Why Think that Cognition is Distributed?

David Spurrett

A number of other papers in this issue of *Alternation* (Blair and Cowley, Dellis and Spurrett, Cowley) pursue their particular topics of interest within what is known fairly generally as 'distributed cognition'. The primary purpose of this short paper is to survey some of the case for distributed cognition, or to say why it is that we think it makes sense to regard cognitive processes as extending beyond the brain.

First, though, it will be worth spending a little time considering why it can seem so obvious that cognitive processes are either co-extensive with the central nervous system, or perhaps with some sub-section of it, i.e. that cognition takes place 'in the head'. After Cowley and Spurrett (forthcoming), we call this view 'cognitive internalism'. There is, after all, ample evidence for the dependence of cognition on the brain. Persons without brains don't do any thinking. Brains certainly *aren't* there for any of the exotic purposes such as cooling the blood proposed by some ancient authors—they are there to do something with the information transmitted via nerves to the sense organs, and to do *something* else to the nerves connecting the brain with muscles and other engines of bodily activity. It makes a lot of sense, at least *prima facie*, to see the brain as being a bounded system with definite inputs and outputs, and to think that what it does is, somehow, to solve a control problem—finding the best outputs given the inputs.

One large body of evidence, collected especially since the nineteenth century adds to this picture the idea that the brain is regionally specialised for particular functions, and another indicates that it has a particular kind of processing architecture. The first body of work painstakingly correlated variations in performance on particular cognitive problems with haphazardly collected instances of localised brain damage, resulting from either surgery or misadventure. Although the tides in the debates over localisation versus global processing moved and continue to move, it is clear that particular regions of the brain, whether or not they handle *all* of the work relevant to particular types of performance, are at least critically important for, *inter alia*, such functions as motor control, speech production, speech comprehension, visual processing, and face recognition[1].

The second body of work concerns the finer structure of the brain, and follows from the 'discovery' of neurones and the ways that they are interconnected. The existence of nerves had been established for millennia and the action of nerves, especially motor nerves, studied extensively from the sixteenth century onwards. It was only in 1873 (through the work of Camillo Golgi) that the apparently homogenous mass of the brain was shown, through a process of treating pieces of brain tissue so that microscopic examination could resolve individual neurones, to have a structure composed of many interconnected neurones, which appeared to operate in similar ways, at least, to nerve cells[2]. The discoveries here fed into and reinforced some of the work on localisation, since they provided a way of thinking about what was happening when some part of the brain was stimulated by an electrical pulse, and by extension when nerves ordinarily stimulated one another. They also enabled more serious thinking about the lower level processing architecture of the brain to begin, just because it was possible now to imagine the brain as being, in some sense, a vast collection of elements which were themselves something like switches.

At another level of abstraction entirely, a significant body of thought was already treating thinking as essentially computational. Central to this was Boole's approach to logic, which formalised propositional reasoning at least in a way apparently amenable to mechanical treatment, as a series of applications of syntactic rules to representational tokens. Boole himself called his 1853

[1] The founding work on localisation took place in the mid nineteenth century, by such pioneers as Paul Broca, Carl Wernicke, David Ferrier, Gustav Fritsch and Eduard Hitzig. Glynn (1999) includes an accessible survey with references to the primary literature.

[2] The conjecture that the brain was a network of some sort was much older than the first method of examining the network. In the 1630s Descartes (1985: 107) had written, in his *Treatise on Man*, that one would 'not be able to imagine anything more plausible than that [the brain] is composed of many tiny fibres variously interlaced ...'.

book *An Investigation of the Laws of Thought on Which are Founded the Mathematical Theories of Logic and Probabilities*, making quite clear that the serial and symbolic computational vision he articulated was precisely intended to capture what was significant about human thought in general.

Although the 'fit' between a roughly Boolean image of computational thinking, and a neurally inspired image of brain activity is not self-evident, the two lines of thinking naturally converged as a research programme, set on its course by founding contributions from, e.g., Turing and Von Neumann. Given the expectation that thought was computational, and the discovery that brains appeared to be some kind of computing device, the task seemed to be to find out how the brain did whatever it did, so that our own thinking turned out to be computational in the right kinds of ways. That it to say that it was almost obvious, although by no means necessary, to suppose, first, that the 'problem' of optimising outputs on the basis of inputs is a primarily computational problem, and second that what the brain is for, is performing those computations in a space between, and bounded by, perception and action[3]. This is pretty much what early, in the words of Haugeland (1985) 'Good Old Fashioned Artificial Intelligence' (GOFAI) did, viewing cognition as a 'computational process that takes and input and produces an output' (1997) and so Simon and Kaplan (1989) could quite fairly say that 'The computer was made in the image of the human'.

GOFAI has produced interesting and important results, and remains a significant research programme. Nonetheless there is a growing sense that real human and animal intelligence is something different entirely. Sketching the reasons for this is the main burden of the remainder of this paper.

The Great Escape

Cognition did not break out of the internalist dungeon all in one go. As is almost typically the case, a variety of voices had been raised against it, from a range of difference backgrounds, disciplinary perspectives and theoretical orientations. Prior to, or roughly contemporary with, the beginnings of artificial intelligence, one could mention Vygotsky, Heidegger, Wittgenstein,

[3] Two of the historical figures being given especially short shrift here are Babbage, with his work in the nineteenth century making calculating machines, and Turing (e.g. 1950), who in the mid twentieth century did a great deal to generalise the notion of a computational architecture.

Merleau-Ponty, Von Uexkull and others[4]. During the height of early GOFAI, there were Gibson, more Vygotsky, followers of Heidegger and others still[5]. All, in different ways and for different although often overlapping reasons saw fault with cognitive internalism, and sought in their work to describe ways in which cognition was bound up with bodily experience and/or the very 'external' environment which cognitive internalists took to be utterly separated from cognition. More than theory, though, the present survey is concerned with evidence. A number of particular empirical studies of different types of cognition show in a range of ways how cognitive processes latch onto, exploit and are intimately connected with the bodily, physical and social environments in which they take place, or that cognition is distributed and nto merely related to brain function.

There are, in the literature, two senses of 'distributed' which it will be worth distinguishing. The first, associated with neural networks and some other parallel processing architectures, indicates that cognition is distributed *within* the brain, or computational system. Neural networks, the most brain-like type form of parallel distributed processing, or PDP, are now a well established technology. Depending on the particular network, and the system mapping its inputs and outputs onto features of the world or some coding system, a neural network can implement a computational solution to a problem which does not involve internal representational symbols in the way expected by traditional artificial intelligence. Systems which are distributed in this way do not typically involve any form of central processing. Even a system which is highly distributed in this sense, though, need not be at all distributed in the second sense[6].

A system which is distributed in the second sense exploits resources besides those of the brain or any other obviously information-processing architecture (which means bodily or environmental resources, or both) for *cognitive purposes*. This is a more radical sense of distributed cognition, since it denies part of what can make it seem so natural and simplistic to think that all cognition is handled by the brain. This concerns the fact that neurones do

[4] See, e.g., Heidegger (1962), Vygotsky (1986); Merleau-Ponty (1962); Von Uexkull (1943); Wittgenstein (1953).

[5] See, e.g., Gibson (1979); Dreyfus (1979, 1991).

[6] Cilliers (1998) is a recent example of a work forcefully urging a view which is distributed in the first sense, yet remains strikingly cognitive internalist.

not connect up to the world, but that they do connect up to sensory and motor organs.

There is little doubt that cognition is distributed in the first sense. That commitment, of course, leaves open just what the actual computational architecture of the brain is, and in particular in what respects and to what degree it is modular, hence allowing for functional decomposition at levels intermediate between individual neurones and the brain as a whole. It is the case for cognition being distributed in the second sense which is of primary concern here.

One set of examples and evidence can be found with reference to locomotion. Thelen and Smith (1994: 8f) point out that traditional cognitive internalist approaches to locomotion had been encouraged by evidence indicating that locomotion in many vertebrates seemed to be controlled by a pattern of activity generated in the spinal cord. This pattern could be generated and co-ordinated locomotion on a treadmill ensue even in animals where the brain had been removed entirely. Furthermore the pattern could be shown to be produced even when the muscles it 'controlled' had been artificially paralysed. It was widely supposed that the spinal column, considered as a part of the brain, was home to a central pattern generator which was the driving structure for locomotive activity.

Let us leave this hypothesis aside for a moment, and look at some recent examples of effective locomotion in robots. The case of robots is illuminating precisely because it is with robots that the most striking examples of artificial or manufactured cases where a computer has access to bodily and environmental resources are to be found.

The hexapod robot built by Espenschied (see Quinn & Espenschied 1993, and also Nelson *et al.* 1997) has processing capabilities limited to a mere 37 artificial neurones, in an array inspired by biological precedents in hexapod invertebrates.[15] Each leg has its own set of neurones, generating motion rhythms. These could easily be mistaken for the hypothesised central pattern generator, except for the fact that these rhythms are in turn modulated by additional connections between the sets of neurones peculiar to each leg, which in turn allow individual legs to inhibit the activity of neighbouring legs. Further, each leg-specific set of neurones received some 'sensory' feedback from feature-detectors built into that leg, which tracked information about joint position and pressure between the foot and the surface beneath. The neural set for each leg controlled the swing (forward and backward) of the leg and, lifting and lowering of the foot.

The overall network of neurones is highly distributed in the first sense noted above, so there is no 'central executive' doing the processing, and, crucially, no stored plan. Incorporating tendencies to some key types of motion, also modelled on biological understanding of insect locomotion, has the result that the robot engages in 'searching' behaviour with individual feet, which enables a single foot to find a secure footing before another is 'permitted' (by the cessation of inhibitory signals from the first leg) to do the same, and 'climbing' behaviour which enables it to navigate obstacles which block its path. The resultant locomotion is capable of dealing with varying terrain and, importantly, exhibits what is known as 'graceful degradation' which is to say that limited and selective 'damage' to isolated parts of the system (i.e. a few neurones, sensory feedback paths, or part of the physical structure of the robot) has relatively little effect on the quality of overall performance. This flexibility and efficiency is achieved by means of exploiting bodily and environmental resources, instead of, for example, time consuming and computationally expensive visual perception followed by modelling of the environment in three dimensions, with the robot placed in that model at a specified position and with a particular orientation, as a method of determining where and how to place the feet.

So even if there is something which works somewhat like a central pattern generator in these cases, it is clearly not doing all of the work. Rather a coalition of resources including non-neural ones *together* account for flexible and responsive locomotion. It turns out to be just the same in vertebrates, as Thelen and Smith found. They argue that there is no evidence for the view that there is some centralised cognitive 'plan' for locomotion, and plenty of evidence against it. They show that varying non-neural factors (by placing the feet of a walking infant on a treadmill where the left foot moves at a rate different to the right, by increasing the effective mass of a limb by means of weights, or by decreasing it by partly placing the infant in a buoyant medium, etc.) significant variation in the patterns of locomotion result. Actual performance, as with the robot case, depends not on the activity of the spinal column alone, nor on any purely neural collection of effector processes, but rather on a coalition of elements only some of them neural. As a result, they maintain that locomotion is modular in nature, depending on a variety of resources, and where successful performance is an emergent phenomenon arising from the interaction of the various roughly modular elements.

In case the example of locomotion does not seem 'cognitive' enough, one may consider the behaviour of human players of the computer game

Tetris[7]. A classical cognitive internalist account of Tetris playing would separate perception of the current situation of the playing area, and the most recently appearing 'zoid'[8] from, on the one hand, computation directed to working out where best to place the new zoid, and on the other, building a motor plan for placement on the basis of this computation. A full account of such an approach, based on the planning literature, is offered by Kirsh and Maglio (1994), and separates processing into four phases:

1 Create an early, bitmap of representation of selected features of the current situation.
2 Encode the bitmap representation in a more compact, chunked, symbolic representation[9].
3 Compute the best place to put the zoid.
4 Compute the trajectory of moves to achieve the goal placement.

In their research on Tetris playing, Kirsh and Maglio (e.g. 1994; 1997) used this model as a source of comparisons with the actual measured behaviour of human players. The more human players handle the cognitive demands of the game in line with classical expectations, the more the data should fit that model. The more difficult it is to square the data with the model, the more likely it is that humans are doing something else. It is important to note that in the standard planning model computation and action are radically separated: the placement problem is solved *first* as an internal computation, and only

[7] Tetris is a game involving the rotation and translation of 'falling' shapes (moving from the top to the bottom of the playing area) in order to achieve optimal fit at the bottom of the playing area. Completely filled horizontal rows are emptied on completion, with any incomplete rows above them moving down into the spaces thus created. The game ends when the accumulation of incomplete rows caused by player error prevents any new pieces from entering at the top of the playing area. Additional reference to Kirsh and Maglio on Tetris can be found in Blair and Cowley (this volume) and Dellis and Spurrett (this volume).
[8] 'Zoid' is the term Kirsh and Maglio use to refer to the individual 'pieces' or 'elements' which fall in the Tetris game.
[9] This would involve translating the raw image into a version that encoded edges, concave corners, convex corners, etc.

once that solution has been worked out is there any role at all for action in the world.

Kirsh and Maglio therefore noted that one consequence of the standard model is that, if it is basically correct, any actions taken *before* the completion of the internal computation would, of necessity, be unplanned. This would indicate that to the extent that human action approximates that prediction, there should be a delay between the appearance of a new zoid and the commencement of rotation and translation. They state, though, that this 'is patently not what we see in the data' (Kirsh & Maglio 1994). The first gross indication of this mismatch between the predictions of a classical planning model and the actions of real human players is that effective players engage in rotations almost immediately, and that the number of rotations and translations made were significantly in excess of what the internalist model predicted as being optimal.

On the basis of a series of carefully constructed additional experiments Kirsh and Maglio demonstrated convincingly that these 'surplus' actions *all* yielded computational advantages; that they facilitated earlier recognition of particular shapes, simplified the problem of detecting edge fit, reduced the number of vertical placement errors, and so forth. Kirsh and Maglio (1994) say the following in their conclusion:

> This way of thinking treats the agent as having a more cooperative and interactional relation with the world: the agent both adapts to the world as found, and changes the world, not just pragmatically, which is a first order change, but epistemically, so that the world becomes a place that is easier to adapt to. Consequently, we expect that a well-adapted agent ought to know how to strike a balance between external and internal computation. It ought to achieve an appropriate level of cooperation between internal organizing processes and external organizing processes so that, in the long run, less work is performed.

Kirsh and Maglio coined the term 'epistemic action' to describe these actions, which are paradigm examples of distributed cognition. They are to be distinguished from 'pragmatic' actions, where pragmatic actions are oriented towards achieving some physical goal. Epistemic actions on the other hand are real physical actions that are performed in order to modify the computational state of the system performing the action, in this case the person playing the game. It is worth pointing out that this reliance on the external world for cognitive support is more than a developmental phase, observed in beginner or intermediate players and less so or not at all in advanced ones. Further work by Maglio and Kirsh (1996) shows that the incidence of epistemic actions actually *increases* as skill advances. Epistemic action, or distributed cognition, is not a developmental stage, but an apparently essential property of increasingly optimal solutions. What this research shows, in one carefully selected area (although additional work by Maglio *et al.* (1999) indicates that the type of result found with Tetris is found with respect to Scrabble playing as well), is that human cognition relies heavily on non-neural resources, that is the manipulative capabilities of our bodies, and the structure observable in the environment, in order to solve essentially computational problems.

Part of what makes Kirsh and Maglio's research so compelling is that by using the game of Tetris, they were able to collect fine-grained data concerning human behaviour which could be evaluated in a rigorous manner, given the computational properties of the game. Good players are those who achieve high scores, and good zoid placements are ones which don't create gaps. This feature of the game meant that the specific patterns of epistemic action characteristic of good play could be, and were, discovered and further studied. Kirsh and Maglio could tell relatively easily who the good players were, and focus on making sense of what they did.

In a similar way, good navigation teams are ones whose ships don't go off course. The documented plan concerning where a ship is *supposed* to be at any given time permits evaluation of the practices which lead to its being either there, or somewhere else. In his extended study of the navigation problem, Hutchins (1995, see also Hutchins and Klausen 1996) brought the skills of a cognitive anthropologist to bear on a particular type of situation in which humans co-operate in the attempt to solve a computational problem. Hutchins (1995:12) glosses navigation as 'a collection of techniques for answering a small number of questions, perhaps the most central of which is "where am I?"' In the case of a large ship, answering this question typically involves fixing the position and orientation of the ship on a chart for a specific moment in time, or reconciling information that can be gathered from on board the ship with that represented on the chart.

The particular ship on which Hutchins gathered most of his data, like many others, could not possibly be navigated by a single person[10]. The physical dimensions of the ship, and the arrangement of the navigation area

[10] That is, except when the ship is in the open sea.

and observation positions for taking bearings on landmarks were such that a co-ordinated team was essential. Hutchins study focussed on the behaviour of the people taking the bearings, the people telling them what bearings to take, the people recording those bearings, plotting them on the chart, and integrating them with other sources of information concerning the speed of the ship, the depth of water beneath it and so forth. Contrary to what a cognitive internalist might have expected, that is a collection of cognitively self-contained agents sharing information about elements of the navigation problem, so that some central executive agent could make decisions, he found that the team as a whole functioned as a kind of extended cognitive agent.

Hutchins argued that not only material labour, but also cognitive labour, can be socially divided, and that the ways in which cognitive division of labour is possible depend partly on cultural and material conditions. The particular social conventions of military navigation that Hutchins studied enabled a collection of people and artefacts (including charts, measuring instruments, and other navigation tools, often used in ways which Kirsh and Maglio would describe in terms of epistemic action) to have properties paradigmatic of human built computer systems, including function-specific 'daemons' (1995: 1991) and protocols for 'buffering' data prior to processing (1995: 195). Two important points need making here. The first is that the type of computational architecture Hutchins identifies in a navigation team has some of the same general properties as the locomotion architecture discovered by Thelen and Smith, including forms of robustness and flexibility arising from what Thelen and Smith call 'soft assembly' (Hutchins 1995: 185, 223; Thelen and Smith 1994: 60) or the absence of fixed control hierarchies. The second point is that the 'level' at which the system being studied most closely approximates a classical computer is that of the distributed system including the world. This is a point which has been made in many different ways recently. For his part Hutchins quotes Dennett (1991: 212, in Hutchins 1995: 361) speculating about how Turing could have considered the formal architecture now known as a Turing machine to be similar to human thinking, by taking a certain view of the activity of a working mathematician as paradigmatic. Hutchins himself points out that a danger here is mistaking the cognitive properties of a larger system in which a mathematician manipulates symbols in some medium such as pencil and paper for the cognitive properties of the brain of the mathematician him or herself (see Rumelhart, Smolensky, McClelland, and Hinton 1986: 46). An encultured brain in a body, with access to cognitive prostheses like writing media and symbol systems has powers different to those of a 'naked' brain in virtue of those external resources. Hutchins' approach also enables some sense to be made of the ways in which cultural products such as measuring instruments and systems of notation can facilitate different types of computation, and permit different forms of division of cognitive labour. His examples concern navigation systems, but for the sake of brevity of explanation I will refer to number systems. The ways in which it is much easier for users of Arabic (or any other place-value system of numerals) to perform arithmetic than it is for the users of Roman (or any other system of numerals not using place-values) has nothing to do with the different properties of the *brains* of individuals from different ancient civilisations. Rather 'hooking up' with a particular external system of signs and rules confers particular advantages and disadvantages.

On the subject of signs, Dennett (e.g. 1991) and Clark (e.g. 1997), among others, both make a great deal of the ways in which language itself can be regarded as a kind cognitive prosthesis, permitting different and vastly more powerful forms of social co-ordination, and sharing of epistemological and other resources. In his analysis Clark's list of cognitive advantages includes the capacity for self-stimulation to improve control and performance at tasks (Clark 1997: 202), being able to use symbolic systems to augment memory, by using non-neural storage media (Clark 1997: 201), using labels and symbols to simplify our environments and learning processes (Clark 1997: 201), and simplifying various other types of problem solving (See also Blair and Cowley in this volume). Clark tends to emphasise the ways in which our native cognitive powers can dovetail with language to extend our cognitive abilities, while Dennett favours the view that in learning language we engage in a process which reprograms our brains, permitting them in some ways to approximate the functioning of a von Neumann machine. What they agree on, though, is that our cognitive powers *with* language, where language is seen as a largely external, public resource, are vastly different to those without.

Rather than accumulate further individual fragments of evidence[11], I want to draw this survey to a close with a sketch of an argument for the *likelihood* of distributed cognition. This argument is by no means a proof that distributed cognition is the only possible type, although a developed form of the argument could perhaps show that distributed solutions to cognitive

[11] There is, at present, no better general survey than Clark (1997) and most of the examples here are also discussed in that work.

problems were at least nomologically necessary in entities designed by natural selection.

In the realm of biology there is a counterpart to cognitive internalism, which one could call 'functional internalism'—the view that the activities of any organism will be performed entirely by 'on board' components. Consider, in this light, the sponge (following Vogel 1981). Sponges feed by filtering materials found in water. It has been known since 1864 that the force driving the unidirectional current internal to a sponge was flangellar. Vogel notes that:

> ... an important related question was not asked for another hundred years. If the flangella were inoperative, would water pass through a sponge anyway? Or, to put the matter in more realistic terms, does flangellar action account for all of the water passing through a sponge, or can ambient water currents make a contribution to filtration? It appears now ... that not only do ambient currents help, but that the structure of sponges is most exquisitely adapted to take advantage of such currents, with clear functions attaching to a number of previously functionless features. Dynamic pressure on the incurrent openings facing upstream, valves closing incurrent pores and downstream, and suction from the large distal or apical excurrent openings combine to gain advantage from even relatively slow currents .... Why did so much time elapse before someone made a crude model of a sponge, placed it in a current, and watched a stream of dye pass through it? (Vogel 1981:190).

The answer to Vogel's question is, as he notes, that biologists been wedded to what I have just called functional internalism. Vogel's project is to reject this dogma, and he proposes a principle which should be adopted to help guide us away from that view: 'Do not develop explanations requiring expenditure of metabolic energy until simple physical effects are ruled out' (Vogel 1981:182). Andy Clark, who refers to Vogel (in 1989 and 1997) draws a moral for cognition, which nods in the direction of a famous fictional opportunistic exploiter of resources he didn't bring with him, the '007 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. That is, know only as much as you need to know to get the job done (Clark 1989: 64).

The more general argument which it is possible to see here goes something like this: The process of natural selection is blind to the boundaries between brains and bodies, and between bodies and world. Not only that, the process of selection always takes place in a particular environment, without the particular scarcity and competition problems of which there would be no selection at all. If distributed solutions to problems in general, whether the water filtering challenge facing sponges or something more cognitive, are more efficient or effective, then they will be favoured. The examples above are all more efficient or effective than internalist counterparts. Whether the argument can be made into more than a sketch, though, would depend on showing that distributed solutions have to be more effective, or are at least considerably more likely to do so. In that case we should expect them, and be forced back to considering internalist options only when we are unable to make sense of some phenomenon in distributed terms. The papers in this collection by Blair and Cowley, Dellis and Spurrett, Cowley and Povall are not motivated by the notion that cognition *has* to be distributed, but they do all take it in different ways as a working hypothesis that it is more likely to be than not. If it isn't, then the investigation of human cognitive capabilities, and the project of maximising them, has far more to do with studying the cognitive properties of artefacts, environments, bodies and cultures than traditionally expected.

Department of Philosophy
University of Natal

References

Boole, G 1853. *An Investigation of the Laws of Thought on Which are Founded the Mathematical Theories of Logic and Probabilities*. London: Macmillan.

Blair, G & S Cowley (this volume). Language in Iterating Activity: Microcognition Re-membered.

Brooks, RA 1997. From Earwigs to Humans. *Robotics and Autonomous Systems* 20,2-4:291-304.

Cilliers, P 1998. *Complexity and Postmodernism: Understanding Complex Systems*, London: Routledge.

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

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

Cowley, S & D Spurrett, D (forthcoming). Putting Apes (Body and Language) Together Again. To appear in *Language Sciences*.

Dellis, A & D Spurrett (this volume). Real Patterns and Distributed Cognition.

Dennett, D 1991. *Consciousness Explained*. Boston: Little, Brown.

Descartes, R 1985. *The Philosophical Writings of Descartes*, Volume 1, Cottingham, JG, R Stoothoff & D Murdoch (trans & ed). Cambridge: Cambridge University Press.

Dreyfus, H 1979. *What Computers Can't Do*. New York: Harper and Row.

Dreyfus, H 1991. *Being-in-the-world: A Commentary on Heidegger's* Being and Time. *Division I*. Cambridge, Mass.: MIT Press.

Gibson, JJ 1979. *The Ecological Approach to Visual Perception*. New York: Houghton Mifflin.

Glynn, I 1999. *An Anatomy of Thought: The Origin and Machinery of Mind*. London: Phoenix.

Haugeland, J 1985. *Artificial Intelligence: the Very Idea*. Cambridge, Mass.: MIT Press.

Heidegger, M 1962. *Being and Time*. Macquarrie, John & Edward Robinson (trans). Oxford: Blackwell.

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

Hutchins, E & T Klausen 1996. Distributed Cognition in an Airline Cockpit. In Engestrom, Y & D Middleton (eds): *Cognition and Communication at Work*. New York: Cambridge University Press.

Kirsh, D & P Maglio 1994. 'On Distinguishing Epistemic from Pragmatic Action. *Cognitive Science* 18:513-549.

Kirsh, D & P Maglio 1997. Some Epistemic Benefits of Action: Tetris, a Case Study. *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society*. Hillsdale, New Jersey.

Merleau-Ponty, M 1962. *The Phenomenology of Perception*. London: Routledge and Kegan Paul.

Maglio, PP & D Kirsh 1996. Epistemic Action Increases with Skill. *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum.

Maglio, PP, T Matlock, D Raphaely, B Chernicky & D Kirsh 1999. Interactive Skill in Scrabble. In *Proceedings of Twenty-first Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum.

Nelson, GM, RD Quinn, A Bachmann, WC Flannigan, RE Ritzmann, JT Watson 1997. *Design and Simulation of a Cockroach-like Hexapod Robot*. Proceedings of the 1997 IEEE International Conference on Robotics and Automation (ICRA '97), Albuquerque, NM, April 22-24, 1997.

Simon, HA & CA Kaplan1989. Foundations of Cognitive Science. In Posner, M (ed): *Foundations of Cognitive Science*. Cambridge, Mass.: MIT Press.

Quinn, R & K Espenschied 1993. Control of a Hexapod Robot using a Biologically Inspired Neural Network. In Beer, R, R Ritzman & T McKenna (eds): *Biological Neural Networks in Invertebrate Neuroethology and Robotics*, New York: Academic Press.

Rumelhart, DE, P Smolensky, JL McClelland & GE Hinton 1986. Schemata and Sequential Thought Process in PDP Models. In Rumelhart, DE & JL McClelland (eds): *Parallel Distributed Processing. Volume 2*. Cambridge, Mass.: MIT Press.

Thelen, E & LB Smith 1994. *A Dynamic Systems Approach to the Development of Cognition and Action*, Cambridge, Mass.: MIT Press.

Turing, AM 1950. Computing Machinery and Intelligence. *Mind* 49:433-460.

Von Uexkull, J 1934. A Stroll through the Worlds of Animals and Men. In Lashley, K (ed): *Instinctive Behavior*. New York: International Universities Press.

Vygotsky, LS 1986. *Thought and Language*, Cambridge, Mass.: MIT Press.

Wittgenstein, L 1953. *Philosophical Investigations*. Anscombe, GEM (trans). Oxford: Blackwell.