding situated cognition. The first is the one we defend here. It embraces a vision of cognitive processing itself as (sometimes) quite literally extending into the extra–organismic environment. Rupert dubs this the Hypothesis of Extended Cognition (HEC) and depicts it as a radical hypothesis apt, if true, to transform cognitive scientific theory and practice and to impact our conceptions of agency and persons. But it needs to be assessed, he argues, alongside a more conservative (though still interesting and important) competitor perspective. This is the perspective dubbed the Hypothesis of Embedded Cognition (HEMC) according to which:

> Cognitive processes depend very heavily, in hitherto unexpected ways, on organismically external props and devices and on the structure of the external environment in which cognition takes place. (Rupert 2004, p. 393)

Why prefer HEMC over HEC? Rupert starts with an appeal to commonsense. Commonsense, he suggests, rebels at the vision of extended cognition, so we need sound theoretical reasons to endorse it. HEMC, by contrast, is much more compatible with common sense.

Two main worries are then raised for HEC. The first worry, similar to one raised by Adams and Aizawa, concerns the profound differences that appear to distinguish the inner and outer contributions. Thus, for example, we read that "the external portions of extended 'memory' states (processes) differ so greatly from internal memories (the process of remembering) that they should be treated as distinct kinds" (2004, p. 407). Given these differences, there is no immediate pressure to conceive the internal and the external contribution in the same terms. But worse still, there is (allegedly) a significant cost. The cost is one that appears briefly in Rupert (2004), and is greatly expanded in Rupert (in press). For taking all kinds of external props and aids as proper parts of human cognitive processing robs us, he fears, of the standard object of cognitive scientific theorizing, viz, the stable persisting individual. Even in cases of developmental theorizing, where what is at issue is not so much stability as change, Rupert argues, one still needs to find an identifiable, though developing, core. Treating the temporary coupled wholes comprising organism and props as our target cognitive systems is thus a recipe for chaos since:

> The radical approach offers developmental psychologists no more reason to be interested in, for example, the series of temporal segments we normally associate with Sally from ages two-to-six rather than to be interested in, say, Sally, aged two, together with a ball she was bouncing on some particular day, Johnny, aged five, together with the book he was reading on some particular afternoon, and Terry, aged seven, plus the stimulus item he has just been shown by the experimenter. (Rupert in press, ms p.9)

More generally then, Rupert worries that cognitive science and cognitive psychology would lose their grip on their subject matter, and with it whatever progress has been made so far, were they to identify human cognitive processing with the activity of these "short-lived coupled systems" (in press, ms p.7). Given this very high cost, and given that all the genuine insights of HEC, so Rupert claims, can be accommodated in the more conservative framework of HEMC, there can be no compelling reason to adopt HEC.

These are good questions to raise, and we find Rupert a thoughtful and engaging critic. Nonetheless, Rupert's worries, including the very idea of a stark, all-or-nothing, contrast between HEC and HEMC, are misplaced, and for two quite deep reasons.

The first is that no part of the arguments for extended cognition turn on, or otherwise require, the similarity of the inner and outer contributions. Part of the confusion hereabouts may be due to a persistent misreading of what is sometimes known as the Parity Claim (originally introduced, though not thus labeled, in Clark and Chalmers 1998). This was the claim that if, as we confront some task, a part of the world functions as a process which, were it to go on in the head, we would have no hesitation in accepting as part of the cognitive process, then that part of the world is (for that time) part of the cognitive process. But far from requiring any deep similarity between inner and outer processes, the parity claim was specifically meant to undermine any tendency to think that the shape of the (present day, human) inner processes sets some bar (as, for example, Adams and Aizawa 2001 suggest) on what ought to count as part of a genuinely cognitive process. It is meant instead to act as a kind of veil of metabolic ignorance, inviting us to ask what our attitude would be if currently external means of storage and transformation were, contrary to the presumed facts, found in biology. (This, by the way, is very much how Turner approaches the idea of an "extended physiology", which we discussed in section 3.) The principle thus appeals to our rough sense of what we might intuitively judge to belong to the domain of cognition (rather than, say, that of digestion), but attempts to do so without the pervasive distractions of skin and skull.

Contrary to any requirement of similarity then, what the friends of extended cognition actually expect, and (as we saw in section 4 above) study, are hybrid processes in which the inner and the outer contributions are typically highly distinct in nature, yet deeply integrated and complementary. As an aside, this complementarity is probably most evident if your vision of the inner realm is a fundamentally connectionist one, as the stability, compactness, arbitrariness and recombinability of public language symbols and encodings contrasts quite dramatically with the fluid, distributed, non-classically compositional representations developed by a connectionist or neo-connectionist engine.

A second reason to resist the easy assimilation of HEC into HEMC concerns the nature of the interactions between the internal and the external resources themselves. Such interactions, it is important to notice, may be highly complex, nested, and non-linear. As a result there may be no viable means of understanding the behaviour and potential of the extended cognitive ensembles by piecemeal decomposition and additive reassembly. To understand the integrated operation of the extended thinking system created, for example, by combining pen, paper, graphics programs, and a trained mathematical brain, it may be quite insufficient to attempt to understand and then combine (!) the properties of pens, graphics programs, paper, and brains. This may be insufficient for just the same kinds of reason advanced by Bingham in the case of the human action system, or, within neuroscience itself, as reasons to study not just the various major neural sub-structures and their capacities, but also the complex (often transient) larger-scale activities in which they combine. The larger explanatory targets here are whole processing cycles, involving soft-assembled coalitions of neural resources recruited for some specific problem-solving purpose. Such soft-assembled neural packages involve the temporally evolving, often highly re-entrant, activity of multiple populations of neurons spanning a variety of brain areas. Why then suppose that the soft-assemblies most relevant to human cognitive achievements are essentially bounded by skin and skull? Why shouldn't the process of recruitment, and the skills of dovetailing the various contributions, yield, at least in our artifact-rich world, a succession of similarly complex hybrid ensembles spanning brain, body and world?

What, finally, of the allegedly intolerable costs of such an enlarged perspective? In one sense, the worry is simply misplaced, and results from a failure to appreciate the two independent dimensions that jointly construct the space of cognitive extensions. With this framework in mind, we see that there is no need, in taking cognitive extensions perfectly seriously, to lose our grip on the more-or-less stable, more-or-less persisting biological bundle that lies at the heart of each episode of soft-assembly leading to a TECS. Occasionally, of course, we may confront genuine (permanent, reliable) extensions of that more-or-less persisting core. Otto's notebook and Kanzi's symbol board are, we think, gestures at examples of this kind: cases where the persisting, mobile resource bundle is augmented in a robustly reliable manner. But in most other cases, we confront the cognitive equivalent of Bingham's Task Specific Devices: soft-assembled, temporary medleys of information-processing resources comprising a dovetailed subset of neural activity and environmentally routed augmentations. The costs of not accepting HEC are thus great indeed. For as cognitive extensions these are, quite literally, the soft-assembled circuitry of a great deal of practical human thought and reason. We ignore or downplay the importance of these ensembles, treating them as merely ersatz cognitive circuitry, at our theoretical peril. For the bulk of real-world problem solving, especially of the kinds apparently unique to our species, may be nothing but the play of representation and computation across these spectacularly transformative mixes of organismic and extra-organismic resources.

Overall, Rupert's strategy of argument rests on the claim that any benefits accruing to the expanded perspective can be as easily accommodated by the more conservative reading according to which all the cognizing goes on in the biological elements, with the rest just a temporarily recruited set of input devices, props and supports. In this respect, it is similar to the claims of behaviorists in the first half of the 20th-century that they could account for all so-called "cognitive phenomena" solely in terms of behavior. Similarly, Griffiths and Scarantino (this volume) think the debate concerning the extended mind merely "semantic" and effectively opt for HEMC on grounds of minimal disruption.

But we should treat such conservative claims with great caution. Consider the following (admittedly slightly brutal) parody.

> Look, there's all this new exciting talk about how the brain is the causal basis for cognitive processing. Call this the Hypothesis of In-brain Cognition (HIC). Poppycock. For there is a more conservative hypothesis available, the Hypothesis of In-neuron Cognition (HINC): for any particular cognitive ability, there is a given neuron, N, that is the real causal basis for that ability. Cognitive processes depend very heavily, in hitherto unexpected ways, on the rest of the brain, but it is only a given individual neuron that is ever genuinely cognitive. Any useful accounts you may develop using HIC can, in any case, be fully accommodated by HINC, and HINC is significantly less radical than HIC in that it requires only that we take into account how N exploits information coming from the rest of the brain and how N, in turn, transmits other information to yet other parts of the brain. Hence, HINC is to be preferred to HIC.

Perhaps there really is no reason to prefer HINC to HIC? Might it simply be that, for largely historical reasons, we have developed research traditions that accept something like HIC rather than HINC? History is no doubt part of the story here, but we trust it is not the whole of it. Rather, it is that our best empirical research tells us that intuitively cognitive acts often involve lots of neurons spread throughout the brain. HIC is then a sort of shorthand that signals this. In fact, our best research has helped us to identify not “the brain” but specific (sometimes temporarily assembled) complexes of neural systems as the causal basis for particular cognitive acts and capacities..

If this kind of substantive justification (of our actual preference for HIC over HINC) is at all on track, then Rupert’s claim that HEMC is preferable to HEC should seem suspect. In our view, there is much research that already fruitfully explores extended cognitive systems, and we have provided at least a preliminary sense of its diversity in the preceding sections. Perhaps it will always be possible to offer a deflationary reading of the ontological commitments of that research along the lines suggested by HEMC, just as it may always be possible to offer a similarly more ontologically frugal reading of current research along the lines suggested by HINC. But the real question is not whether we can always translate between these frameworks but where it is that we find functionally integrated systems that allow their bearers to perform cognitive tasks. We think that some of these are found solely in the head, and that some of them cross that boundary and incorporate cognitive resources in an individual’s environment. That's nature's way.

7. Conclusion: Letting Nature Take Its Course

Human agents exhibit both a metabolic and a cognitive organization. But whereas the former depends heavily on expensively maintained and policed organismic boundaries, the latter looks prone to repeated bouts of seepage in which cognitive processes productively loop through surrounding environmental structures. These structures may be natural, socio-cultural, or technological, or any combination thereof. And the resulting wholes may be one-off, repeated, or relatively permanent. This two-dimensional matrix limns the space and structure of cognitive extensions. The study of situated cognition, we have argued, is the study of the many forms of cognitive extension that appear in this complex space. To study such systems is not perversely to focus on some strange mish-mash of the cognitive and the non-cognitive. It is not wantonly to pursue the mind into some place it does not belong. Rather, it is to corral cognition in its den: to track nature taking its cognitive course.

* Thanks to Andrew Wilson for the pointer to Bingham's work on Task-Specific Devices, and to Mark Rowlands, John Sutton, Michael Wheeler, and Susan Hurley for useful discussions on many of the topics. The authors can be contacted at rob.wilson@ualberta.ca and andy.clark@ed.ac.uk.

References

Adams, F. and K. Aizawa, 2001, "The Bounds of Cognition", Philosophical Psychology 14(1): 43-64.

Adams, F. and K. Aizawa, in press, "Defending the Bounds of Cognition", in R. Menary (ed.), The Extended Mind. Aldershot: Ashgate.

Adams, F. and K. Aizawa, this volume, "Why the Mind is Still in the Head", in M. Aydede and P. Robbins (eds.), Cambridge Handbook of Situated Cognition. New York: Cambridge University Press.

Ballard, D., M.M. Hayhoe, P.K. Pook, and R.P.N. Rao, 1997, "Deictic Codes for the Embodiment of Cognition", Behavioral and Brain Sciences 20: 723-767.

Bingham, G., 1988, "Task–specific devices and the Perceptual Bottleneck", Human Movement Science 7: 225-264.

Brooks, R.A., 1991a, "Intelligence Without Reason", Proceedings of the 1991 International Joint Conference on Artificial Intelligence, pp.569-595. Reprinted in R.A. Brooks Cambrian Intelligence: The Early History of the New AI. Cambridge, MA: MIT Press..

Brooks, R.A., 1991b, "Intelligence Without Representation", Artificial Intelligence 47:139-160. Modified version reprinted in J. Haugeland (ed.), Mind Design II: Philosophy, Psychology, Artificial Intelligence. Cambridge, MA: MIT Press, 1997.

Brooks, R.A., 1999, Cambrian Intelligence: The Early History of the New AI. Cambridge, MA: MIT Press.

Burge, T., 1979, "Individualism and the Mental", in P. French, T. Uehling Jr., and H. Wettstein (eds.), Midwest Studies in Philosophy, vol.4, Metaphysics. Minneapolis, MN: University of Minnesota Press.

Burge, T., 1986a, "Individualism and Psychology" , Philosophical Review 95:3-45.

Burge, T., 1986b, "Intellectual Norms and Foundations of Mind", Journal of Philosophy 83:697-720.

Clancey, W., this volume, "Psychological and Anthropological Antecedents to Situated Cognition", in M. Aydede and P. Robbins, (eds.), Cambridge Handbook of Situated Cognition. New York: Cambridge University Press.

Clark, A., 1989, Microcognition: Philosophy, Cognitive Science and Parallel Distributed Processing. Cambridge, MA: MIT Press.

Clark, A., 1997, Being There: Putting Brain, Body, and World Together Again. Cambridge, MA: MIT Press.

Clark, A., 2003, Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence. New York: Oxford University Press.
Clark, A., 2005, “Beyond The Flesh: Some Lessons from a Mole Cricket”. Artificial Life 11: 233-244.
Clark, A., in press-a, “Word, Niche and Super-Niche: How Language Makes Minds Matter More”, Theoria.
Clark, A., in press-b, “Re-inventing Ourselves: The Plasticity of Embodiment, Sensing, and Mind”, Journal Of Philosophy And Medicine.
Clark, A., in press-c, "Coupling, Constitution and the Cognitive Kind: A Reply to Adams and Aizawa”, in R., Menary (ed.), The Extended Mind. Aldershot: Ashgate.
Clark, A., this volume, "Active Externalism and the Extended Mind", in M. Aydede and P. Robbins (eds.), Cambridge Handbook of Situated Cognition. New York: Cambridge University Press.
Clark, A., and D. Chalmers, 1998, "The Extended Mind", Analysis 58:10-23.
Ebbs, G., 1998, Rule-Following and Realism. Cambridge, MA: Harvard University Press.
Egan, F., 1992, "Individualism, Computation, and Perceptual Content", Mind 101: 443-459.
Fodor, J.A., 1975, The Language of Thought. Cambridge, MA: Harvard University Press.
Fodor, J.A., 1982, “Cognitive Science and the Twin-Earth Problem”, Notre Dame Journal of Formal Logic 23:98-118.
Fodor, J.A., 1983, The Modularity of Mind. Cambridge, MA: MIT Press.
Fowler, C., and M.T. Turvey, 1978, "Skill Acquisition: An Event Approach with Special Reference to Searching for the Optimum of a Function of Several Variables", in G. Stelmach (ed.), Information Processing in Motor Control and Learning. New York: Academic Press.
Gallagher, S., this volume, “Philosophical Antecedents to Situated Cognition”, in M. Aydede and P. Robbins (eds.), Cambridge Handbook of Situated Cognition. New York: Cambridge University Press.
Goody, J., 1977, The Domestication of the Savage Mind. Cambridge, England: Cambridge University Press.
Griffiths, P.E. and A. Scarantino, this volume, “Emotions in the Wild: A Situated Perspective on Emotion”, in M. Aydede and P. Robbins (eds.), Cambridge Handbook of Situated Cognition. New York: Cambridge University Press.

Haugeland, J., 1998, "Mind Embodied and Embedded" in J. Haugeland Having Thought: Essays in the Metaphysics of Mind. Cambridge, MA: Harvard University Press. (Originally appeared *in* Acta Philosophica Fennica 58 (1995) pp.233-267.

Hutchins, E., 1995, Cognition in the Wild. Cambridge, MA: MIT Press.

Hurley, S., 1998, Consciousness in Action Cambridge, MA: Harvard University Press.

Jablonka, E., and M. Lamb, 2005, Evolution in Four Dimensions. Cambridge, MA: MIT Press.

Loar, B., 1988, "Social Content and Psychological Content", in R. Grimm and D. Merrill (eds.), Contents of Thought. Tucson: University of Arizona Press.

Newell, A., 1980, "Physical Symbol Systems", Cognitive Science 4: 135-183.

Newell, A., and H., Simon, 1976, "Computer Science as Empirical Enquiry" in J. Haugeland (ed.), Mind Design. Cambridge, MA: MIT Press

Norman, D., 1993, "Cognition in the head and in the world" Cognitive Science 17 (1): 1-6.

Odling-Smee J., K. Laland, and M. Feldman, 2003, Niche Construction: The Neglected Process in Evolution. Princeton, NJ: Princeton University Press.

Oyama, S., 1985, The Ontogeny of Information. 2nd ed., Durham, NC: Duke University Press, 2000.

Oyama, S., P.E. Griffiths, and R.D. Gray (eds.), 2001, Cycles of Contingency: Developmental Systems and Evolution. Cambridge, MA: MIT Press.

Pettit, P., 1993, The Common Mind: An Essay on Psychology, Society, and Politics. New York: Oxford University Press.

Putnam, H., 1975, "The Meaning of 'Meaning'", in K. Gunderson (ed.), Language, Mind and Knowledge. Minneapolis, MN: University of Minnesota Press. Reprinted in H. Putnam, Mind, Language, and Reality: Philosophical Papers vol.2. New York: Cambridge University Press.

Pylyshyn, Z., 1984, Computation and Cognition. Cambridge, MA: MIT Press.

Rogers, Y., and M. Scaife, 1997, "Distributed Cognition", available at http://www-sv.cict.fr/cotcos/pjs/TheoreticalApproaches/DistributedCog/DistCognitionpaperRogers.htm).

Rowlands, M., 1999, <u>The Body in Mind</u>. New York: Cambridge University Press.

Rupert, R. 2004. “Challenges to the Hypothesis of Extended Cognition,” <u>Journal of Philosophy</u> 101: 389-428.

Rupert, R., in press, "Representation in Extended Cognitive Systems: Does the Scaffolding of Language Extend the Mind?" in R., Menary (ed.), <u>The Extended Mind</u>. Aldershot: Ashgate.

Salzman, E and J.A.S. Kelso, 1987, "Skilled Actions: A Task Dynamic Approach", <u>Psychological Review</u> 94: 84-106.

Savage-Rumbaugh, S., and R. Lewin, 1994, <u>Kanzi: The Ape at the Brink of the Human Mind</u>. New York: Wiley and Sons.

Savage-Rumbaugh, S., S.G. Shanker, and T.J. Taylor, 1998, <u>Apes, Language, and the Human Mind</u>. New York: Oxford University Press.

Scaife, M. and Y. Rogers, 1996, "External Cognition: How Do Graphical Representations Work?" <u>International Journal of Human-Computer Studies</u> 45: 185-213.

Segal, G., 1989, “Seeing What is Not There”, <u>Philosophical Review</u> 98: 189-214.

Shapiro, L., 1997, “A Clearer Vision”, <u>Philosophy of Science</u> 64: 131-153.

Stenning, K. and J. Oberlander, 1995 "A Cognitive Theory of Graphical and Linguistic Reasoning: Logic and Implementation", <u>Cognitive Science</u> 19: 97-140.

Turner, J. Scott, 2000, <u>The Extended Organism: The Physiology of Animal-Built Structures</u>. Cambridge, MA: Harvard University Press.

Wilson, R.A., 1994, “Wide Computationalism”, <u>Mind</u> 101: 351-372

Wilson, R.A., 1995, <u>Cartesian Psychology and Physical Minds: Individualism and the Sciences of the Mind</u>. New York: Cambridge University Press.

Wilson, R.A., 2000, "The Mind Beyond Itself", in D. Sperber (ed.), <u>Misrepresentations: A Multidisciplinary Perspective</u>. New York: Oxford University Press, pp.31-52.

Wilson, R.A., 2003, “Individualism” in S. Stich and T.A. Warfield (eds.), <u>The Blackwell Companion to Philosophy of Mind</u>. New York: Blackwell.

Wilson, R.A., 2004a, <u>Boundaries of the Mind: The Individual in the Fragile Sciences: Cognition</u>. New York: Cambridge University Press.

Wilson, R.A., 2004b, “Recent Work in Individualism in the Social, Behavioural and Biological Sciences”, <u>Biology and Philosophy</u> 19: 397-423.

Wilson, R.A., 2005, Genes and the Agents of Life: The Individual in the Fragile Sciences: Biology. New York: Cambridge University Press.

Wilson, R.A., in press, “Meaning Making and the Mind of the Externalist”, in Richard Menary (editor), The Extended Mind. Aldershot: Ashgate.

Zhang, J. and D.A., Norman, 1994, "Representations in Distributed Cognitive Tasks", Cognitive Science 18: 87-122