Do socio-technical systems cognise?

Olle Blomberg [1]

Abstract. The view that an agent's cognitive processes sometimes include proper parts found outside the skin and skull of the agent is gaining increasing acceptance in philosophy of mind. One main empirical touchstone for this so-called *active externalism* is Edwin Hutchins' theory of *distributed cognition* (DCog). However, the connection between DCog and active externalism is far from clear. While active externalism is one component of DCog, the theory also incorporates other related claims, which active externalists may not want to take on board. DCog implies a shift away from an organism-centred cognitive science to a focus on larger socio-technical-cum-cognitive systems. In arguing for this shift, proponents of DCog seem to accept that socio-cultural systems have some form of agency apart from the agencies of the individuals inside them. I will tentatively suggest a way in which such a notion of agency can be cashed out.

1 Introduction

In "The Extended Mind" [8], Andy Clark and David Chalmers ask where the mind stops and the rest of the world begins. They argue that bits of our environment sometimes become proper parts of our cognitive processes. In other words, contrary to the received view in cognitive science, cognitive processes sometimes loop out beyond the skin and skull. This is a claim about the location and boundaries of cognition.

A closely related issue is what the "unit of analysis" should be in cognitive science. The unit of analysis is the system or set of interactions that needs to be analysed in order to reach a correct understanding of how organisms cognize and behave. Presumably, if cognitive processes extend beyond the skin and skull, so should the unit of analysis. But an extended unit of analysis may be recommended on less radical grounds too, as Robert Rupert has pointed out [27]. It is enough to claim that cognition is deeply embedded in the world — without looping out into it — in order to conclude that "we can properly understand the traditional subject's cognitive processes only by taking into account how the agent exploits the surrounding environment to carry out her cognitive work" [27, p. 395].[2] While the boundaries of the unit of analysis and the boundaries of cognition are not necessarily the same then, they are clearly connected.

The *distributed cognition* approach (henceforth DCog) is probably the approach in cognitive science that has widened the unit of analysis the most. In DCog, socio-technical systems (made up of socially organized individuals equipped with tools and technologies) are treated as cognitive systems. [3] For example, the cognitive anthropologist Edwin Hutchins, in what is arguably the canonical account of DCog [17], provides a detailed ethnography of a navigation team steering a large military vessel into port. He analyses the navigation team as a cognitive system in which mental-cum-cultural representations are created, transformed and propagated.

While Hutchins' work is one of the main empirical touchstones of the philosophical extended mind movement, the relation between DCog and the philosophy that has drawn on it is unclear. This is partly due to the promiscuous use of the 'distributed cognition' label as more or less synonymous with 'the extended mind', 'active externalism', 'vehicle externalism', 'locational externalism' etc. [8, 15, 31]. In this paper, I will use Clark and Chalmers' label *active externalism* to refer to these philosophical positions collectively. Clark and Chalmers [8] refer to Hutchins' research as an example of empirical work that "reflects" their active externalism. Other active externalists (such as Susan Hurley and Robert Wilson) as well as "active internalists" (such as Fred Adams, Ken Aizawa and Robert Rupert) also refer to DCog as a sort of empirical counterpart of active externalism [15, 31, 2, 27].

There is a clear focus on "socially distributed cognition" in the DCog literature. This is a phenomenon that is largely absent in discussions about active externalism. Clark and Chalmers' [8] mention the possibility of socially *extended* cognition, where one thinker's mental state is partly constituted by the state of another thinker, but in such a case, the cognitive system is still firmly centred on the brain of an individual human organism. However, the socio-technical systems that are typically studied using the DCog framework are *not* centred on an individual organism. DCog thus departs from Clark's "organism-centered" [7, p. 139] active externalism, in the sense that there is often no clear locus of control which can be attributed to any one organism (but not in the sense that there may be no organisms involved at all).[4]

I will not enter the debate between active externalists and active internalists. My aim is rather to clarify what the relation is between (Hutchins' version of) DCog and active externalism.[5] In the next section, I present the DCog approach and tease out four theoretical-philosophical claims that make up the approach's theoretical backbone. One of these claims is tantamount to a commitment to active externalism. In the following sections, I consider whether some of the arguments that have been used to support active externalism can

[1] University of Edinburgh, United Kingdom, email: K.J.O.Blomberg@sms.ed.ac.uk

[2] Rupert also makes use of the concept 'unit of analysis', but perhaps in a slightly different way. He characterises active externalists as claiming that "the unit of analysis should be the organism and certain aspects of its environment treated together, as a single, unified system." [27, p. 395] My concept 'unit of analysis' is intended to be separate from 'cognitive system' so that even an active internalist can claim that "the unit of analysis should be the organism and certain aspects of its environment", although she would reject that they in the end should be treated "together, as a single unified system."

[3] Such systems are sometimes also referred to as 'socio-cultural systems' or 'distributed cognitive systems' in the DCog literature.

[4] However, Christine Halverson, a former student of Hutchins, states that "*DCog focuses on the socio-technical system*, which usually (but not necessarily) includes individuals." [11]

[5] Rupert [27, pp. 391–2, 425n59] also briefly discusses the relation between DCog and active externalism.

be used to support a widening of the unit of analysis to cover socio-technical systems. I argue that this is doubtful and, considered as theory of human cognition, DCog seems to rest on a contentious claim about socio-technical systems having a form of agency.

2 The distributed cognition approach

DCog grew out of ethnographic studies of people interacting with each other and with various tools in organisational settings. Such socio-technical systems are conceptualised through the theoretical lens of DCog as both computational and cognitive. In Hutchins' analysis of naval navigation, the activity of the navigation team is described using the symbol-shuffling framework of traditional cognitive science, but applied to a unit of analysis that includes not only several mariners, but also various representational artifacts.[6]

To give some flavour of research informed by DCog, I here provide a brief summary of Hutchins' analysis of a type of navigation activity. When Hutchins did his fieldwork, a navy ship that was near land and in restricted waters had to have its position plotted on the chart (map) at intervals of a few minutes. In such situations, a team of about five people had to be involved in "the fix cycle". To fix a ship's position, two lines of sight from the ship to known visual landmarks have to be drawn on the chart (the ship should be where the lines intersect on the chart). Simplifying slightly, the fix cycle ran as follows: with the help of special telescopic sighting devices called alidades, two "bearing takers" determined the bearing (direction) of one landmark each; they reported the bearings over a telephone circuit to a "bearing timer-recorder" who jotted them down in the bearing log; the "plotter", standing beside the bearing time-recorder, then plotted the lines of sight on the chart to determine the ship's position.

Hutchins glosses this fix cycle in a computational framework drawn from traditional cognitive science:

> The task of the navigation team [...] is to propagate information about the directional relationships between the ship and known landmarks across a set of technological systems until it is represented on the chart. Between the situation of the ship in the world and the plotted position on the chart lies a bridge of technological devices. Each device (alidade, phone circuit, bearing log etc.) supports a representational state, and each state is a transformation of the previous one. Each transformation is a trivial task for the person who performs it, but, placed in the proper order, these trivial transformations constitute the computation of the ship's position. [16, pp. 206–7]

From a DCog perspective, the members of the navigation team together with their tools and social organisation make up a *cognitive system* that keeps the ship on track. It seems appropriate to think of the navigation activity as instantiating a form of computation, but why think of the distributed computational process as a *cognitive* process? Are all computational processes cognitive? Or just those that are somehow hooked up in the right way to a biological organism? I do not question that it may be fruitful to conceptualise and study a socio-technical system as computational systems for various reasons. But is it fruitful for cognitive science to adopt the socio-technical system as a unit of analysis? Will this increase our understanding of the nature and manifestation of human *cognition*? Why should these systems be studied by *cognitive science* rather than, say, social science?

Later on, I will engage with these questions. But first, I present four distinct theoretical-philosophical claims that are part of DCog and relate them to the current philosophical debate about active externalism.

2.1 Into the wild

Empirical research informed by DCog has primarily been descriptive and based on ethnographic observation. According to Hutchins, much research on cognition "in the lab" (arguably a highly atypical setting for human cognition) is based on unexamined assumptions about what a human mind is for. One of the supposed pay-offs of ethnographic studies of cognition "in the wild" is to expose these assumptions and provide "a refinement of a functional specification for the human cognitive system" [17, p. 371]. According to Hutchins, cognitive science needs to get a richer empirically grounded conception of its explananda.

Hutchins calls the approach he recommends *cognitive ethnography*. A cognitive ethnography is a description of a "cognitive task world" of some specific setting. Hutchins claims that we in fact know very little about such everyday cognitive task worlds since "our folk and professional models of cognitive performance do not match what appears when cognition in the wild is examined carefully." [17, p. 371] One systematic such mismatch that cognitive science is suffering from, according to Hutchins, consists in mistaking the cognitive properties of socio-technical systems for cognitive properties of individuals considered in isolation [17, p. 355]. Hutchins argues that recognition of this mistake should lead one to suspect that the performance of cognitive tasks such as navigation "requires internal representation of much less of the environment than traditional cognitive science would have led us to expect." [17, p. 132]

In sum, DCog incorporates a methodological commitment that I call the ETHNOGRAPHY claim:

> **ETHNOGRAPHY**: Cognitive science is operating with an inadequate functional specification of the mind. Ethnographic descriptions of cognitive activities in the wild can provide a better specification for cognitive science in the lab to work with.

Note that this claim in itself does not touch on the issue of where cognitive processes are to be found, it merely points out there is a gap in our knowledge about the range, variety, and constitution of everyday activities in which cognitive processes are somehow involved.

2.2 Computation in socio-technical systems

While DCog departs from traditional cognitive science in many ways, its core, the computational model of mind, is retained. Computation is broadly conceived as the "creation, transformation, and propagation of representational states" so that it can be applied both to what happens inside and outside the heads of individuals [17, pp. xvi, 49]. Hutchins actually argues that while the notion of computation as symbol manipulation was metaphorically applied to the individual mind (in the head), it is a *literal* description of what occurs within (some?) socio-technical systems [17, pp. 363–4].

[6] Other settings studied under the auspice of DCog include the cockpit of a commercial airliner [18], a telephone hotline group [1], software programming teamwork [9], a neuroscience laboratory [3], work practice in an engineering company [26], and trauma resuscitation teamwork [29] In Hutchins' terminology, these are all cases that exemplify social distribution of cognition. While Hutchins usually presents DCog as a theory about the nature of human cognition [17, 20], it should be noted that DCog is also used as an analytical framework in Human-Computer Interaction (HCI) and in Computer-Supported Cooperative Work (CSCW) [33, 14, 11]. Interestingly, Clark [7, p. 96] actually takes HCI to be a field that house "nascent forms" of a science of the extended mind.

> [T]he computation observed in the activity of the larger system can be described in the way that cognition had been traditionally described that is, as computation realized through the creation, transformation, and propagation of representational states. [17, p. 49]

It is a bit unclear whether Hutchins believes DCog to be a theory of socio-technical systems in general or only of a symbol-shuffling subset of them. In [17, p. 363] and [19, p. 67] Hutchins sometimes suggests that it is a framework restricted for describing a subset of systems, but in later writings DCog "refers to a perspective on all of cognition, rather than a particular kind of cognition" [14, p. 3] (see also [20, p. 376]).[7]

We thus get the COMPUTATION claim:

> **COMPUTATION**: (i) A socio-technical system is a computational system, in which "representational states are created, transformed and propagated", and (ii) cognitive science should take it as a unit of analysis.

I take COMPUTATION to constitute the core of DCog. Its first part, (i), sets DCog apart from other socio-cultural approaches to cognition, while the second part (ii) sets it apart from traditional internalist cognitive science. Note that the claim in (ii) is not that cognitive science should *exclusively* take socio-technical systems as its unit of analysis.

2.3 Crossing old boundaries

Hutchins typically does not merely construe socio-technical systems as computational systems, but also as cognitive systems.[8] In calling socio-technical systems "cognitive", Hutchins seems to accept *something like* Clark and Chalmers' Parity Principle.[9] It is the functional-computational contributions of a process that makes it cognitive, not whether it occurs on one side or the other of a skin or skull boundary. In an article co-authored with James Hollan and David Kirsh, he writes:

> Distributed cognition looks for cognitive processes, wherever they may occur, on the basis of the functional relationships of elements that participate together in the process. A process is not cognitive simply because it happens in a brain, nor is a process noncognitive simply because it happens in the interactions among many brains. [14, p. 175]

This in itself need not imply that the boundaries of the cognition of individuals need to be redrawn. One can imagine several brain-bound cognitive agents interacting in such a way, with each other and with their tools, that they collectively make up a larger cognitive system (so that there are several brainbound cognitive systems nested within a larger one). However, Hutchins argues that this is the wrong picture. One important advantage of having a single framework for describing both what goes on inside and outside the heads of individuals, Hutchins argues, is that this highlights that "the normally assumed boundaries of the individual are not the boundaries of the unit described by *steep gradients in the density of interaction among media*." [17, p. 157, my emphasis] He claims he has "developed a language of description of cognitive events that is unaffected by movement across old boundaries." [19, p. 65]

As I interpret Hutchins, the location of these steep gradients determines where the boundaries of cognitive systems. The criterion can be used to analyse the relevant boundaries of socio-technical systems, as well as individuals working with tools. For example, in his analysis of how a bearing taker finds a specific landmark to read and report the bearing, the system is not restricted by the bearing taker's biological boundaries. Instead, it includes, at one point, the degree scale and the tick hairline presented to the bearing taker as he aligns the alidade with the landmark, and it then shifts as activity progresses: "The active functional system thus changes as the task changes. A sequence of tasks will involve a sequence of functional systems, each composed of a set of representational media." [17, p. 157]

Hutchins' criterion for determining system boundaries is similar to John Haugeland's proposed *bandwidth criterion* for deciding whether the mind is a distinct component in a brain-body-world system [12]. Following Herbert Simon [30], Haugeland suggests that systems should be decomposed according the pathways and bandwidth of information flow in the systems. A system is made up of components that interact with each other over interfaces. Interfaces are points of well-defined low-bandwidth interaction between components. Components are made up of parts that interact at much higher bandwidth and in ill-defined ways (relative to the interaction that is mediated by the components interfaces). If a system can be analysed as made up of components and interfaces in this way, then the system's behaviour can be made more intelligible. However, a mind is not a component that can be partitioned off from the world in this way according to Haugeland.[10]

Hutchins, I take it, clearly embraces some form of active externalism. DCog thus incorporates what I will call the EXTENDED claim:

> **EXTENDED**: Cognitive processes are not bound by the skin and skull of an individual but may loop out and include bits of the environment as proper parts.

Note that EXTENDED is different from the first part (i) of COMPUTATION. Active internalists can certainly accept that (some) socio-technical systems are computational systems. Adams and Aizawa, for example, argue that DCog is best seen as a theory of "naturally occurring computation" rather than of cognition, on the ground that processes that exhibit the "mark of the cognitive", all occur the brains of people "as a matter of contingent empirical fact". [2, pp. 46, 59]. Rupert takes a similar stance: socio-technical systems may "act as computational systems, of a sort" but there is no explanatory benefits of treating them as cognitive systems [27, p. 392]. One can of course also accept EXTENDED without accepting the socio-technical systems are computational systems.

[7] As the HCI researchers Victor Kaptelinin and Bonnie Nardi [22, p. 205] have noted, DCog seems to be suited for studying certain highly structured socio-technical systems which some kind of overarching system-level goal can be attributed to. Without such system-level goals it becomes difficult to interpret system activity as a form of problem solving.

[8] I write "typically" since occasionally, Hutchins uses the term 'functional system' instead of 'cognitive system'. His broad conception of computation certainly leaves room for important differences between internal and external computational processes. Hutchins can thus argue that even if what happens inside an individual's head is not a component according to the bandwidth criterion, internal and external processes might be different in such a way that only internal processes ought to be called "cognitive".

[9] The Parity Principle: "If, as we confront some task, a part of the world functions as a process which, *were it done in the head*, we would have no hesitation in recognizing as part of the cognitive process, then that part of the world *is* (so we claim) part of the cognitive process." [8, p. 8]

[10] Note that Haugeland is only using the bandwidth criterion negatively to argue that the mind cannot be partitioned off from the body and the world. He is not using it to partition off some other component (made up of bits of brain, body and world), which could be identified with the mind.

2.4 Socio-technical systems and agency

DCog seems to incorporate yet another claim, which takes it even further away from traditional brainbound cognitive science. Hutchins claims that a social-technical system considered as a whole can have cognitive properties of its own. In discussing the navigation team and its tools as a cognitive system, Hutchins [17] attributes several cognitive capacities to this system, such as perception (p. 182), error-detection (p. 182), self-reflection (see p. 182), remembering (p. 196), and confirmation bias (p. 239). Hollan and Hutchins claim that "[f]rom a distributed cognition perspective, goals may be properties of institutions, but need not necessarily be properties of individuals." [21]

I will call this claim, which is independent of EXTENDED, the AGENCY claim:

> **AGENCY**: A socio-technical system can have a form of agency and be the locus of cognitive capacities such as memory, perception and reasoning.

Many will probably take AGENCY to be highly counterintuitive. Given this, should not AGENCY be read metaphorically, as a claim that it might be *fruitful* to view a socio-technical system as *sort of* an agent? To treat a socio-technical system as an agent, it could be argued, is no more misleading than to treat a subsystem in the brain as an intentional system (one that, say, "interprets" incoming information from other neural subsystems).[11] In this vein, Mark Perry suggests that DCog should be seen as a "representational tool for systems analysis, and not as a true description of activity" and system boundaries should be taken as "artificially defined" [25]. Hutchins is not entirely consistent on this issue, but when takes up the issue of whether mentalistic terms such as 'remembering' are only metaphorically applied to socio-technical systems, he argues that they are not [17, pp. 363–4].

Despite the fact that Hutchins [17] is frequently cited in the extended mind debate, only Rupert [27, 28] and Wilson [32] seem to have picked up on the fact that AGENCY is part of DCog.

3 Outline of the argument

COMPUTATION, which is the core of DCog, suggests a radical reorientation in cognitive science. To include the workings of socio-technical systems among the explananda of cognitive science would amount to a significant widening of the discipline's scope. The second part (ii) of COMPUTATION is therefore in need of some kind of defence. While the boundaries of the unit of analysis need not be restricted to the boundaries of cognition, the relevance of the workings of socio-technical systems for our understanding of cognition needs to be argued for or demonstrated in some way.

There seem to be two routes that proponents of DCog can take to defend COMPUTATION, a direct route or an indirect one. The computational processes of a socio-technical system must either themselves be cognitive (the direct route), or else it must be the case that the unit of analysis needed to make sense of the cognitive processes has to be widened to cover the socio-technical system in which the processes are embedded.

I will argue that EXTENDED cannot help establish COMPUTATION, at least not when EXTENDED is arrived at by appeal to the bandwidth criterion. While EXTENDED can be used to motivate the study of socially *extended* cognition, it cannot, or so I will argue, justify treating whole socio-technical systems as cognitive systems. COMPUTATION therefore needs some other (or further) supporting consideration. I will therefore argue that COMPUTATION depends on AGENCY being true. If AGENCY is accepted, then the claim that cognitive science should study socio-technical systems — the second part (ii) of COMPUTATION — follows naturally.

Should AGENCY be accepted? I will not give an answer to this question, but I will argue that the principles that may lead one to accept EXTENDED cannot be straightforwardly carried over into an argument in support of AGENCY. Towards the end of the paper, I will suggest one way in which AGENCY at least can be made intelligible.

4 From EXTENDED to COMPUTATION

Before considering the direct and the indirect route from EXTENDED to COMPUTATION, I want to briefly consider whether COMPUTATION can be established without the means of EXTENDED or AGENCY.

4.1 Embedded cognition

In an attempt to deflate active externalism, or what he calls *the hypothesis of extended cognition* (HEC), Rupert argues that all the empirical results and observations that active externalists appeal to in order to defend their position can be accounted for equally well (or better) by a *hypothesis of embedded cognition* (HEMC).

> **HEMC**: "[C]ognitive processes depend very heavily, in hitherto unexpected ways, on organismically external props and devices and on the structure of the external environment in which cognition takes place." [27, p. 393]

HEMC indirectly takes the study of how organisms interact with tools and their immediate environment into the purview of cognitive science, although, what cognitive science should ultimately explain is the (internal) cognitive processes. An example will be helpful here. In an ethnographic study of cooperative work in the control room of a London Underground line, the sociologists Christian Heath and Paul Luff [13] note how the two personnel in the control room constantly peripherally monitor each other's activities and design their actions not only to achieve the action's primary goal but also to communicate to each other what they are doing.[12] Such multi-tasking and mutual coordination is ubiquitous in all kinds of settings. Yet, how people manage to do this is hardly something that cognitive science has advanced our understanding of very far.

Such mundane but overlooked patterns of interaction are important phenomena that cognitive science arguably ought to investigate. What cognitive abilities and capacities enable people to smoothly engage in such temporally fine-grained social interaction and monitoring? This example certainly suggests that it may be fruitful for cognitive scientists to pay more attention to what is going on in parts of sociology. However, it seems to me to fall short of making the case that the information flow and "behaviour" of the control room system should be taken as a unit of analysis in cognitive science.

4.2 The direct route

Perhaps the bandwidth criterion (which I take Hutchins to be endorsing) can be used to establish COMPUTATION. According to the

[11] Of course, some think that such explanations are very much misleading [4].

[12] Heath and Luff's study was not informed by DCog, but by ethnomethodology, a theoretical framework in microsociology.

bandwidth criterion, if a socio-technical system is not decomposable into components, which interact through relatively well-defined interfaces, but itself interacts with its environment through such interfaces, then cognitive science ought to, it seems, treat that whole socio-technical system as an explanandum.

In the case of Hutchins' navigation team, this would be plausible if the visual input of the landmarks as presented in the alidade and the auditory output of commands to change are low-bandwidth interaction when compared to the interaction happening inside the system. However, this does not seem to be the case. There are clearly some well-defined low-bandwidth interfaces inside the system. For example, the communication of landmark and bearing information over the telephone circuit between bearing takers and the bearing time-recorder and plotter is clearly a low-bandwidth and well-defined one. In addition, the low-bandwidth interaction of the system as a whole with the wider world of the sea is probably a special feature of this particular socio-technical system (Hutchins does not, I think, claim that *all* socio-technical systems are fruitfully taken as objects of study in cognitive science, but if COMPUTATION turns out to be true only of a very small set of socio-technical systems, then the claim is considerably less interesting).

The bandwidth criterion, it should be made clear, is not the only criterion for determining the boundaries of cognition that has been proposed by active externalists. Andy Clark, for example, rejects the bandwidth criterion as a criterion for determining cognitive system boundaries [7, pp. 156–9]. The existence of genuine interfaces between the brain/body and the world does not, he argues, threaten the claim that cases of genuine cognitive extension are fairly common. What is important is instead that people's cognitive performance often results from "rich temporal integration" of internal and external processes and events [7, sect. 2.6, 4.7]. In such cases, the "fine structure [of internal processes and events] has been selected (by learning and practice) so as to *assume* the easy availability of such and such information" from the external world [7, p. 74]. The emergence of such "subpersonal interweaving" [7, p. 240n11] of internal and external threads is (sometimes?) reflected in personal-level experience, such as when a tool or some other bit of the world become "transparent equipment through which you confront a wider world." [7, p. 74]

Clearly, the proponent of DCog cannot rely on personal-level phenomenology to argue for the second part (ii) of COMPUTATION (unless they are willing to claim that socio-technical systems have experiences). What about the subpersonal-level phenomena of rich temporal integration and interweaving? It certainly seems possible in principle that a whole socio-technical system or practice may emergence in such a way that the all the processes that occur in the system are highly dependent on each other and their organisation. Perhaps Hutchins' navigation team and Heath and Luff's control room are actually examples of such systems. The question is if such a subpersonal (or should it be intersubpersonal?) organisation counts for anything in the absence of personal-level phenomena (superpersonal-level phenomena?).

4.3 The indirect route

The indirect route from EXTENDED to COMPUTATION is analogous with the way in which HEMC leads to the adoption of a larger unit of cognitive analysis (without extending the boundaries of cognition). Assuming that EXTENDED is true, might it not be the case that the extended cognitive processes are deeply dependent on environment of the extended cognitive system. Adapting HEMC somewhat, the proponent of DCog might appeal to the following *hypothesis of embedded extended cognition* (HEMEC):

> **HEMEC**: Extended cognitive processes depend very heavily, in hitherto unexpected ways, on props and devices external to the *extended* cognitive system and on the structure of the *wider* environment in which the extended cognition takes place.

If we assume that the dependency that HEC (EXTENDED), HEMC and HEMEC are concerned with is understood in terms of bandwidth (so that two components are heavily dependent on each other just in case they are coupled in high-bandwidth interaction), then it becomes difficult to argue for HEMEC. If one has already accepted EXTENDED on "bandwidth profile grounds", then all the props and devices that are coupled with an agent in high-bandwidth interaction will already be part of that (extended) agent. HEMEC will therefore not help extend the unit of analysis further. Perhaps there is some other (better) way to unpack dependency without relying on bandwidth profiles, which could justify a further widening of the unit of analysis. As I have mentioned, it is possible to argue for active externalism in other ways than by relying on the bandwidth criterion.

5 From AGENCY to COMPUTATION

To motivate the inclusion of socio-technical systems among the explananda of cognitive science, some notion of group agency seems to have to be made cogent. If some socio-technical systems are agents, then it seems plausible that the computational processes in these systems should be thought of as their cognitive processes. Admittedly, this looks like putting the cart before the horse, since AGENCY is arguably in as much need of justification as COMPUTATION. However, I think looking at the relation between COMPUTATION and AGENCY may throw some light on what would be needed in order to show that they are true.

5.1 Subsystemic representations

Arguments for the existence of group agency, or socio-technical system agency, usually appeal to the explanatory benefits of treating groups or socio-technical systems as agents (see [28]). However, as critics are quick to point out, it seems that the behaviour of groups or socio-technical systems — their "agency" — can be reductively explained by appeal to the behaviour of the people that participate in the system and how they communicate among themselves. For example, one can argue that while a whole navigation team is needed to correctly plot the passage of a ship, the knowledge of the ship's position is found in the head of the plotter, never literally on the navigation chart or diffused in the team and its tools. Similarly, while the organisation of the team must be considered when making sense of the actions of its members, it is redundant to attribute agency to the organisation itself.

Rupert [28] argues, correctly in my view, that to make the case of what he calls "group cognitive systems", it must minimally be shown that the representations used in/by such systems are *mental* representations, not merely cultural/conventional representations that sometimes prompt mental representations in the minds of individual group members. Rupert then argues that according to a number of well-known theories of mental representations, the cultural/conventional representations that are propagated in group cognitive systems fail to count as mental representations.

In cognitive science and the philosophy of mind, one commonly distinguishes between the personal-level of explanation and the subpersonal-level of explanation. Folk psychological accounts of human conduct, often couched in terms of the beliefs and desires, are examples of person-level explanations. Computational and information-processing models in cognitive psychology, on the other hand, are examples of subpersonal-level explanations. I propose that we make a similar distinction when discussing socio-technical systems. Systemic explanations refer to the "behaviour" of the entire system, in terms of its goal for example (e.g. "navigation into port"), while subsystemic explanations refer to the computational processes that occur in the system.

Rupert presupposes that all representations in a group cognitive system are personal-level representations. Now, this is a plausible presupposition, and as far as I know, it is shared by most philosophers who have defended some group agency thesis. Moreover, when proponents of DCog are out on the field, they are supposed to trace the trajectories and transformations of personal-level representations. However, these representations are ultimately of interest in virtue of their functional roles in the socio-technical system they are trying to understand, in virtue of them being *subsystemic* representations. In DCog, public representations thus have a dual role. They are personal-level representations and — when "functionalised" — they are also subsystemic representations.[13]

Now, I tentatively propose, that one way of cashing out the idea of group (or socio-technical) system agency, is in terms of computations over subsystemic representations that are *not* personal-level representations for any member in the system. To understand the details of how such a system "behaves", a reductive explanation is unlikely to be adequate. If there exist such subsystemic representations inside a system it seems that there might be some explanatory benefit in treating the whole system as a cognitive system, as a kind of agent.

6 Discussion

Much of the previous discussion hangs on the idea that there is a proper domain of explananda for cognitive science. This explananda will consist of the behaviour of various cognitive systems, such as human beings and other animals, and more controversially, the behaviour of robots, software agents, or socio-technical systems. In this paper, when I have referred to explanatory targets or explananda, I have primarily done so by appealing to loose intuitions about what cognition is. I take it to be uncontroversial that cognition is at least primarily an activity of biological organisms, and when we want to extend our notion of cognition to other entities, we have to appeal to similarities to these paradigmatic systems.

The intuition, deep-seated in many, that socio-technical systems simply cannot be agents or cognitive systems may have its roots in the fact that socio-technical systems lack many features of biological organisms. Biological organisms are autopoietic systems, "self-producing" systems, that continuously reproduce their own internal components and boundaries. While there have been attempts to apply such concepts from biology to socio-technical systems, such attempts are I think best seen as metaphorical (see [24] for a brief discussion). The fact that it is so easy to extend notions of mind and cognition from a computational perspective, should perhaps be taken as a sign that the perspective is missing something important.

An alternative way of understanding DCog is to read it as a proposal to revise our very concept of cognition. If this is correct, then the objection that DCog does not fit our intuitions about cognition appears moot. Ronald Giere suggests that such a reading is the most charitable one. He argues against the application of everyday mentalistic notions such as 'believing' and 'remembering' to socio-technical systems but he does not find the notion of 'distributed cognition' objectionable since 'cognition' is a term used primarily by specialists: "We are thus free to develop it as a technical term of cognitive science". [10, p. 318] For Giere, a socio-technical system qualifies as a cognitive system simply by producing or outputting knowledge. The socio-technical system of the navigation team and its tools studied by Hutchins thus make up a cognitive system since it repeatedly produces a fix of the ship's position. On Giere's view, the knowledge of the position is found in the head of the plotter (and possibly one or two other persons), but not on the chart or somehow diffused in the system.

It is possible to read Hutchins [17] as proposing such a revision as well. In a way, he points out that the phenomenon that traditional brainbound cognitive science took as characteristic of cognitive processes, namely the sequential manipulation of symbols, actually manifests itself in various socio-technical systems. So if cognitive science is the science of systems that manipulate symbols or process information, then it should look elsewhere than in the heads of individuals. Such a revision of the concept of 'cognition', Hutchins can argue, allows us step inside the cognitive system and observe symbol manipulation directly [17, pp. 128–9]. I have no objection against such a revision in principle. However, one might argue that it is both arbitrary and redundant [23, 5]. After all, frameworks for studying socio-technical systems and modern organisations are already available within the social sciences.

7 Conclusions

Proponents of DCog, whose works sometimes cited as empirical work that reflect active externalism, seem to be pressed to embrace the idea that some socio-technical systems should be considered to be agents. Appeals to a bandwidth criterion for determining the boundaries of cognitive system do not establish that socio-technical systems should be taken as a unit of analysis in cognitive science. However, many active externalists do rely on the bandwidth criterion to determine the bounds of cognition, but rely on other considerations. It is possible that the bandwidth criterion is not the right one, and that a better criterion will in fact show that (many) socio-technical systems are cognitive systems after all.

Finally, I want to note that these are conclusions about DCog as a framework for studying human cognition. DCog is also widely used in Human Computer Interaction (HCI) and Computer-Supported Cooperative Work (CSCW) research. If the background motivation for our use of DCog is to understand a specific work setting and the (potential) role for information technology in it, then it seems unproblematic to focus on the propagation of representational states in the system. Many information technology systems (especially those deployed in organisational settings) are used to create, transform and propagate various representational states.

ACKNOWLEDGEMENTS

I would like to thank Andy Clark, Julian Kiverstein and Matteo Colombo, and two anonymous reviewers for helpful comments. Thanks also to Microsoft Research for financial support.

[13] Clark [6, pp. 292–3] argues, in the context of active externalism, that speech, writing and other "material symbols" play such a dual role in human cognition.

REFERENCES

[1] Mark S. Ackerman and Christine Halverson, 'Considering an organization's memory', in *Proceedings of ACM 1998 Conference on Computer Supported Cooperative Work*, pp. 39–48. ACM Press, (1998).

[2] Fred Adams and Ken Aizawa, 'The bounds of cognition', *Philosophical psychology*, **14**(1), 43–64, (2001).

[3] Morana Alac and Edwin Hutchins, 'I see what you are saying: Action as cognition in fmri brain mapping practice', *Journal of Cognition and Culture*, **4**(3-4), 629–661, (2004).

[4] Max R. Bennett and Peter M. S. Hacker, *Philosophical Foundations of Neuroscience*, Blackwell Publishing, 2003.

[5] Graham Button, 'Review of edwin hutchins' 'cognition in the wild'', *Computer Supported Cooperative Work*, **6**, 391–395, (1997).

[6] Andy Clark, 'Material symbols', *Philosophical Psychology*, **19**(3), 291–307, (2006).

[7] Andy Clark, *Supersizing The Mind: Embodiment, Action, and Cognitive Extension*, Oxford University Press, 2008.

[8] Andy Clark and David// Chalmers, 'The extended mind', *Analysis*, **58**(1), 7–19, (1998).

[9] Nick V. Flor and Edwin Hutchins, 'Analyzing distributed cognition in software teams: A case study of team programming during perfective software maintenance', in *Empirical Studies of Programmers: Fourth Workshop*, eds., Jurgen Koenemann-Belliveau, Thomas G. Moher, and Scott P. Robertson, chapter 5, 36–64, Ablex Publishing Corporation, (1991).

[10] Ronald N. Giere, 'Distributed cognition without distributed knowing', *Social Epistemology*, **21**(3), 313–320, (July–September 2007).

[11] Christine A. Halverson, 'Activity theory and distributed cognition: or what does cscw need to do with theories?', *Computer Supported Cooperative Work*, **11**(1-2), 243–267, (2002).

[12] John Haugeland, 'Mind embodied and embedded', in *Having Thought: Essays in The Metaphysics of Mind*, chapter 9, 207–237, Harvard University Press, (1998).

[13] Christian Heath and Paul Luff, 'Collaborative activity and technological design: Task coordination in london underground control rooms', in *Proceedings of Second European Conference on Computer-Supported Cooperative Work*, eds., Liam Bannon, Mike Robinson, and Kjeld Schmidt, 65–80, Kluwer, (1991).

[14] James Hollan, Edwin Hutchins, and David Kirsh, 'Distributed cognition: Toward a new foundation for human-computer interaction research', *ACM Transactions on Computer-Human Interaction*, **7**, 174–196, (2000).

[15] Susan L. Hurley, *Consciousness in Action*, Harvard University Press, 1998.

[16] Edwin Hutchins, 'The technology of team navigation', in *Intellectual Teamwork: Social and Technological Foundations of Cooperative Work*, eds., Jolane Galegher, Robert E. Kraut, and Carmen Egido, chapter 8, 191–220, Lawrence Erlbaum Associates, (1990).

[17] Edwin Hutchins, *Cognition in The Wild*, MIT Press, 1995.

[18] Edwin Hutchins, 'How a cockpit remembers its speeds', *Cognitive Science*, **19**(3), 265–288, (July 1995).

[19] Edwin Hutchins, 'Response to reviewers', *Mind, Culture, and Activity*, **3**(1), 64–68, (1996).

[20] Edwin Hutchins, 'The distributed cognition perspective on human interaction', in *Roots of Human Sociality: Culture, Cognition and Interaction*, eds., Nick J. Enfield and Stephen C. Levinson, chapter 14, 375–398, Berg Publishers, (2006).

[21] Victor Kaptelinin, Bonnie Nardi, Susanne Bodker, John Carroll, Jim Hollan, Edwin Hutchins, and Terry Winograd, 'Post-cognitivist hci: second-wave theories', in *CHI '03: CHI '03 extended abstracts on Human factors in computing systems*, pp. 692–693, New York, NY, USA, (2003). ACM.

[22] Victor Kaptelinin and Bonnie A. Nardi, *Acting with Technology: Activity Theory and Interaction Design*, MIT Press, 2006.

[23] Bruno Latour, 'Cogito ergo sumus! or psychology swept inside out by the fresh air of the upper deck… (review of ed hutchins' cognition in the wild)', *Mind, Culture, and Activity*, **3**(1), 54–63, (1996).

[24] John Mingers, 'An introduction to autopoiesis–implications and applications', *Systems Practice*, **2**(2), 159–180, (1989).

[25] Mark Perry, 'The application of individually and socially distributed cognition in workplace studies: two peas in a pod?', in *Proceedings of European Conference on Cognitive Science*, pp. 87–92, (1999).

[26] Yvonne Rogers and Judi Ellis, 'Distributed cognition: an alternative framework for analyzing and explaining collaborative working', *Journal of Information Technology*, **9**, 119–128, (1994).

[27] Robert Rupert, 'Challenges to the hypothesis of extended cognition', *Journal of Philosophy*, **101**(8), 389–428, (2004).

[28] Robert Rupert, 'Minding one's cognitive systems: When does a group of minds constitute a single cognitive unit?', *Episteme*, **1**(3), 177–188, (2005).

[29] Aleksandra Sarcevic, Ivan Marsic, Michael E. Lesk, and Randall S. Burd, 'Transactive memory in trauma resuscitation', in *CSCW '08: Proceedings of the ACM 2008 conference on Computer supported cooperative work*, New York, NY, USA, (2008). ACM Press.

[30] Herbert A. Simon, *The Sciences of the Artificial, 3rd Edition*, MIT Press, 1996.

[31] Robert A. Wilson, *Boundaries of the Mind: The Individual in the Fragile Sciences*, Cambridge University Press, 2004.

[32] Robert A. Wilson, 'Collective memory, group minds, and the extended mind thesis', *Cognitive processing*, **6**(4), 227–236, (December 2005).

[33] Peter Wright, Bob Fields, and Michael Harrison, 'Analysing human-computer interaction as distributed cognition: The resources model', *Human Computer Interaction*, **15**, 1–42, (2000).