Externalism, Epistemic Artefacts and The Extended Mind

To appear in: Richard Schantz (ed) The Externalist Challenge. New Studies on Cognition and Intentionality de Gruyter, Berlin & New York.

Kim Sterelny
Philosophy
Australian National University
Victoria University of Wellington
Kim.Sterelny@vuw.ac.nz

Draft of July 24, 2003

I <u>Niche Construction and Epistemic Engineering</u>

A common picture of evolution by natural selection sees it as a process through which organisms change so that they become better adapted to their environment. However, agents do not merely respond to the challenges their environments pose. They modify their environments, filtering and transforming the action of the environment on their bodies A beaver, in making a dam, engineers a stream, increasing both the size of its safe refuge and reducing its seasonal variability. Beavers, like many other animals, are ecological engineers. They act to modify the physical challenges posed by their environment. Nests, burrows and other shelters reduce the impacts of adverse weather and of other agents. Animal also modify their exposure to biological risks. Hygienic behaviour reduces the impact of disease. Intensive grooming; moving to new roosts; using a "latrine burrow"; disposing excrement in faecal sacs; these all improve an animal's prospects of avoiding disease. So many organisms are like the beaver; they partially construct their own niches. They are ecological engineers, and, as John Odling-Smee and Kevin Laland have shown, niche construction is often of great evolutionary significance, transforming the effects of natural selection on both the ecological engineers and their descendants.[1]

Ecological engineering is visible to selection, for such alterations often have fitness effects that are stable across generations. So niche-constructing behaviour itself evolves. Yet until fairly recently, the evolutionary analysis of behaviour has tended to underplay the environment-altering aspects of agent action by treating treat environment as a fixed set of constraints on action. The great exception to this approach has been the social environment. Frequency dependent models of social behaviour, game theory, and sexual selection theory all treat the environment as an ensemble of other agents also making choices, and hence the social environment is not a fixed background to action. But with this exception, niche construction continues to be underplayed. For example, (Krebs and Davies 1997) is the field-defining anthology of behavioural ecology, now in its fourth edition. Yet this collection contains no paper primarily on physical or biological niche construction.

---

[1] See (Odling-Smee, Laland et al. 1996; Laland and Odling-Smee 2000; Laland, Odling-Smee et al. 2000; Odling-Smee, Laland et al. 2003), and also (Jones, Lawton et al. 1997).

This practice of treating the environment as fixed background is particularly a feature of optimality modelling. Optimality models share a set of fundamental elements. They specify the set of strategies available to an agent in a given situation; they include a measure of the costs and benefits that attach to each strategy; and they specify the features of the environment that determine those costs and benefits. Costs, benefits and their environmental determinants are fixed. The agent can only choose the right strategy given these fixed factors, the one that best trades benefit against cost. How, for example, should a bird feeding its young trade the number of trips back to the nest against the load carried on each trip? What will maximise the rate of food delivery to its chicks; an important consideration since better fed chicks have better survival prospects. Starling foraging has been analysed with this question in mind. Since the journey from foraging site to nest takes time and energy, the optimum strategy would seem to be that of bringing back the heaviest load possible, thus cutting down on travel costs. But starlings must keep what they have caught in their beak and so their efficiency declines as they forage. They must trade declining efficiency against transit costs. The optimum behaviour for starlings depends both on the distance between nest and foraging site, and the rate at which their efficiency declines. When travel costs are low because the bird is foraging close to its nest, a lighter load is more efficient. When travel costs are high, it is worth paying the cost of foraging less efficiently and carrying more per load (Kacelnik 1984).

The cost of information is quite often neglected in such models. In the case of fairly simple decision problems like that of the starling, this simplification seems reasonable. For the starling decision model has <u>a low cognitive load</u>. The information required for correct strategy choice is not extensive and may well be acquired as a by-product of their routine activity. Starlings maximise their fitness by maximising the rate at which they deliver food to their chicks, so they need to know the distance from foraging site to nest and the rate at which their efficiency declines. That information is available (though not necessarily registered) as a by-product of feeding. In other circumstances, birds must take into account both risks to themselves in foraging and in transit, and risks to the nest occasioned by their visits. For such birds, the cognitive load might be heavy. Agents often need information that is not ready to hand, not generated as an automatic by-product of their ecological activity. So they are epistemic as well as ecological agents, acting to change the informational character of their environment. Many animals are epistemic engineers.

Low-tech epistemic agency is ubiquitous. We see it in such routine action as a raptor choosing a hidden lookout post which nonetheless offers good views of its hunting ground. However, epistemic action is often more extensive and systematic than this. Many birds, once they have built their nests "decorate" them with moss and other greenery. The nests of small passerines, in particular, are often very hard to find, both by interested humans and (presumably) the birds' predators. Such nests are almost certainly cryptic by design. Their unobtrusiveness is no mere side-effect of the methods and materials from which they are constructed. These birds are engaged in epistemic counter-measures against their enemies, attempting to render their predators' informational environment opaque. Social interaction can have the opposite effect. Many species of bird give "contact calls" advertising their location to their mates while foraging, and hence making this information available for free.

This paper considers epistemic action in the human lineage, and in particular the ideas of Andy Clark, Dan Dennett and Stephen Mithen, who all take the invention and use of epistemic artefacts — tools for thinking — to be central to the explanation of human intelligence and human culture. I agree with this line of thought. But I shall argue that Clark, in particular, in concentrating on epistemic tools used only by a single agent, has somewhat mis-characterised our relation to our tools, and hence has mis-characterised the relationship between this program and more traditional representational theories of mind.

In contrast to Clark's view, I take this relationship to be one of mutual support,. For our use of epistemic artefacts explains the elaboration of mental representation in our lineage, and this elaboration explains our ability to use epistemic artefacts. In discussing our abilities to represent aspects of our environment, Andy Clark sometimes talks of "portable" representational resources: resources that an agent can deploy in responding to a range of challenges. I shall argue that the effective use of epistemic tools often depends on such portable resources. Elsewhere, I have called these states “decoupled representations”. For though they track features of the external world they are not locked into driving specific actions (Sterelny forthcoming). When decoupled representations track the world accurately they are, in Godfrey-Smith's apposite phrase, "fuels for success" (Godfrey-Smith 1996). For whatever an agent's plans, accurate decoupled representation of the world improves that agent's prospects of success. In sections III and IV I shall argue that our use of epistemic tools depends on our ability to use a rich store of decoupled representation.

II Human Epistemic Agency

For hundreds of thousands of years we have engineered our physical and biological environments. But epistemic agency has been an equally pervasive feature of hominid life styles. Indeed, Stephen Mithen has argued that tools are often both material and epistemic artefacts. A fish trap traps fish, and hence it plays a direct economic role in the life of its user. But it can also serve as a guide for the construction and location of further fish traps. The products of material culture — tools, shelter, vehicles, weapons, clothes — serve both to rework our physical and biological circumstances, and as templates, props and prompts for their own reproduction. As Mithen sees it, this is both a deep and a crucial feature of human history: perhaps 100,000 thousand years deep (Mithen 2000). Co-opting physical tools for epistemic purposes long predates using specialised epistemic artefacts. But these too have a deep history. With the invention and elaboration of pictorial representation, humans came to be makers of specialised epistemic artefacts. It is very difficult to date the first appearance of specialised epistemic artefacts, but unmistakable, superbly executed paintings are over 30,000 years old (Mithen 1998a). In Mithen's view, the use and elaboration of epistemic artefacts explains the extraordinary acceleration in both the richness and the variability of human cultures over the last 50,000 years or so. He thinks our archaeological record shows the marks of a cognitive breakthrough. Yet in all probability human cognitive powers have expanded without any significant change in human genomes. Instead, their expansion is a result of the proliferation and transmission of epistemic artefacts; of tools for thinking. As he sees it, we have got smarter, even though our naked brains have not[2].

In their recent work, both Andy Clark[3] and Dan Dennett[4] have also stressed the role of these cognitive tools in the explanation of human intelligence. Human intelligence needs explanation, for our minds are qualitatively different from those of our closest kin; even (if Mithen is right) our extinct kin with similar-sized brains. On the Dennett-Clark-Mithen view, we are not distinctively intelligent in virtue of being innately endowed with a superbly powerful, autonomous internal computational engine. Our special cognitive powers derive from our ability to extend our minds' capacities through interacting with our environment. They do not doubt that there was some internal cognitive difference between those humans who formed the base of the cognitive explosion of our lineage, and other primates with similar sized brains. But the difference was not that of having fully human intelligence. Rather, it was a small change that acted as a trigger (a "seed", in Clark's terminology) of a runaway coevolutionary process between humans and their artefacts. I think this idea is right, and shall sketch some of these ability-enhancing interactions before turning to the implications of this picture.

1. Most obviously, we alter our environment to ease memory burdens. We do so by storing information in the environment, by recoding it, and through social

---

[2] I have two major reservations about this hypothesis. First, I have been convinced that the idea of an explosive jump in human culture from about 50,000 years ago is much oversold (see McBrearty, S. and A. Brooks (2000). "The Revolution That Wasn't: A New Interpretation of The Origin of Modern Human Behavior." Journal of Human Evolution **39**(5): 453-563.. And I think that though human genomes are unlikely to have changed much, human brains almost certainly have, as a consequence of changes in our developmental environment. But these reservations do not undercut Mithen's basic point that epistemic artefacts massively enhance our cognitive capacities.

[3] See especially Clark, A. (1997). Being There: Putting Brain, Body, and World Together Again. Cambridge, MIT Press.
Clark, A. and D. Chambers (1998). "The Extended Mind." Analysis **58**(1): 7-19.
Clark, A. (1999). "An Embodied Cognitive Science?" Trends in Cognitive Science **3**(9): 345-350.
Clark, A. (2001). "Reasons, Robots and the Extended Mind." Mind and Language **16**(2): 121-145.
Clark, A. (2002). Minds, Brains and Tools. Philosophy of Mental Representation. H. Clapin. Oxford, Oxford University Press**:** 66-90.
Clark, A. (forthcoming). Mindware: An Introduction to the Philosophy of Cognitive Science. Oxford, Oxford University Press..

[4] See especially Dennett, D. (1993). "Learning and Labeling." Mind and Language **8**(4): 540-547.
Dennett, D. C. (1995). Darwin's Dangerous Idea. New York, Simon and Shuster.
Dennett, D. (1996). Kinds of Minds. New York, Basic Books.
Dennett, D. (2000). Making Tools for Thinking. Metarepresentation: A Multidisciplinary Perspective. D. Sperber. Oxford, Oxford University Press**:** 17-29..

organisation. Environmental storage is most obvious of these techniques. Our contemporary environment is full of purpose-built tools for easing burdens on memory. These include diaries, notebooks and other “organisers”. Indeed, Clark and Chalmers rely heavily on the functional similarity of internal and external memory in arguing that human cognitive systems literally include (some of) these external tools. (On this, more later). Filofaxes are new tools, but purpose-built aids to memory are certainly ancient. Not only is pictorial representation over 30,000 years old; there are cryptic objects in the archaeological record of similar age — incised bones — which may well be systems for the long-term storage of information (Mithen 1998b). Furthermore, and deeper still in the past, ecological tools have informational side-effects. Trail marking — making easily-noticed alterations to the bush while walking through it — may be older still, as there is some evidence that bonobos mark trails in this way. Written records are obviously wonderful memory aids. Writing is a recent invention, but humans have long re-coded information in public language to make it easier to recall. In songs, stories and rhyme, the organisation of the information enables some elements to prime others. Such re-coding enables us to partially substitute recognition for recall. Finally, as human societies became technically and socially more complex, the division of cognitive labour eased memory burdens; for then no member of the group had to know everything the group as a whole needed to know.

2. We transform difficult cognitive problems into easier perceptual problems Andy Clark has discussed a nifty example of this transformation. It turns out that even those with a good visual imagination find it hard to extract information from an image which is not made salient in its initial description. For example, it is much harder to notice the second interpretation of an ambiguous figure in imagination that it is when confronted with the physical image. Sketches reveal unanticipated ways of seeing the image. This, he suggests, is why initial sketching is so important for artists; why they do not just construct the finished work in imagination and then execute it. (Clark 2001),pp 132-133). Likewise, even the strongest chess players prefer to analyse positions with a real set and pieces. But though Clark's example is apt, it is important to emphasise that turning cognitive tasks into perceptual tasks is not recondite, it is

routine. We do it (for example) when we re-present quantitative information as a pictorial pattern, in pie-charts, graphs, maps.

3. We transform difficult perceptual problems into easier ones. For example, in shaping wood with a chisel and hammer, it is useful to mark the spot which you wish to strike, making it easier to focus attention on the exact working surface. Epistemic action of this kind could easily be ancient. Handaxes need to be made by reductive techniques, flaking the tool out of a larger matrix. This process would certainly be eased by scratching the point on the rock face you needed to strike. Contemporary human environments are often transformed in this way. Pedestrian crossings are painted in bright contrasting colours so they stand out against a background; likewise other warning signs. Students highlight text to help focus their attention and to aid search for later study. I have argued (Sterelny forthcoming) that in learning contexts, language is a very powerful tool for making small perceptual differences salient. Different linguistic labels for similar-looking birds alert individuals to the existence of a diagnostic perceptual difference between them.

4. We transform difficult learning problems into easier ones. For we alter the informational environment of the next generation. Non-human animals mostly alter the epistemic environment of their young as a passive byproduct of their own activities. We too alter our children's learning environment as a byproduct of our own economic and ecological activity. The development of a rich vocabulary that matches the biological, physical, artefactual and social kinds of the local environment is very important for effective action in those domains. But those taxonomies also make the informational environment of the next generation more transparent. Unlike other animals, we also invest heavily in teaching. We deliberately provide information to the next generation: “that is a horsefield’s bronze-cuckoo; you can tell if from a shining bronze-cuckoo by its more prominent white eyebrow and the fact that its chest bands are incomplete”. But we do not just provide information verbally: learning is scaffolded in many other ways. Skills are demonstrated in a form suited for learning. Completed and partially completed artefacts are used as teaching props. Practice is supervised and corrected. The decomposition of a skill into its components is made obvious; subtle elements will often be exaggerated, slowed down or repeated.

Moreover, skills are often taught in an optimal sequence, so that one forms a platform for the next.

5. We engineer workspaces so that frequent tasks can be completed more rapidly and reliably. Kitchens, for example, are organised so that the tools are handy, and so that there are safe and suitable workspaces (chopping boards and the like). Cooks often organise their ingredients before they begin cooking so that they are ready to hand when needed, and so that their physical sequence and grouping corresponds to the recipe being used (Kirsh 1995; Kirsh 1996). Clark too describes a cute example of workspace engineering. Skilled bartenders use the distinctive shapes of glasses and their sequence to cue recall for customers' orders and to code the order in which they will be served. Their ability to respond accurately to multiple simultaneous orders plummets if they are forced to use identically shaped glasses ((Clark forthcoming) chapter 8). Organised workspaces are now a ubiquitous feature of human life. Cognitive tools, too, are simplified and standardised to enhance performance on repeated tasks. Improvements in notation systems — the switch from imperial to decimal currency and measurement — makes many routine calculations easier, faster, and less error-prone.

III. Dumb Brains, Smart Minds?

Clark and Dennett both emphasise the importance of epistemic agency in explaining human intelligence, though with an important difference in emphasis. Dennett stresses the developmental role of tools for thinking. The symbol-soaked, cognitively engineered environments in which we now develop have profound implications for our internal cognitive systems. In particular, the experience of learning a public language and using sentential representations makes thought somewhat language-like. In contrast, Clark emphasises mind-artefact linkages. He has been deeply impressed by "embodied cognitive science"; intelligent agency is not explained via the computational transformation of sentence-like mental representations. Rather, embodied cognitive science favours models of adaptive behaviour that emphasise the use of information freely available in the world[5] and which emphasise the use of heuristics with low information demands. However, despite his sympathy for this line of thought, Clark suspects that some human cognitive traits remain "representation-hungry". So his best guess is that a full theory of human intelligence cannot be given by incremental improvements on these models. Nonetheless, he thinks the profound change in human cognitive evolution is the development of effective epistemic artefacts[6]. He is thus inclined to back the following horse:

> "to depict much of advanced cognition as rooted in the operation of the same basic kinds of capacity used for on-line, adaptive response, but tuned and applied to the special domain of *external and/or artificial cognitive aids* – the domain, as I shall say, of *wideware or cognitive technology*."((Clark forthcoming) chapter 8)

To a significant extent we are serial processing inference engines. But we are so only because of our reliable, designed, adjusted coupling with a staggering array of cognitive artefacts, including linguistic and quantitative systems of serial

[5] In slogan form: "the world is its own best model"

[6] It is important not to exaggerate Clark's views here. He explicitly acknowledges that humans have coevolved with their epistemic artefacts, and this changes human brains. Nonetheless, he does not emphasise the role of these internal changes in his explanation of the qualitative differences between humans and other primates.

representation. Our extended mind uses and processes linguiform representations even though (quite likely) naked human brains do not.

Furthermore, some of these artefacts can literally be parts of an agent's cognitive system. In The Extended Mind, Andy Clark and David Chalmers develop a thought-experiment about an Alzheimers sufferer (Mr T, as I shall call him). Mr T cannot unaided remember the location of an exhibit he wants to visit. But he manages such problems by writing down in a notebook crucial information for his daily plans notebook. He then acts by consulting this book. Clark and Chalmers argue that the information in the plays the same functional role for Mr T that an ordinary (non-occurrent) belief plays for ordinary human agents. They conclude we should count the notebook as part of the patient's mind, and the location of the exhibit as one of Mr T's beliefs (Clark and Chambers 1998). Clark is careful not to trivialise this extension of the boundaries of the mind. He insists that agents' minds include only those external tools to which they have regular, unfettered access: "the props and aids which can count as part of my mental machinery … are at the very least, reliably available when needed and used or accessed pretty much as automatically as biological processing and memory" (2001, p139).

While agreeing with Clark on the fundamental role of epistemic agency in explaining human rationality, I have reservations about this picture. Even when there is a reliable link between user and tool, there are important differences between internal and external cognitive resources. The external storage of information is very important, but the psychological and evolutionary dynamics of mind /filofax relations are critically different from those of mind/memory interactions. So I do not think it is helpful to think of epistemic artefacts as literally parts of the minds of the agents that regularly use them. Moreover and more importantly, Clark underplays the importance of nonexclusive use of epistemic artefacts. Many of our most important cognitive tools are common-use tools, not parts of coupled systems.

Let me begin with reliable mind-artefact linkages and the differences between internal and external storage. The effective use of tools for the external storage of information depends on the agent's perception of them. Yet there are important differences between perception and internal monitoring. Organisms are communities of co-operative and co-adapted parts. Within the agent, there is no war of deception and counter-deception. Internal routes of information flow — from memory; feedback loops from action — will be under selection for reliability. Consider the internal signals generated as an agent tries to grasp and lift an object. Over time, we would expect these signals to become cleaner and less noisy; and internal devices to become more efficient in picking them up and using them to drive appropriate responses. Internal signals keep information available on-line. Moreover, as there is no danger of deception, they do not have to be epistemically vetted. All else equal, over evolutionary time the internal informational environment of an agent will become more transparent. None of this is true of perception. An agent's perceptual systems, like their introspective systems, are under selection to deliver reliable and veridical judgements about the environment. However perception operates in an environment of active sabotage by other agents. So despite the superb design of some of our sensory systems, perception of the external environment often delivers signals that are noisy, somewhat unreliable, and functionally ambiguous (was that really an "s", we ask ourselves, puzzling at a notice or even our own handwriting).

In sum, the world we perceive is a world made in part by other agents who are not in the job of making our life as easy as possible. Moreover, perceptual tasks are not as predictable as introspective tasks. The features of the external environment we need to notice have changed, sometimes rapidly. We live in an unstable world. An agent's access to externally stored information is neither as reliable nor as uncontaminated as access to internally stored information. Thus Mr T's external memory is less reliable after dark; when he forgets his glasses; when his pen leaks or his pencil breaks; when it rains and his book gets wet. And we have not yet considered the issue of epistemic sabotage by other agents. For Mr T is at risk of thought insertion and deletion to the extent that others have access to his notebook. He may be able to guard against thought insertion by recognising his own handwriting and disregarding anything in his notebook not in his own hand. But that does not solve the problem of having his beliefs stolen. This set of problem simply does not arise for such of Mr T's information that he still codes internally.

In assessing the parallel between internal and external memory, we cannot disregard the various threats to which external memory is especially subject, even though in Mr T's particular circumstances some of these are probably unlikely. Unless there is something we have not been told, he is probably not a likely target for active sabotage. But even more important is the fact that external epistemic artefacts are used in shared and sometimes contested space. Indeed, our most important cognitive tools — language, well designed notation systems and the like — are the multi-generational product of many minds. Clark's favoured examples of the use of tools to extend our cognitive abilities tend to be of solitary activities: working through a large multiplication task; an artist using initial sketches in producing a painting; an academic writing a paper by revising drafts; cutting, pasting and annotating his way from one version to the next. Cognition, for him, remains paradigmatically a solitary vice though one prosthenically enhanced by wideware.

How typical are such cases of the role of external tools in extending our cognitive powers? Think instead of conversations, discussions, brainstorming. In such activities the public symbols you use as props and prompts are not your own, and yours too are taken up (sometimes half way through) by others. Likewise, scientific labs are shared spaces, and the tools are often shared tools. Notebooks, experiments, programs and papers are more often than not the result of many hands and minds. The same is true of decision and action in many commercial and administrative organisations. Files, for example, are often joint products. Teaching too is interactive, and the props used in teaching will not the exclusive cognitive product or resource of one individual.

External tools are in the public domain, and that has implications for the cognitive resources needed to make effective use of those tools. Jointly used epistemic artefacts are often less than optimal for any of their users: they need to be <u>individualised</u> at each use. When my department is making an appointment and I am considering a candidate's file, I need to check the file to make sure everything is there and to see what others have added. I may need to re-order the material or extract what I need. Moreover, though human interactions are often co-operative, they are not exclusively so. Deception and the hidden agendas of others are issues that arise in the use of such tools.

These complications matter. To parody Clark's picture, he thinks reliable access to dovetailed tools explains how agents with dumb brains can nonetheless be smart. It is his Alzheimer's example writ large: external tools finesse the grave limits of the naked brain. The use of epistemic tools in a common and often contested space increases the cognitive demands on the naked brain. Epistemic artefacts are typically (i) jointly produced, (ii) they exist in a common and contested space, (iii) and they often have more than one user. As a consequence, the use of epistemic tools is often cognitively demanding. Agents using common tools cannot afford to be dumb. This fact forges a link between the Clark-Dennett-Mithen picture of minds enhanced by epistemic artefacts and representational externalism. I take that externalism to be the idea that internal states are representations because they are tracking states: their representational properties are constituted by relations to the agent's environment[7]. An externalism worth defending also claims that these representational properties explain the behaviour, or the behavioural capacities, of the representing agent. I shall argue that such representations explain our ability to use of epistemic artefacts successfully.

[7] The different versions of naturalistic semantics (for example, the different versions of indication theory) give differing accounts of this tracking relation. I shall not attempt to adjudicate between these versions, largely because I think it is unlikely that any single tracking relation corresponds cleanly to a folk psychological notion of content. Nor is it likely that only one tracking relationship is explanatorily important.

<u>IV Two Kinds of Externalism</u>

In their 1998, Clark and Chalmers contrast their extended mind with representational externalism:

> "In these cases, the human organism is linked with an external entity in a two-way interaction, creating a *coupled system* that can be seen as a cognitive system in its own right. All the components in the system play an active causal role, and they jointly govern behavior in the same sort of way that cognition usually does … Our thesis is that this sort of coupled process counts equally well as a cognitive process, whether or not it is wholly in the head.
>
> This externalism differs greatly from standard variety... When I believe that water is wet and my twin believes that twin water is wet, the external features responsible for the difference in our beliefs are distal and historical, at the other end of a lengthy causal chain. Features of the *present* are not relevant: if I happen to be surrounded by XYZ right now (maybe I have teleported to Twin Earth), my beliefs still concern standard water, because of my history. In these cases, the relevant external features are *passive*. Because of their distal nature, they play no role in driving the cognitive process in the here-and-now. This is reflected by the fact that the actions performed by me and my twin are physically indistinguishable, despite our external differences.
>
> In the cases we describe, by contrast, the relevant external features are *active*, playing a crucial role in the here-and-now. Because they are coupled with the human organism, they have a direct impact on the organism and on its behavior. In these cases, the relevant parts of the world are *in the loop*, not dangling at the other end of a long causal chain. Concentrating on this sort of coupling leads us to an *active externalism*, ... " (1998 p11)

I do not think this is the right picture of the relationship of representational externalism to the extended mind. In rejecting this picture I emphasise two themes. One is the importance of portable information for the use of epistemic artefacts, The second is the importance of tracking relations in explaining the successful use of these artefacts. As a consequence of these two themes, in my view our ability to use epistemic tools effectively is not just <u>enhanced</u> by decoupled representations, it <u>depends</u> on our ability to form and use accurate decoupled representations. Embodied cognitive science often explains adaptive action by models in which an agent's representational resources are tightly fused to specific acts and problems. Webb's robot cricket recognises and finds an appropriate mate by being able to follow a sound signal to its source. It does so through having ears structured to channel sound but only of the right frequency, and by having response mechanisms dedicated to driving the cricket to the source of the sound. The system represents (or at least tracks) the source of the sound, but not in any way that is a general fuel for success ((Clark 2001) pp 127-128). The use of epistemic artefacts cannot, I claim, be supported by information tightly coupled to a particular task. Rather, the evolution of multi-purpose representation is tied to the evolution of our capacity to use our epistemic artefacts. I shall return to the role of multi-purpose information shortly; let me now turn to the explanatory salience of tracking.

Externalists take representational properties to be some species of systematic, natural relationship between an internal state of the agent and some aspect of that agents' environment. Representational properties are relational properties, and in particular, they are relational properties that explain the existence of cognitive states by explaining our dispositions to behave adaptively. For the representational properties of thoughts explain the success and failure of agents as they try to achieve their goals on the basis of their representations of their environment. It may be the case, as Clark and Chalmers note, that we can explain the execution of a particular act without appealing to these representational properties of the agent's thoughts. Likewise, it may even be true that in explaining the success of any particular act, we do not need to appeal to the representational properties of thought. For example, in explaining Michael's success in finding the wine, it might be enough to explain his orientation at, and movement towards a particular cupboard. The execution of that action sequence, together with the fact that wine was in the cupboard, explains why Michael found the wine. But a disposition to succeed cannot be explained in the same way. For such dispositions are open-ended and counterfactually robust. If (within limits) Michael would have found the wine wherever it was, this can be explained only on the basis of a capacity to accurately track the wine's location (Sterelny 1990; Godfrey-Smith 1996).

Epistemic tools support open-ended and counterfactually robust dispositions to succeed. Thus one distinguishing feature of good field naturalists is their identifications skills about the plants and animals of their local patch. But another distinguishing feature is their ability to use key guides, floras, handbooks, field guides

— in short, relevant epistemic artefacts — in unfamiliar patches to identify unfamiliar plant and animals. They use these tools successfully because they know which characters are typically diagnostic; because they know how these epistemic tools are organised; because they can identify habitat types and types of behaviour. Without understanding how these tools are internally structured, and without accurate representations of such biological phenomena, a flora is just dead weight. Field biologists have to learn to use these tools. The standard field guide to the birds of Venezuela illustrates well over a thousand species (Hilty 2002), and describes many more forms of these species (for example, immature birds) using a set of standard terms for describing the parts of a bird's plumage (scapulars, primaries, secondaries, vent etc). A user who does not understand the significance, organisation and presentation of this information would be overwhelmed by its quantity. But once these tools are mastered, they effect a large and open-ended expansion of their user's skill base. Thus a skilled field biologists has a disposition to succeed in identifying plants and animals in a wide variety of circumstances. We cannot explain this disposition without appealing to the accuracy of her ecological and taxonomic representations and her grasp of the organisation of biological data bases.

Field biologists thus have an open-ended disposition to succeed in a wide variety of biological projects. Moreover, those dispositions depend on information that is not fused to specific acts in the manner of the cricket's information about the nature and location of a potential mate's call. The biologist's information is multi-purpose. Agents acquiring it typically do not know when, whether, or how this information will be used. A field botanist learning the taxonomically-informative varieties of leaf shape will not know the circumstances in which that information will become relevant to her actions. Agents who live in unstable environments face challenges whose demands cannot be predicted in advance. They are disposed to succeed only if their inner cognitive states both accurately track their environment (often enough) and support a variety of plans.

I have argued that the use of epistemic tools depends on such rich representational resources, and that the disposition to use these tools successfully depends on the existence of appropriate tracking relations between inner resources and environment.

But this argument vindicates no particular version of externalism: no particular hypothesis about the relation between internal state and external environment which constitutes tracking by the state of the environment[8]. Indeed, several different kinds of tracking relationships may have played important roles in the evolution of rich representational capacities (Godfrey-Smith 2002; Godfrey-Smith forthcoming; Sterelny forthcoming). Let me now connect these ideas about the explanatory importance of tracking to the fact that the information needed to use epistemic tools is often multi-purpose.

The cricket's information about the direction and identity of a sound is tied to a very specific purpose. This is not true of the information needed to use biological tools and in this respect they are typical of epistemic tools. Agents with no information about botanically diagnostic characters; agents with misinformation about such characters and agents able to recognise the plants of her own patch only through unanalysed gestalts will all struggle to use a flora. Only informed agents can use a flora. Moreover, agents typically acquire this information piecemeal, without knowing its functional significance. This information is not learned in a form which ties it to specific uses, routines or skills. It is not embedded in executing specific tasks. While this is typical of the informational demands on the use of epistemic artefacts, it is not typical of the informational demands on nonhuman animals[9]. Decoupled representation has coevolved with the use of epistemic artefacts because agents need a rich information base to effectively use epistemic artefacts, and because they acquire that information piecemeal, without knowing its functional significance. If this co-evolutionary hypothesis is right, we could not use epistemic tools without internal states that both accurately track features of our world and which can contribute to an

---

[8] However, it does rule some options out. As Peter Godfrey-Smith notes, we can explain an agent's successful action on the basis of its accurate representation of its environment, but only if "accurate tracking" is defined independently of success. The facts which constitute tracking the world must be independent of the behavioural triumphs of the mind-owning organism. Godfrey-Smith argues that Millikan's ambitious version of teleosemantics does not define what an agent's thoughts are about independently of the agent's dispositions to succeed in his or her behavioural plans Godfrey-Smith, P. (1996). Complexity and the Function of Mind in Nature. Cambridge, Cambridge University Press.. If so, the correspondence of those thoughts with the world cannot explain the agent's successes.

[9] I doubt that decoupled representation is uniquely hominid, but if this line of thought is right, the extensive use of information not tied to fairly specific tasks may be uniquely hominid This conjecture is supported by the fact that It is surprisingly difficult to find unequivocal evidence of the use of spatial and social maps, or evidence of broad-banded understanding of the causal properties of physical objects in nonhuman animals.

open-ended array of action. And if we did not use such tools, we would have relatively little need of such states.

Let me sketch the case for thinking that biological epistemic tools are typical rather than exceptional. First, public representations and other epistemic tools have to be interpreted. Perhaps photos, drawings, maps and the like can be used by treating them as natural signs of the environmental structures they map. But even with such symbols, interpreters need to master the regularities of their fit to the world. A map, for example, maps some aspects of its domain. Thus maps of an underground system typically represent the order of the stations and the connections between the various lines, but they do not map the distance between stops. When distances are mapped, an agent needs to learn the scale of the map, and the arbitrary conventions maps use; for example, using colour to code hight above sea level. Moreover, these features of the use of maps and similar representations must be genuinely learnt. They are variable and contingent, so (in contrast, perhaps, to internal symbols) they cannot simply be implicit in the automatic routines for the use of a representation. Yet, equally, there are regularities: having learnt to use a map of the Sydney suburban network, using that for London is less intractable. Map reading depends quite extensively on portable information. For maps are read with different purposes, and this makes different aspects of a map important. The same map of a chunk of New Zealand's fiordland will be used very differently by hikers, hunters and fishermen.

Models and templates, likewise, require interpretation. A fish-trap carries information about how and where to make other fish-traps. But the template cannot be blindly copied, even by an agent who could commit every detail to memory. For the length, depth and shape of the trap must be sensitive to the size and depth of the inlet the trap closes at low tide, and to the strength of the tidal water that will run over it. When another agent makes an artefact for his own purposes, it is not automatically a perfect template for my use: the other agent may be larger or shorter; weaker or stronger; a left-hander. I shall need to modify as well as copy his production. Our information-intensive ways of life often relies on reverse-engineering the tools and skills of others. But reverse engineering depends on a rich prior background that enables an agent to select a model, filter noise from signal (not everything the model does while demonstrating a skill is part of the skill), and adjust for their differences with the model. So the reverse-engineering agent needs a rich, relevant and largely accurate system of decoupled representation. Imitation is sometimes discussed as if it were a cognitively cheap form of learning, but the very limited imitative capacity of the great apes is a signal that it is cognitively demanding, though the exact nature of that demand remains controversial.

Symbol systems are now amongst our most important epistemic artefacts. These tools make possible forms of human reasoning that would otherwise be impossible. Without positional notation and without algorithms which decompose large arithmetic operation into elementary ones, accurate quantitative reasoning would be impossible. In using these notation systems we do indeed "store information in the world". Yet the appropriate use of these symbol structures is cognitively demanding. The innumerate are not rare even in western societies making serious attempts to make numeracy skills universal. The genuinely arbitrary symbol systems of language impose greater demands still. In particular, there are persuasive argument for accepting the Grician view that to understand an utterance, we have to represent its speaker meaning (Origgi and Sperber 2000). And while analyses of speaker meaning vary widely, they all involve representations of thoughts; beliefs about beliefs and/or intentions. Perhaps an agent can acquire the concept of a tiger from a tiger-photo, without representing that photo as a representation. I do not see how an agent could acquire the concept of a tiger from tiger-conversations without representing “tiger” as a representation.

Furthermore, if an agent must guard against deception or manipulation, then the cognitive load of the use of shared epistemic tools is greater still. For an agent then needs not just bare metarepresentational capacities but assessment tools. Dan Sperber has argued, very plausibly that the meta-tools of folk logic — the notions of truth and evidence, of sound and unsound reasoning — evolved as a response to this vetting problem (Sperber forthcoming). Counter-deception is a problem whose informatinal load is both heavy and unpredictable: there is no telling in advance what you will need to know in order to know that another is a liar. As I argue in (Sterelny forthcoming), the vetting problem is particularly pressing for linguistically coded information. The arbitrariness of linguistic encoding and the fact that language is not stimulus bound are two of the features which make language such a powerful system. But they also make it a deception-subject-system. Other informational vehicles — especially templates and models, but to some extent diagrams and the like — have intrinsic marks of reliability. It is much harder to fake a tool than a story.

The use of epistemic tools in a public space involves quite complex problems of co-ordination. A recipe is a fairly standard example of an epistemic artefact. So consider a group of friends jointly producing a meal by following a recipe. Each agent must (a) monitor what others are doing; (b) negotiate a division of tasks; (c) negotiate a division of shared space and shared work surfaces; (d) negotiate a division of shared tools — who gets to use which chopper when. There will often be reasonably generous margins of error, but they must form and follow a co-ordinated plan. None of this can be done by reflex. It will all depend on what they know of one another, their materials and their tools. Who gets flustered easily? Who thinks that if some is good, more is better? Without a rich stock of background information, many joint activities will unravel. Of course, if there are repeated interactions of the same kind, as in a commercial restaurant, then these co-ordination decisions will be standardised and will not be re-litigated case by case. But when friends get together to cook a joint feast, all this must be done on the fly, and usually is, without any great drama. However, the fact that we usually solve such problems without great stress does not show they have a light cognitive load.

A lesson of all these examples is that human problem solving makes intensive demands on memory. Very often, when we solve problems, the information we use is not available on line at the site of action. Each agent in the cooking co-ordination problem brings to the kitchen a good deal of knowledge of food and cooking, and (very likely) of the quirks and idiosyncrasies of their partners in gastronomy. The apprentice maker of fish-spears is likely to have acquired information about fish and the water, and about the properties of wood and twine, by trial and error exploration of his local environment. We soak up a good deal of social, technical, psychological and ecological information in contexts which do not signal the specific relevance of that information for later action. Sometimes we acquire information for specific functional purposes. We learn the characteristics of a program like Endnote only because we want to use it in composing documents. But much of the information we acquire and then use in the general exploration of our world is not labelled for its relevance for action. At the time we find out that George thinks there is no such thing as too much garlic, we do not know how or whether this quirk will be relevant to our later actions.

So the use of epistemic tools often requires the recall of information not initially acquired to facilitate its use. Moreover, the use of epistemic tools — especially jointly used epistemic tools — requires agents to integrate information across informational domains. A particularly clear example of this integrative demand is the use of artefacts as templates or models. Imagine a boy learning how to make a fish spear from an adult, using a completed example to guide his own construction. To acquire this skill in this way, the boy will need to grasp (a) ecological information: he needs to understand why the fish spear works, and to do this, he will need to understand something about fish, how to spot them, and how to make allowances for misleading perceptual appearances; (b) he will need to be able to interpret the adult's behaviour; for example, to understand that a demonstration is slowed down or exaggerated; (c) he will need some understanding of the physical properties of the materials with which he is working. Learning to make and use artefacts from artefacts will typically require an agent to integrate information about the physical properties of the tool; information about its uses; and information about the intentions of the model. In short, the effective use of epistemic artefacts depends on informational resources internal to the agent. Moreover, for the most part these resources are portable. They are not tied to particular functions

Time to sum up the state of play. Dennett, Clark, Mithen (and of course others) have argued that we cannot understand the evolution, operation or development of human intelligence without understanding the multitude of ways in which we act as epistemic agents, transforming the informational load on our own decision making and that of others, and the ways we act as epistemic tool makers, constructing devices that help us meet those transformed informational challenges. These ideas are true and important. In this paper, my own twist has been to emphasise the social rather than the individual aspects of the use of epistemic tools, and (as consequence) the informational load on their effective use. A consequence of this shift in emphasis (it is no more) is that I see the extended mind perspective on human cognition as more nearly complementary with some of the research agendas associated with externalism and the representational theory of mind. This too is only a change in emphasis. Even if I am right in my assessment of the informational load on the use of wideware, that would not vindicate a picture of the mind as an autonomous, sentence-crunching computational engine. But it would vindicate the idea that portable cognitive resources (Godfrey-Smith's "fuels for success"; my "decoupled representations") play a central role in human cognitive life, probably a much more central role than they play in the cognitive life of other agents. Moreover, the importance of portable resources cannot be explained without taking into account their tracking profiles. Portable cognitive resources are fuels for success only because they are both world-tracking and action-driving.

References

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

Clark, A. (1999). "An Embodied Cognitive Science?" Trends in Cognitive Science **3**(9): 345-350.

Clark, A. (2001). "Reasons, Robots and the Extended Mind." Mind and Language **16**(2): 121-145.

Clark, A. (2002). Minds, Brains and Tools. Philosophy of Mental Representation. H. Clapin. Oxford, Oxford University Press**:** 66-90.

Clark, A. (forthcoming). Mindware: An Introduction to the Philosophy of Cognitive Science. Oxford, Oxford University Press.

Clark, A. and D. Chambers (1998). "The Extended Mind." Analysis **58**(1): 7-19.

Dennett, D. (1993). "Learning and Labeling." Mind and Language **8**(4): 540-547.

Dennett, D. (1996). Kinds of Minds. New York, Basic Books.

Dennett, D. (2000). Making Tools for Thinking. Metarepresentation: A Multidisciplinary Perspective. D. Sperber. Oxford, Oxford University Press**:** 17-29.

Dennett, D. C. (1995). Darwin's Dangerous Idea. New York, Simon and Shuster.

Godfrey-Smith, P. (1996). Complexity and the Function of Mind in Nature. Cambridge, Cambridge University Press.

Godfrey-Smith, P. (2002). "On the Evolution of Representational and Interpretive Capacities." The Monist **85**(1): 50-69.

Godfrey-Smith, P. (forthcoming). On Folk Psychology and Mental Representation. Mental Representation. H. Clapin, P. Staines and P. Slezak, Greenwood Press.

Hilty, S. L. (2002). Birds of Venezuela. Princeton, Princeton University Press.

Jones, C., J. Lawton, et al. (1997). "Positive and Negative Effects of Organisms as Physical Ecosystems Engineers." Ecology **78**: 1946-1957.

Kacelnik, A. (1984). "Central Place Foraging in Starlings (*Sturnus vulgaris)*. I Patch Residence Time." Journal of Animal Ecology **53**: 283-299.

Kirsh, D. (1995). "The Intelligent Use of Space." Artificial Intelligence **73**: 31-68.

Kirsh, D. (1996). "Adapting the Environment instead of Oneself." Adaptive Behavior **4**(3/4): 415-452.

Krebs, J. and N. Davies, Eds. (1997). Behavioural Ecology: An Evolutionary Approach. Oxford, Blackwell.

Laland, K. N. and F. J. Odling-Smee (2000). Niche Construction, Ecological Inheritance and Cycles of Contingency. Cycles of Contingency. R. Gray, P. Griffiths and S. Oyama. Cambridge, MIT Press**:** 117-126.

Laland, K. N., J. Odling-Smee, et al. (2000). "Niche Construction, Biological Evolution and Cultural Change." Behavioral and Brain Sciences **23**: 131-175.

McBrearty, S. and A. Brooks (2000). "The Revolution That Wasn't: A New Interpretation of The Origin of Modern Human Behavior." Journal of Human Evolution **39**(5): 453-563.

Mithen, S. (1998a). The Supernatural Beings of Prehistory and The External Storage of Religious Ideas. Cognition and Material Culture: The Archaeology of Symbolic Storage. C. Renfrew and C. Scarre. Cambridge, McDonald Institute**:** 97-106.

Mithen, S. (1998b). A Creative Explosion? Theory of the Mind, Language and The Disembodied Mind of the Upper Palaeolithic. Creativity in Human Evolution and Prehistory. S. Mithen. New York, Routledge**:** 165-186.

Mithen, S. (2000). Mind, Brain and Material Culture: An Archaeological Perspective. Evolution and the Human Mind: Modularity, Language and Metacognition. P.

Carruthers and A. Chamberlain. Cambridge, Cambridge University Press**:** 207-217.

Odling-Smee, F. J., K. N. Laland, et al. (1996). "Niche Construction." American Naturalist **147**: 641-648.

Odling-Smee, J., K. Laland, et al. (2003). Niche Construction: The Neglected Process in Evolution. Princeton, Princeton University Press.

Origgi, G. and D. Sperber (2000). Evolution, Communication and The Proper Function of

Language. Evolution and The Human Mind. P. Carruthers and A. Chamberlain. Cambridge, Cambridge University Press**:** 140-169.

Sperber, D. (forthcoming). "An Evolutionary Perspective on Testimony and Argumentation." Philosophical Topics.

Sterelny, K. (1990). The Representational Theory of Mind: An Introduction. Oxford, Blackwell.

Sterelny, K. (forthcoming). Folk Logic and Animal Rationality. Rational Animals? S. Hurley and M. Matthew Nudds. Oxford, Oxford University Press.

Sterelny, K. (forthcoming). Thought in a Hostile World. New York, Blackwell.