Cogn Process (2006) 7: 3–10
DOI 10.1007/s10339-005-0015-9

REVIEW

Mitch Parsell

The cognitive cost of extending an evolutionary mind into the environment

Received: 16 June 2005 / Revised: 13 July 2005 / Accepted: 18 July 2005 / Published online: 26 August 2005

Abstract Clark and Chalmers (1998) have argued that mental states can be extended outside an organism's skin. In response to some worries about the availability, reliability and portability of such extended resources, Clark (2005) offers a set of rough criteria that non-biological objects must fulfil to legitimately ground mental states. One such criterion is that the information retrieved from these non-biological sources be (more or less) automatically endorsed. But Sterelny (2003, 2005) has persuasively argued that the extended sphere is epistemologically opaque: a domain of contested truth and deliberate deception. As such, retrieving information from this domain requires the deployment of social guards for the information to remain reliable. But deploying such guards would seem to endanger endorsability by increasing cognitive load. Here I demonstrate that deploying social guards does not increase cognitive load if the guards are implemented in a highly distributed connectionist economy or off-loaded to the external environment.

Keywords Extended mind · Evolution · Connectionism

Clark and Chalmers (1998) advance the most outrageous thesis in the philosophy of mind since eliminative materialism (Churchland 1988). But unlike eliminativism, which is clearly false, their position is, in retrospect at least, obviously true. The thesis is the Extended Mind, which Clark (2005) more recently describes as the view "that mental states, including states of believing, could be grounded in physical traces that remain firmly outside the head".

The main argument for extending mental states into the environment is based on parity of reasoning. Once the biological body is accepted as an arbitrary barrier, we are forced to accept parity between outer epistemic artefacts and inner informational states. Of course, Clark and Chalmers (1998) recognise a range of apparent disanalogies between inner and outer informational states. The most significant ones concern functional issues of portability and availability, concerns that have led to the proposal of a set of rough criteria for which an extended object counts as grounding a mental state. The most critical from the present perspective is endorsability: that information coming from external sources should be pretty much automatically endorsed. The reasoning is as follows: to the extent that a state is treated with endorsement by default, it is on par with an internal source of information and thereby properly considered part of a coupled mental process. If, on the other hand, the information is subject to much cross-checking, it is unlike an inner information flow and thus not part of a mental process. When endorsement is (almost) automatic, extended resources can be legitimately considered part of the mind. But why would an agent want to bleed their mind into the environment? One reason is to reduce cognitive load.

External epistemic resources have long been recognised to decrease cognitive load. Consider cases such as the use of pen and paper to aid complex mathematical exercises and the general and broad range of cultural artefacts like language, books and diagrams (Clark and Chalmers 1998). For some, the ability to off-load is the centre of human intelligence. Dennett (1996, pp 134–135; emphasis in the original), for example, claims we humans are so intelligent precisely because we are in the

> habit of off-loading as much as possible .... into the environment .... where a host of peripheral devices we construct can store, process, and re-represent our meanings, streamlining, enhancing and protecting the processes of transformation that are our thinking.

This seems almost trivially true of modern people. Few people bother remembering phone numbers any-

Communicated by John Sutton

M. Parsell
Department of Philosophy, Macquarie University,
Sydney, 2109 NSW, Australia
E-mail: mparsell@scmp.mq.edu.au
Tel.: +61-2-98508826
Fax: +61-2-98508892

more, but rely on their mobiles to fulfil such mundane memory functions. But the use of epistemic scaffolding has been around far longer than mobile phones. Indeed, such scaffolding seems central to human evolution. Evolution favours naked brains geared to off-loading memory-intensive problems to the environment and, in at least some situations, to off-loading complex computational problems by means that allow the transformation of those problems into more tractable ones. Visual processing has evolved to exploit contingent facts about the structure of the visually perceptible environment. Our perceptual processes seem to make use of bodily and locomotive motion to reduce computational load on the visual system.[1] Indeed, it would be disadvantageous for organisms to develop inner resources if the external environment can be exploited to achieve the task. Thus an evolutionary eye demands attention to extended epistemic resources. This has recently been strongly argued by Sterelny (2003).

Sterelny (2003) presents a wonderfully detailed and highly original account of human evolution, placing epistemic scaffolding in the front and centre. He argues that selective pressure operates at a group level for humans because of our unsurpassed cooperative nature. This cooperation sees humans form extended social groups that rely for their existence and maintenance on the development and employment of a range of epistemic artefacts. But, while Sterelny's account relies upon extended epistemic artefacts, he has recently questioned whether such extended states can reduce cognitive load (Sterelny 2005). The use of epistemic artefacts does not come cheap because the external environment is epistemically opaque, a place of contested truth and deliberate deceit. As such, the external world requires intelligence to navigate. Sterelny (2005) draws two conclusions from this: (1) that extended states do not reduce the cognitive load on the naked brain, and (2) that there is a non-arbitrary boundary between inner cognitive processes and external epistemic artefacts. Clark (2005) denies both conclusions, but restricts his response to the second. My focus is on (1). I will, however, demonstrate that should Sterelny be successful in establishing (1), then (2) will follow. To demonstrate this implication, one must establish that increased cognitive load endangers endorsability. Briefly, the opacity of the social realm demands the deployment of what I will term "social guards": guards used to detect cheating by members of one's social group. Deploying such guards seems to increase the load on the naked brain. If true, this would undermine endorsability by requiring increased attention when employing external resources. Thus extended resources would not be automatically endorsed and, as such, would not by parity of reasoning qualify as part of mental economy.

In this paper, I demonstrate that deploying social guards need not increase cognitive attention. I accept that Sterelny has demonstrated that a reliance on extended artefacts demands the deployment of social guards. But I will show that this does not undermine endorsability if either the social guards reside in a non-symbolic cognitive economy or it is possible to off-load the guards to the environment. In either case, the deployment of social guards will not require increased attention and, consequently, endorsability will not be threatened. The discussion is divided into three sections. In the first two, I briefly rehearse the extended mind thesis and Sterelny's account of human evolutionary history, respectively. In the final section, I establish that although Sterelny has demonstrated the need for social guards such guards do not undermine endorsability.

The extended mind

Although, upon first encounter, extended mental resources may seem weirdly exotic, the extended mind thesis is neither as original nor as bizarre as it may first appear. Theories in both psychology and philosophy have come close to positing extended mental resources. In psychology, behaviourism attempted to articulate and individuate mental state in terms of purely external language: as input, output and learning histories. In philosophy, externalism of content places the meaning of mental states in the external realm. Of course, neither position really extends the mind outside the body, but both are steps in that direction. Nor are extended mental resources as exotic as first impressions would suggest. We have already seen strong reasons for extending epistemic resources into the environment. Cases of extended epistemic action like using pen and paper require spreading epistemic credit to the environment (Clark and Chalmers, 1998, p 2). But no one prior to Clark and Chalmers (1998) extended mental credit into the world. That is, although the individuation, meaning and epistemic resources implicated in mental life may previously have been extended, Clark and Chalmers (1998) are the first to see external artefacts as mental resources. Their reasoning, in hindsight, is both obvious and persuasive. If one is prepared to bleed epistemic credit into the world, then parity of reasoning demands a parallel bleeding of the mind: if the brain depends on external resources to function intelligently and those resources are reliably available, it is prima facie plausible to include those resources within the domain of the mental. Still, it could be objected that external resources, while causally relevant in some situations, lack the portability of in-the-head processes.

[1] Clark and Chalmers (1998) cite Blake and Yuille (1992), but there are a plethora of other studies that could be drawn on to establish the same point. Indeed, research coupling cognition and the environment in visual perception goes back at least to Gibson's classic (1979) study *An Ecological Approach to Visual Perception*. For more recent examples of research along these lines, see Shaw and Turvey (1999a, b). Of course that phrase "extended mind" would be anathema to those working in the ecological psychology tradition following Gibson. These researchers tended to view their programme as a means of bringing the environment inwards rather extending out cognition.

Brains make their resources pretty much constantly available.[2] Some external resources, particularly physical ones like diaries and mobile phones, lack such portability. Nevertheless, the brain has evolved to take advantage of those external environmental resources that are constantly available and it makes perfect evolutionary sense for the brain to do so. Furthermore, some abstract, non-physical human creations, like language, do seem on pretty much constant demand. And language has had a significant impact on the nature of the brain itself. This is true for the human species over evolutionary time and for each individual in the process of acquiring cognitive abilities. Therefore language (at least) seems crucial "in constraining the evolution and development of cognition" (Clark and Chalmers 1998, pp 11–12). Thus some external resources are core cognitive abilities: reliable, portable and central to a range of skills that are clearly cognitive. As such, we gain explanatory power and ease by treating these extended resources on a par with inner resources. But as extended resources exist in shared space, the endorsement of these resources is likely to be less than automatic.

Consider the Internet. The Internet is a vast repository of information. Unfortunately, it is also a vast repository of misinformation. When I submit a search to Google, I do not automatically endorse the information on the first site listed, but cross-check that information with other sites. A personal website seems different: as only the owners can legitimately upload information, they will be likely to endorse that information. Still, as they exist in shared space, even personal websites are open to corruption by others. Indeed, all extended resources seem more liable to corruption than inner resources. There is, therefore, a case for treating extended resources with a different degree of endorsement than one treats inner informational states. In seems an undeniable fact that external resources are open to other actors in ways one's brain is not. This is not to deny that a brain could be manipulated. The point is merely that the extended sphere, due to our recognition of its essentially social and contested nature, is unlikely to be more or less automatically endorsed in the way that inner resources typically are. This is even true of personal resources. It is surely true that the prospects of someone corrupting my diary, mobile phone or Internet site are greater than those of someone corrupting my brain. Thus, all other things being equal, extended resources are open to more scrutiny and cross-checking than inner resources. Indeed, Sterelny (2003) has recently argued that human evolution has designed us to include a number of protocols for checking the reliability of extended artefacts and, more recently (Sterelny 2005), that deploying these protocols increases the demands on the naked brain. It seems to me that this threatens endorsability and thereby endangers the entire extended mind thesis.[3]

Human evolutionary trajectory

Sterelny's (2003) account of human evolutionary trajectory emphasises the importance of cooperation in human evolutionary history. Selective forces have operated at a group level for humans because of our cooperative nature and reliance on non-genetic inheritance, such as by imitation and the construction of epistemic artefacts. Importantly, non-genetic inheritance and cooperation form a tight explanatory loop: the development and employment of epistemic artefacts, such as language, depends on cooperation and cooperation is itself enhanced by such artefacts. Still, although we rely on cooperation, we cannot assume others will be cooperative. Other people have different life projects that, although dependent on the success of the extended group, need not coincide with our own particular projects. Thus we need to employ social guards to warn against possible deception. Such social guards are elaborations of non-social deception detectors.

Sterelny (2003) makes a plausible case for supposing that interaction with a physical environment without predators is a relatively simple epistemic task that can be handled by simple tracking devices without the need for internal representations. But predation significantly complicates matters. In the quest to eat you, predators will often attempt to mislead. They will hide and mimic in the hope of getting close enough to kill. The result of such deceptive acts is an environment in which information is deliberately misrepresented. To live successfully in such an opaque realm our cognitive innards need to be more complex than simple tracking devices. We require a robust inner representational economy.

Predators pollute the epistemic world, making information less reliable. When no cue is sufficiently reliable to signal the presence of potential threat, evolution favours the ability to call upon multiple cues, by, for example, relying upon multiple perceptual channels (see Sterelny 2003, pp 11–29). For animals that live in extended social groups, a parallel process favours employing multiple individuals. If it is better to use two or more perceptual channels to signal predation, it is similarly better to use two or more perceptual systems. Spreading the cognitive load for detecting the deception of predators across groups of organisms leads to more successful organisms. This spreading of predator detection is well documented in a large number of animal

[2]To claim that brain processes are always available is to over-state the brain's abilities. Sleep, intoxication and other factors can significantly reduce the cognitive resources that can be reliably called upon.

[3]As will become clear, I am pointing to a prima facie conflict between reliability and endorsability for resources extended into a social world. This is not to deny that reliability and endorsability are distinct criteria. Indeed, the apparent conflict between the two trades on the fact that they can come apart, for it is in ratcheting up reliability that endorsement is lost.

species. It has even been observed to cross species. Rainey et al. (2004) have recently established that Diana monkeys (*Cercopithecus diana*) and Yellow-casqued hornbills (*Ceratogymna elata*) recognise the warning calls made by each other. But even in the case of signalling within a species, such cooperative efforts have costs. Most importantly from the present perspective, it increases one's reliance on others, which increases cognitive costs if there is the possibility of cheating within the group.

Cooperation cannot be assumed in predator detection. Announcing the presence of a predator also announces one's own location to the predator, which thereby increases the likelihood of oneself becoming prey. Thus there is a pull for individuals to cheat on cooperative detection tasks: to seek a free ride and gain the detector advantage without paying the cooperative costs. For the group it makes good evolutionary sense to protect itself against such cheating, or over evolutionary time the group will be swamped by cheaters and the cooperative advantage lost. There seem to be two general ways to protect against such cheating: (1) reward cooperation or (2) punish cheating. We need only be concerned with (2). The first step to achieve (2) is by the detection of deception. But such a task is more difficult than predator detection itself. It is easier in most instances to hide what you are aware of than to hide your actual presence. Thus, evolution will tend to produce social cheaters that attempt to hide their knowledge, proto-beliefs, motives and the like. Indeed, such social deceivers—organism that attempt to deceive members of their own extended social group—will be favoured whenever the costs of deception are outweighed by the reduction in costs associated with not cooperating.[4] Of course, the failure to signal the presence of predators and the attempt to hide such cheating is just the fist step in social deception.

Generally, social deception is the attempt to deceive members of one's own extended social group, as opposed to, for example, in attempting to deceive one's prey. Such deception often consists in not merely misrepresenting external facts—the presence of a predator—but also one's mental states—beliefs, motivations and the like. Consider a simple illustration of social deception taken from Sterelny (2003, pp 58–59). A young baboon notices an adolescent digging up a tuber and screams as if attacked once the tuber is uncovered. The youth's mother arrives to chase the adolescent away and thereby leave the young baboon to enjoy the fruits of her deceptive labour. The detection of such deception carries a far higher cognitive burden than the detection of the deception typically displayed by predators. It requires not merely a representational, but a meta-representational economy. The need for such complex cognitive innards is obvious in human affairs. In order to know when another person may be attempting to mislead, and thereby to know which claims need cross-checking, one needs an understanding of the beliefs, desires and possible motives of that person. Consider such cross-checking as phoning named referees to guard against exaggerations on job applications or seeking collaborating evidence before trusting an email correspondent with one's bank details. Because of the centrality of cooperative interactions to human affairs—both in the here-and-now and during our evolutionary history—the ability to deploy guards against such social deception must occupy a central role in human cognition. Indeed, one of the most widely recognised theories of evolutionary psychology posits a mental module devoted to cheater-detection. Social Contract Theory (SCT) posits the existence of an expert system within the human brain devoted to social exchange. This module is further supposed to contain a sub-routine designed specifically for the detection of cheaters (see Cosmides 1989; Cosmides and Tooby 1989, 1992, 1997). Thus SCT is in essence a modular, innate account of social guarding. I will return to SCT presently.

To pull the preceding threads together, spreading epistemic tasks across a group has costs. By reducing one's own attention to predation, for example, one increases reliance on others. But others may not be playing fair. They may be attempting to gain a free ride and it makes good evolutionary sense for them to do so: they will gain the benefits without paying the costs. Moreover, it makes just as much evolutionary sense for them to hide their cheating or they are likely to face some form of punishment. Thus, adding even such a simple social dimension as predator detection to life further contaminates the environment, for social actors not only want to hide themselves, but also their desires, motivations and the like. The end result of this process of increased epistemic pollution is, according to Sterelny (2003), the need for the development of inner representational systems that includes meta-representational capacities and a range of cheater-detection protocols. For humans who depend on group cooperation for far more than predator detection, such capacities and protocols are doubly important. Indeed, Sterelny (2003) argues that evolution has designed our cognitive economy so that social guards are central. But the deployment of social guards seems to increase the demands on our cognitive economy, increasing attention and, thereby, endangering automatic endorsement. Thus, we seem to have a tension between reliability and endorsability: to retain reliability, extended resources will be subject to checking, but to count as mental, such resources need to be (almost) automatically endorsed.

[4] There are a number of circumstances in which this will not be the case and where honesty can be assumed. The most obvious candidates, according to Sterelny (2003, pp 21–22), are male sexual advertisements. Males cannot afford not to advertise and in many species only those adverts that are expensive will attract females. Consider, for example, the peacock's tail. Such sexual advertisements are not only attractive to females, but also they are highly visible to predators. Peahens can thus rely upon the validity of these signals, as they are extremely costly to the peacock.

Non-symbolic and extended social guards

Saving the substantive conclusion of the extended mind thesis from Sterelny's attack entails resolving the tension between reliability and endorsability. This in turn demands establishing the fact that deploying social guards does not demand increased cognitive attention. I will be arguing that: (1) if the extended system is supported by a non-symbolic inner economy, then the cognitive guards may well be a natural by-product of the system requiring no extra cognitive attention; and (2) agents can (in principle) employ a number of extended mechanisms to guard against deception.

Non-symbolic social guards

Sterelny (2003, 2005) has established that employing extended resources depends on a robust representational economy that includes meta-representational abilities. But nothing claimed by Sterelny requires the representational economy to be symbolic in the sense entailed by classical artificial intelligence (AI). Classical AI models of cognition are symbolic in the precise sense that the items that encode content are identical to the items over which computational processes are defined. This coincidence of computational rules and conceptual content implies a formal account of cognition in which intelligent behaviour must result from the application of rules that apply to symbols because of their syntactic form. Sterelny's general non-nativist and anti-modular account of cognition does not fit well with such an account of cognitive life. A highly distributed connectionist network seems far more amenable to Sterelny's (2003) account of the mind. I have recently demonstrated that such a representational economy can perform cheater-detection without the need for a specialised module devoted to the task (Parsell 2005). Parallel reasoning suggests that social guards may come for free in such networks: that social guards may be deployed without increased cognitive attention. Here I will briefly rehearse the argument of Parsell (2005) and attempt to establish that the implicated networks will automatically deploy social guards.

Parsell (2005) responds to the SCT interpretation of the Wason selection task. The task, developed by Wason (1966, 1968), tests the ability to reason with the 'if p then q' locution. Extensive study has demonstrated that people are not terribly good at reasoning with this locution. But performance significantly improves when the task is framed in social terms, when it involves assessing altruism and cheater-detection (see Brown and Moore 2000; Frydman et al. (1999) for recent examples). Cosmides has used this preference for social presentation as grounds for SCT (see Cosmides 1989). The claim is that to explain the preference for social presentations of the task 'one is forced to invoke content-specified inferential machinery, including social contract algorithms' (Fiddick et al. 2000; p 5). This interpretation is the most plausible if the underlying reasoning system is symbolic (although, see Fodor 2000 for a different interpretation).

Symbolic architectures produce intelligent behaviour by the rule-governed manipulation of symbols that are non-semantically individuated. Reasoning is formal in that inferential rules apply according to the physically instantiated properties of the symbol structures (namely, their syntactic form). To enable meaningful inference syntax mirrors semantics. Thus it is possible to instantiate syntactic processes that respect content without appeal to content. On this model of intelligent behaviour, systematic variation in inference due to content cannot result from systematic variation in the inferential process, as the process of inference is insensitive to content. Thus it seems necessary *on the assumption of formal processing* to suppose that inference is modular. But this is not the only possible explanation of the Wason data. If the inferential system is non-symbolic, then systemic variation in reasoning abilities can occur within a single inferential system.

Highly distributed connectionist nets are non-symbolic. They display context-sensitivity at the input, hidden and output layers. Hence inference is also content-sensitive, as it is the total input, together with the learning environment and exact computational structure, which provides the content to the network. As such, content-sensitive inference, even within the one module (network), is to be expected. Thus, in principle, recognising a situation as a social exchange can impact the interpretation of the conditional statement and consequently lead to improved inferential performance. As such, the preference for social presentations of the Wason task is explicable without needing to posit a separate inferential module for social exchange. Leighton and Dawson's (2001) connectionist model of the Wason task clearly demonstrates the critical importance of context: the mediating factor in producing the correct response to the selection task is highly susceptible to contextual effects, indicating that the specific context is likely to be the major determinant of selection task performance. Highly distributed systems that employ novel recursive representations, such as recursive auto-associative memory (RAAM; developed by Pollack 1990) architectures, seem especially promising models of inferential tasks requiring such context-sensitive processing. Thus, there is every reason to believe that cheater-detection can be served by non-modular connectionist nets.

Connectionist computational processing is essentially parallel in nature. In highly-distributed nets there is no modularisation of particular beliefs or pieces of information. The representational burden is superimposed across the entire computational resource base of the model, such that each node (computational primitive) contributes to multiple representational items and each representational item is distributed across multiple nodes. This entails that it is impossible to

isolate the impact of any particular belief or piece of information on the network's processing.[5] Indeed, all information is in a real sense brought to bear in all processes of reasoning. Thus, if a net contains social guards, these will be naturally deployed whenever the network encounters social exchanges. Thus if our interactions with the world are supported by a cognitive system that is not modularised form social cheater-detection, the deployment of social guards will not imply an increase in cognitive attention. As such, automatic endorsement will not be jeopardised. Forgoing all assumes that the underlying (non-modular) distributed system contains social guards. But if Sterelny (2003) is correct in his account of our evolutionary history, then social guards will be central features of human cognitive life.[6]

In sum, evolution has made the development of social guards central to our success as a species. Thus we can expect any successfully socialised human to have developed protocols to guard against social deception. If the system that underlies our cognitive abilities is a highly distributed net which is not modularised for social exchange, then deploying such guards will be automatic. Thus social guarding will not increase cognitive load and endorsability will not be endangered. Of course, it remains possible, though highly unlikely to my mind, that we are symbolic engines. In which case social guarding will not come for free, but will require, if SCT is true, the activation of a system devoted to social guarding or, even if SCT is not true, the activation of symbolic protocols that will by their nature imply increased cognitive costs. Still, even if this is the case, it is possible to deploy social guards without increasing cognitive costs if the guards can be off-loaded to the environment.

Off-loaded social guards

Sterelny (2005) admits the possibility of off-loading social guards. He accepts, for example, that diary entries can be guarded against insertions by the recognition of one's own handwriting. Similarly, it is also possible to employ a code or from of encryption that is opaque to others, such that it will be impossible to manipulate the information contained in the resource without this being transparent. More sophisticated resources allow more sophisticated guarding. Consider, for example, mobile phones that typically include password protection. Here the possibility of the resource being corrupted is essentially removed because even if another should gain access to one's phone they will not have the ability to access to the information contained within the phone. The same is true of a broad range of information and communication technology (ICT) resources, such as digital personal organisers, laptop computers and Internet-based information accounts. When corruption is made more or less impossible by such protection, one need not worry that the information stored therein will be contaminated and, thus, one need not employ inner social guards to protect against such corruption.

Even this off-loading of social guards by coding or password protection does not stop outright theft. The complete deletion of inner resources does not seem possible. Someone can steal my phone, laptop or diary, but not my brain. Thus it may be argued that the full-blown stealing of the external medium is a relevant difference. Even this extreme form of theft can be protected against. If there are multiple copies of the resource, then even if one copy is stolen the information will remain available to the legitimate owner. Most mobile phones and computers offer this possibility and there are a variety of programs that will automatically perform such backing-up with literally no effort on the part of the owner of the resource. Still, a significant natural disaster could destroy all copies of one's data if they reside in a localised geographical area. Backing-up of data to a web-based server removes even this remote possibility of loss. This form of data-protection has significant benefits: (1) being non-physical it is immune to theft; and (2) in-built redundancy of the Internet—with information existing in multiple locations throughout the entire world. When the information is destroyed, it would lead to a global catastrophe. Such a catastrophe not only seems so remote as to warrant none, and hence no increase in cognitive attention, but also it would no doubt severely impact one's biological mental states by most probably causing one's death.[7] All of the above, Sterelny may grant and still claim that I have failed to get to the heart of his critique: that essentially social resources like shared note books and work spaces cannot be guarded against in the same uncostly manner. To respond to this charge, I will focus on the extreme case of extending my mind into another's brain. What I have to say here is more speculative than anything to this point.

[5] See Ramsey et al. (1991) for greater detail. Note that Ramsey et al. (1991) use this to support a connectionist account of folk psychological eliminativism, so the use of the term "belief" above would be impermissible for them.

[6] This need not imply that social guards are innate. Indeed, Sterelny's (2003) account of non-genetic inheritance shows how even such a central cognitive concern as cheater-detection can be acquired by individual in a single life cycle. Further, Rogers and McClelland (2004) have recently provided significant support for supposing even more fundamental belief structures, like those that underlie all reasoning about physical objects, which can be acquired in a connectionist net without prior biasing. Thus, a nonnativist account of the development of social guards seems at least possible. Nevertheless, nothing rules out an innate distributed account of social guards. Thus, the present defence of the extended mind can remain silent on the issue of innateness.

[7] The obvious objection here is to question portability. To worry, that is, that such information will not be contactable when needed. Clark and Chalmers (1998) voice this concern. Third generation mobile phones, however, make the Internet contactable whenever the device is available.

Extending mental states across individuals is where Sterelny's case is the strongest. Once other social actors are recognised as (at least) potentially deceitful, extending one's mind into another's brain becomes an extremely risky business that surely demands deploying social guards. But few of the guards I have so far highlighted would seem applicable here. Certainly one cannot rely upon passwords or handwriting. I will very briefly sketch one possible line of reply to this, namely, that one can rely on an established trust relation.

Trust has been characterised in a range of ways in the literature. Its defining characteristics have been variously identified as a positive belief or expectation (Lahno 2001), an affective attitude (Jones 1996) and a form of gullibility (Fricker 1987). My use of trust is neutral between these analyses. The significant element for the extended mind thesis is that trust demands avoiding conduct that smacks of distrust. To treat others as trustworthy involves refraining from the type of cross-checking that endangers endorsability. In a relationship of reliance, it is legitimate to seek evidence to confirm the wisdom of the reliance, but in a relationship of trust performing such cross-checking, it can harm the relationship itself.

Trusting relations most obviously hold in social relations where one has a substantial history with the other, such as with partners and in well-maintained teacher–student relations. These relations can be harmed if one party feels that the other is untrustful.[8] An unwillingness to take on any degree of risk makes that relationship less, not more, secure. Making the relationship less secure entails that the reliance on the other itself becomes less secure. Thus checking behaviour threatens rather than increases reliance. Thus, rather than being in tension in a trust relationship, (almost) automatic endorsement and reliance form a self-supporting pair. Of course, a trust relation does not preclude the possibility of deceit. Indeed, despite the variety of analyses of trust, there is consensus within the literature that trust necessarily involves risk. The critical question for the extended mind thesis is whether this possibility is on par with the possibility of error in one's own mind? It seems to me that in a well-maintained trust relation, for example, the risk of deception is actually less likely than the possibility of mis-remembering.

In sum, I have argued that if social guards are non-symbolically implemented, they will not increase cognitive load. My argument paralleled my earlier case against the SCT interpretation of the Wason selection task. In addition to this, I have demonstrated that if social guards can be off-loaded, no increased cognitive attention will be associated with their deployment. This is the case even if our underlying cognitive architecture is symbolic. Finally, I have suggested that trust may allow an extension of minds across individuals.

[8] This has been most persuasively argued by Baier (1986), but also see Townley (2005). For a discussion of the perniciousness of global distrust in teacher–student relations, see Townley and Parsell (2005).

Conclusion

Intelligent minds can reduce the cognitive load on the naked brains in which they reside in a number of ways. Most obviously, extended resources can ease the memory burden on the brain—not all information needs to be biologically stored—and transform difficult tasks (perceptual, learning or cognitive) into more tractable problems. Sterelny (2003, 2005) has demonstrated that advanced intelligence is required to successfully retrieve information from the extended sphere. He has not established, however, that this retrieval necessarily increases cognitive load. Thus he has not established that (almost) automatic endorsement is endangered. Thus if the extended mind thesis is restricted to intelligent, representational creatures, Sterelny's position has not damaged the central conclusion of the extended mind thesis. He has provided us with a new evolutionary perspective, but it is a perspective amenable to, not in conflict with, the extended mind.

Acknowledgements Thanks to Andy Clark, Kim Sterelny and John Sutton for many helpful discussions on evolution and the extended mind, to Cynthia Townley for discussions on trust and to Bruce Wilson, Jane French and Catharine Abell for reading earlier versions of this paper.

References

Baier A (1986) Trust and anti-trust. Ethics 96:231–260
Blake A, Yuille A (eds) (1992) Active vision. MIT Press, Cambridge
Brown WM, Moore C (2000) Is prospective altruist-detection an evolved solution to the adaptive problem of subtle cheating in cooperative ventures? Evol Hum Be 21:25–37
Churchland PM (1988) Eliminative materialism and propositional attitudes. J Philos 78:67–90
Clark A (2005) Memento's revenge: The extended Mind, extended. In: Menary R (ed) The extended mind. Aldershot, Ashgate
Clark A, Chalmers D (1998) The extended mind. Analysis 58:7–19
Cosmides L (1989) The logic of social exchange: has natural selection shaped how human reason? Studies with the Wason selection task. Cognition 31:187–276
Cosmides L, Tooby J (1992) Cognitive adaptations for social exchange. In: Barkow J, Cosmides L, Tooby J (eds) The adapted mind. Oxford University Press, New York
Cosmides L, Tooby J (1997) Towards an evolutionary taxonomy of treatable conditions. J Abn Psych 108:453–464
Cosmides L, Tooby J (1989) Evolution psychology and the generation of culture, Part II. Case study: a computational theory of social exchange. Ethol Sociobiol 10:51–97
Dennett DC (1996) Kinds of minds: towards an understanding of consciousness. Weidenfeld and Nicolson, London
Fiddick L, Cosmides L, Tooby J (2000) No interpretation without representation: the role of domain-specific representations and inferences in the Wason selection task. Cognition 77:1–79
Fodor JA (2000) Why are we so good at catching cheaters? Cognition 75:1–79
Fricker E (1987) The epistemology of testimony. Aristotelian Soc LXI (suppl):57–84

Frydman O, Light P, Alegria J (1999) Pragmatic determinants of children's responses to the Wason selection task. Psychologia 42:59–68
Gibson JJ (1979) An ecological approach to visual perception. Houghton Mifflin, Boston
Jones K (1996) Trust as an affective attitude. Ethics 107:4–25
Lahno B (2001) On the emotional character of trust. Ethical Theory and Moral Pract 4:171–189
Leighton JP, Dawson MRW (2001) A parallel distributed processing model of Wason's selection task. J Cogn Syst Res 2:207–231
Parsell M (2005) Content-sensitive inference, modularity and the assumption of formal processing. Philos Psyc 18:45–58
Pollack J (1990) Recursive distributed representations. Artif Intell 46:77–105
Rainey HJ, Zuberbühler K, Slater PJB (2004) Hornbills can distinguish between primate alarm calls. Proc Roy Soc Biol Sci 271:755–759
Ramsey W, Stich S, Garon J (1991) Connectionism, eliminativism, and the future of folk psychology. In: Ramsey W, Rumelhart D, Stich S (eds) Philosophy and connectionist theory. Erlbaum, Hillsdale
Rogers TT, McClelland JL (2004) Semantic cognition: a parallel distributed processing approach. MIT Press, Cambridge
Shaw RE, Turvey MT (1999a) Ecological foundations of cognition. I: Symmetry and specificity of animal-environment systems. J Consci St 6:95–110
Shaw RE, Turvey MT (1999b) Ecological foundations of cognition. II: Degrees of freedom and conserved quantities in animal-environment systems. J Consci St 6:111–124
Sterelny K (2003) Thought in a hostile world. Blackwell, Malden
Sterelny K (2005) Externalism, epistemic artefacts and the extended mind. In: Schantz R (ed) The externalist challenge: new studies on cognition and intentionality. De Gruyter, Berlin
Townley C (2005) Towards a revaluation of ignorance. Hypatia 21
Townley C, Parsell M (2005) Technology and academic virtue: plagiarism through the looking glass. Ethics Inf Technol
Wason PC (1966) Reasoning. In: Foss B (ed) New horizons in psychology. Penguin, London
Wason PC (1968) Reasoning about a rule. Q J Exp P 20:273–281