

Distributed Cognition and Memory Research: History and Current Directions

Kourken Michaelian · John Sutton

© Springer Science+Business Media Dordrecht 2013

Abstract According to the hypotheses of distributed and extended cognition, remembering does not always occur entirely inside the brain but is often distributed across heterogeneous systems combining neural, bodily, social, and technological resources. These ideas have been intensely debated in philosophy, but the philosophical debate has often remained at some distance from relevant empirical research, while empirical memory research, in particular, has been somewhat slow to incorporate distributed/extended ideas. This situation, however, appears to be changing, as we witness an increasing level of interaction between the philosophy and the empirical research. In this editorial, we provide a high-level historical overview of the development of the debates around the hypotheses of distributed and extended cognition, as well as relevant theory and empirical research on memory, considering both the role of memory in theoretical debates around distributed/extended ideas and strands of memory research that resonate with those ideas; we emphasize recent trends towards increased interaction, including new empirical paradigms for investigating distributed memory systems. We then provide an overview of the special issue itself, drawing out a number of general implications from the contributions, and conclude by sketching promising directions for future research on distributed memory.

K. Michaelian (✉)

Bilkent Üniversitesi, Felsefe Bölümü, FA Binası 113A, Ankara 06800, Turkey
e-mail: kmichaelian@bilkent.edu.tr

J. Sutton

Department of Cognitive Science, Macquarie University, Macquarie, NSW 2109, Australia
e-mail: john.sutton@mq.edu.au

1 Distributed Cognition and Memory Research

When different people remember how to get to a particular place, or share their memories of an event they all experienced, or locate a specific and complex sequence of facts among some large body of data, they may employ significantly distinct methods in achieving their aims. While some people just immediately remember the relevant material all by themselves, others will ask for help, either relying entirely on their peers or hoping to cross-cue each other's recall; many will consult external aids, whether general-purpose technologies or idiosyncratically-organized personal systems; in turn, others may not manage the task at all until they put themselves back into the right context, or re-enact a certain procedure or sequence of actions. In ordinary thinking about memory, the notion that the processes of remembering can be thus hybrid, involving differently-balanced deployments of internal and external resources, can seem unproblematic.

From this point of view, it may seem arbitrary to restrict attention to processes within the individual brain: yet this has been the strategy of many long-standing traditions in the philosophy, psychology, and cognitive science of memory (compare Butterfill and Sebanz 2011 on individual and joint action). It has often been assumed that remembering is exclusively a neural process, occurring inside an individual's brain; that whatever the precise nature of the memory trace which carries or is the vehicle of each enduring memory content, it is contained within and possessed by that individual; that any psychological kind or kinds to be identified by the sciences of memory will be wholly internal; and that external resources are at best cues or triggers to the real memory processes inside the head.

In philosophy, the most notorious challenge to these assumptions arises from Andy Clark and David Chalmers' paper "The Extended Mind" (Clark and Chalmers 1998). They argued that both cognitive processes (such as remembering) and standing cognitive states (such as memories) can, in certain circumstances, literally spread beyond the boundaries of skull and skin. In resisting individualism about memory and cognition, Clark and Chalmers built their "active externalism" directly on an array of earlier empirical and philosophical movements, which we survey briefly below. The widespread influence of their controversial and impressive paper has occasionally led contemporary philosophers to neglect both that broader field of existing research on distributed cognition and the methodological component of their project in favour of exclusive focus on metaphysical issues. A natural methodological consequence of individualism is that psychology should be allied only with the neurosciences, that mind and memory are to be studied by way only of individual behaviour and individual brains. In contrast, Clark called for "a new kind of cognitive scientific collaboration involving neuroscience, physiology, and social, cultural, and technological studies in about equal measure" (Clark 2001a, p. 154). But his critics saw such moves as a misguided attempt "to turn psychology into a kind of anthropology or sociology or ecology; and it just won't fit. There already are sciences whose topic of inquiry is the interpersonal and environmental. . . . There is no room for an expanded psychology, no motivation for it, and no need for it" (Butler 1998, p. 222). Ongoing debates over both metaphysical and methodological tensions, and especially their

consequences for empirical research on memory, are the springboard for the eight papers in this special issue of the *Review of Philosophy and Psychology*.

1.1 What is Distributed Memory?

By remembering, we can—fallibly, imperfectly—keep track of states of the world and sequences of events, including our own actions. We can, more or less explicitly, use past experience to guide or direct present decisions and future plans. We can maintain a more or less coherent and integrated sense of our own history and of our place in social groups. While some of these complex and interacting functions of memory are, arguably, unique to humans, they build on capacities for learning and tracking which are more widely shared across species. Whatever their evolutionary history, activities of remembering in human beings are (as we noted above) notably diverse, often involving not only multiple interconnected neural systems, but also bodily practices and external resources of many distinctive kinds. Such is, at least, the view of theorists who see remembering as “distributed”, as spread over, coupled with, situated among, or incorporating heterogeneous resources beyond the brain.

Yet it is a striking feature of human memory that we sometimes remember facts, episodes, or procedures in isolation, while entirely alone, or independent of any relevant current stimuli. Our capacity thus to remember what is absent was often highlighted by early classical theorists of the cognitive revolution (Miller et al. 1960; Anderson 1976), who saw memory as demonstrating that each individual carries all the information we can recall around with us in the form of internal representations. While some theories of distributed cognition have been pitched as radical and thoroughgoing critiques of such forms of cognitivism, in other cases they have been developed more as natural extensions, or as critiques of mainstream cognitive theory from within. But in either case theories of distributed cognition have offered metaphysical and methodological challenges to standard forms of individualism and internalism: they have claimed either that remembering does not occur within the brain alone, or that it should not be studied by exclusive attention to neural processes, or both. They have therefore had to argue that even when memory appears to take us out of the current situation, in fact the processes involved can still often be distributed across disparate resources (Sutton 2009).

1.2 Distributed Cognition: Background and History

A full history of distributed cognition and memory research would recover conceptual links between three great precursors in the first half of the 20th century who addressed this challenge in different but arguably compatible ways: the English psychologist (Bartlett 1995/1932), the French social theorist (Halbwachs 1992), and the Soviet developmental psychologist (Vygotsky 1978). But the modern history of ideas about distributed memory dates from the 1970s and especially the 1980s. In that period, for independent reasons in cognitive psychology, sociology, and developmental psychology respectively, the works of these three predecessors began to be integrated afresh into mainstream debates (Sutton 2009). Bartlett’s account of

individual remembering as the context-dependent reconstruction of momentary patterns from fragmentary, interest-ridden traces and schemas was taken up by both cognitivist and ecological psychologists interested in the limits and consequences of constructive processes (Neisser 1997; Saito 2000). A multifaceted “memory boom” in social and cultural theory brought Halbwachs’ difficult ideas about the social frameworks of memory and the role of social groups in reconstructing the past back to the heart of the social sciences (Olick and Joyce 1998). And, as well as Vygotsky’s direct and ongoing influence on developmental psychology, his accounts of thinking and “self-instruction” were taken up by the “new connectionists”. In addition to their thoroughly constructivist models of memory, one influential section of the 1986 volumes on Parallel Distributed Processing linked Vygotsky to a discussion of the role of external representations in reasoning. To solve complex multiplication problems, for example, we use our trained-up, culturally-sanctioned skills to write down the sums in particular ways, manipulating and then iteratively processing environmental representations (Rumelhart et al. 1986, pp. 44–48). As well as turning problems into simpler pattern-matching or pattern-transformation tasks which we can perform unaided, such systematic practices of (external) symbol processing are the source of any capacities we develop to do mental arithmetic, or perform other complex symbolic operations, without those symbolic and environmental supports. The imagined representations or models on which we then come to be able to operate derive from our experience with actual models in the external environment, which can thus be seen as “a key extension to our mind” (Rumelhart et al. 1986, p. 46).

This connectionist movement was one of a number of independent developments in empirical, computational, and theoretical cognitive science which by the 1980s saw researchers increasingly liberalize their conceptual and methodological toolkit, as classical individualist Artificial Intelligence failed to make progress as swiftly and successfully as had been hoped (Boden 2006). Distributed cognition emerged from a range of traditions which were differently placed in relation to mainstream cognitive science. Alliances with dynamical systems theory arose, for example, within the cognitive sciences. Developmental and cultural psychologists had pushed further on Vygotskian ideas about the role of adults in helping children negotiate the zone of proximal development, until Bruner and colleagues introduced the notion of “scaffolding”. This was intended to capture the ways in which external resources (whether physical, technological, or social) can mould or orchestrate actions beyond a learner’s current level of unassisted performance (Wood et al. 1976; Greenfield 1984). Into the 1990s, this tradition merged with developmental research on domains from motor performance to reasoning couched in the framework of dynamical systems theory (Thelen and Smith 1994; Clark 1997, pp. 35–47). It also shaped a continuous body of social-interactionist work on children’s language and memory, which argued that specific qualities of parents’ talk about the past with pre-schoolers plays direct and constructive roles in sculpting their children’s autobiographical memory capacities (Nelson and Fivush 2004). Our more sophisticated capacities for thinking and remembering, on such views, soak in from the socially-available models to which our biological mechanisms are especially attuned.

Other fields which had been in some ways more marginal during the earlier phases of the cognitive revolution also contributed to a new focus, in the 1980s and early

1990s, on the rich interactions between thinkers and their environments, and on the contexts in which cognitive processes occurred. Anthropology, education, media theory, and design had been to differing degrees “missing disciplines” (Boden 2006) in the putatively multidisciplinary project, despite their relevance to the study of human thinking, remembering, problem-solving, learning, and communicating. Theorists in these domains, and in emerging hybrid disciplines such as human-computer interaction (HCI), sought to integrate the new directions in cognitive theory with the substantial and independent bodies of empirical and ethnographic research in their home fields (Agre and Chapman 1989; Clancey 2009; Lave 1988; Norman 1993; Rogoff 1991; Rogoff and Lave 1984; Rogers 1997; Salomon 1993; Suchman 1987; Zhang and Norman 1994). The development of “distributed cognition” as both theoretical approach and method was pushed ahead most fully over this period by the cognitive anthropologist Edwin Hutchins and his colleagues at the University of California, San Diego.

1.3 Hutchins and Clark

In a series of studies of practices of decision-making across domains such as navigation in both maritime and aviation settings, software development, and air traffic control, Hutchins and his colleagues analysed flexible and intelligent action as involving the cooperation and coordination of disparate individual, interpersonal, and technological activities. In more or less stable organized settings, groups and artifacts form distributed cognitive systems which are collectively responsible for the emergent overall behaviour of the whole. The cognitive scientists’ aim then is to examine the microprocesses of interaction across the diverse components of these distributed and heterogeneous cognitive systems, tracing for example the propagation and transformations of particular representational states across distinct (internal and external) media. Often no individual actor controls or even knows every part of the intelligent process of (for example) navigation, let alone possesses an internal blueprint or plan of each of its component: expertise is not a property of individual agents, but is built in to the constraints of the system.

Hutchins published a forthright and striking exposition of distributed cognition, in the context of his cognitive ethnography of navigation on a US Navy frigate, in his book *Cognition in the Wild* (Hutchins 1996). The framework was then brought sharply to the attention of philosophers of cognitive science when Andy Clark adopted it as a central plank in his rich synthesis of the new anti-individualist movements across the cognitive sciences, *Being There: putting brain, body, and world together again* (Clark 1997). Clark saw distributed cognition as a natural extension of the point made in much connectionist literature that order and systematicity in human cognition and action can derive in part from the stability of our environments, rather than as a direct product or reflection of exhaustively-specified internal recipes. Cognition might thus be multiply distributed, both within neural networks and across bodies, artifacts, and social groups. As well as enthusiastically adopting Hutchins’ cognitive anthropological approach, Clark brought together evidence and ideas from robotics, developmental psychology, dynamical systems theory, connectionist modelling, phenomenological philosophy, and from strands of work in

biology, neuroscience, linguistics, and economics in a dizzying integrative manifesto for an embodied and distributed multidisciplinary cognitive science. As in the case of Hutchins, Clark's visions of new agendas for cognitive science were nevertheless couched in terms of tracing the computational transitions across representational states, even if the transitions and states of interest were now often external. This has occasionally seen both Hutchins and Clark attacked as insufficiently radical by critics of the very idea of cognitive science, who have seen them as exporting or projecting the apparatus of internalism onto the social and public realm: but on the other hand it has arguably facilitated the gradual adoption of at least parts of the distributed cognition framework in mainstream cognitive science.

As Hutchins and Clark have each continued to develop their research programs in new directions, there have recently been signs of some divergences concerning just how and at which timescales culture and mind interact (Hutchins 2011; Clark 2011). But their earlier canonical presentations of distributed cognition have some further common features which are worth noting here. Firstly, although both Hutchins and Clark have been particularly interested in the specifically technological mediation of cognitive processes, sometimes focussing on highly sophisticated computational artifacts, they both clearly intend the framework also to apply to other dimensions of distributed cognitive systems, both noting that most complex real-world cognitive practices involve social and interpersonal interaction in conjunction with technologies and physical systems. Secondly, in some contrast to the way the philosophical literature has developed in the wake of Clark and Chalmers' defence of the extended mind, both Hutchins and Clark took up the burden of showing that the primary contribution of distributed cognition should come in the shape of the methodological difference it can make to the explanatory practices of the cognitive sciences. They each argued that the utility and productivity of the framework would be best demonstrated by its success in encouraging and assisting scientists to study changing interactions between neural, bodily, material, and social processes in driving cognition and intelligent action (Clark 2001b; Hutchins 2005).

Finally, the central route by which both Hutchins and Clark argued for distributed cognition was by stressing the disparate nature of the coordinated resources in a distributed cognitive system. Neural, bodily, physical, and social components of such complex coupled systems need not be alike: rather, their distinctive features and capacities complement each other in combination so as to realize the relevant processes collectively. In particular, where the external resources in question are symbolic in nature, the representational formats and the computational dynamics of such external symbol systems will typically differ from whatever states and processes are at work in the brain. Since both Hutchins and Clark worked within a broadly connectionist understanding of cognitive neuropsychology, they did not argue for any deep isomorphism of structure or function between neural and external resources: external resources are alien but complementary, as Clark put it, to the brain's style of storage and computation, so "the brain need not waste its time replicating such capacities. Rather, it must learn to interface with the external media in ways that maximally exploit their particular virtues" (Clark 1997, p. 220; compare Clark 1998, p. 99). To the extent that philosophical work on extended cognition has focussed primarily on issues about the parity of, or functional isomorphism between, internal and external

resources (Adams and Aizawa 2008; Rupert 2009; Sprevak 2009; Wheeler 2011), it has arguably thus shifted the central line of thought away from the original directions of argument for distributed cognition.

1.4 Memory Research and Distributed Cognition

Both Hutchins and Clark discussed remembering at many points in these works of this period, but it was not part of their projects to engage in detail, either critically or constructively, with existing psychological views on the forms and mechanisms of human memory. Research in the sciences of memory, indeed, had a quite independent history in these decades, often somewhat apart from other mainstream movements in the cognitive sciences. It was only slowly, for example, over the course of the 1980s and especially the 1990s, that autobiographical (or personal or experiential) memory became a widely acceptable topic in cognitive psychology (Conway 1990). And while debates about ecological validity in memory research had been particularly heated (see Barnier 2012), one of their effects was to forge tight links between basic experimental work on memory in the laboratory, and applied or “everyday” memory research: the influential Society for Applied Research in Memory and Cognition (SARMAC), for example, was founded in 1994 after the 3rd major international conference on Practical Aspects of Memory.

In certain respects, then, some methodological and institutional preconditions were in place early for potential interactions between cognitive approaches to memory and the distributed cognition frameworks. However, in practice, these interactions were slow to build. We can see this by sketching the fate of two theoretical initiatives which might have had earlier influence on mainstream psychologies of memory than they did. Daniel Wegner and his colleagues, in a series of ambitious papers published in the 1980s, developed a theory of Transactive Memory Systems (Wegner 1987; Wegner et al. 1985). Focussing especially on long-term intimate couples, Wegner and colleagues argued that over the course of a rich shared history, dyads and other small groups can develop complex forms of cognitive interdependence. A transactive memory system, in Wegner’s terms, is a set of individual memory systems along with the established systems of communication between those individuals. Transactive memories reside in the memories of both individuals when considered as a combined system, given their shared awareness of the way information is distributed within the group, which can take more or less differentiated and specialized forms. Although Wegner suggested a number of natural directions for testing these ideas, they had surprisingly little influence in the cognitive or social psychology of memory, instead being taken up for many years almost exclusively in organizational psychology and in studies of memory and performance in work groups and teams. As a number of papers in this special issue show, however, transactive memory systems theory has taken on a new lease of life in recent years (see especially Theiner and Tollefsen, Dale, and Paxton). It has come back into cognitive psychology, in part, after being discussed in philosophy and social ontology (Tollefsen 2006; Barnier et al. 2008), while Wegner and colleagues have extended their own original focus on socially distributed remembering to look at our cognitive interdependence with the extended bodies of information available to us online (Sparrow et al. 2011).

A second, distinct line of thought in which human memory was seen within something like a distributed cognition framework was in Merlin Donald's book *Origins of the Modern Mind* (Donald 1991). Donald coined the word "exogram", by analogy with the brain's "engrams" or neural memory traces, to describe the external symbols which we as a species have come to combine, manipulate, and rely on. Changes in such external symbol systems since the upper Palaeolithic period have dramatically altered the capacity and operation of human memory (Donald 1991, pp. 308–333; Donald 1998). Identifying certain common features across the diverse history of external representations—body markings, grave decorations, hieroglyphics, maps, musical scores, writing systems, architectural diagrams and so on—Donald focussed our attention on the new cognitive profiles that characterise creatures (and societies of creatures) who can draw on these exograms in addition to neural engrams. Thoughts and memories, for example, become more durable and more easily transmissible and reformattable across media and contexts and are plugged in to vastly larger databases of inherited knowledge (Donald 1991, pp. 314–319). In an important early philosophical development of ideas about distributed cognition, Mark Rowlands built on Donald's work to argue that much of human memory is essentially (not accidentally) environment-involving, and primarily consists in our ability to interface with a range of different collective memory networks (Rowlands 1999, pp. 119–147, also drawing on important work by Rubin 1995). Yet, again, Donald's primary influences were outside the mainstream cognitive sciences of memory, as he was instead drawn in for example to extensive debates in cognitive archaeology with radical enactivists and phenomenologists who thought that his focus on symbolic systems was still far too close to classical cognitivism (Renfrew and Scarre 1998; Malafouris 2004; see also Sutton 2008). This is a continuing dynamic in studies of distributed cognition, with more ecumenical attempts to nudge mainstream cognitive theory and practice out of its individualist methods rubbing up against various waves of much more critical anti-representationist movements. Theorists with backgrounds in long-standing traditions of Wittgensteinian, phenomenological, direct realist, processual, or dynamical/enactivist critiques of the cognitive scientific enterprise as a whole have shown some ambivalence towards distributed cognition, which is not revolutionary enough for their liking; such issues are visible in at least two of the papers in this special issue (Dahlbäck et al., Loader).

But to return briefly to mainstream paradigms in the sciences of memory, further impetus towards an integration of distributed cognition into the experimental cognitive psychology of memory has come in recent years, by way of fruitful debates and improvements in a number of specific experimental paradigms. Significant examples are the collaborative recall paradigm (Weldon 2000; Harris et al. 2008; Rajaram and Pereira-Pasarin 2010), and work on socially-shared retrieval-induced forgetting (Cuc et al. 2007, and see the paper by Fagin, Yamashiro, and Hirst in this special issue). Where some other experimental work on false memory has painted social interaction as primarily a source of misinformation or memory contagion, work in these paradigms has encouraged careful exploration of conditions under which systematic reliance on external (especially social) resources transforms the processes and products of memory, for better or worse (Meade et al. 2009; Barnier 2010; Stone et al. 2012). The fact that there are natural extensions of these paradigms to study memory

processes in ageing individuals and couples is another advantage. Clark and Chalmers rested their case for extended cognition partly on the notorious Otto thought experiment, in which the systematic use of external aids transforms the memory tasks and processes of a man suffering memory decline: the thought experiment was not intended to substitute for investigation of the actual compensatory roles of technologies and other people in scaffolding us against memory loss (Harris et al. 2011; Drayson and Clark 2013; and see the paper by Dahlbäck, Kristiansson, and Stjernberg in this special issue). Partly as a result of these recent movements, a number of well-established cognitive, developmental, and neuropsychologists of memory now have active research interests in distributed cognition and related theoretical frameworks: in a recent review paper in the *Annual Review of Psychology*, William Hirst and Gerald Echterhoff discuss work by Hutchins, Clark, and others, noting that “the extended mind approach . . . places the study of conversational remembering front and center in a way more traditional approaches to the study of memory do not” (Hirst and Echterhoff 2012, p. 58).

2 Rationale for the Special Issue

In planning this special issue, we sought to reinstate the breadth of topics and concerns apparent in this historical sketch to contemporary debates about memory and distributed cognition. We wanted to build both on our own previous work (Barnier et al. 2008; Michaelian 2012a; Sutton 2010), and on prior special issues of journals such as *Philosophical Psychology* (Sutton 2006) and *Memory* (Barnier and Sutton 2008). Earlier versions of three of the eight papers in the current issue were presented as part of a symposium on memory and distributed cognition at ICOM-5, the 5th International Conference on Memory, at the University of York (UK) in August 2011. The largest of the ICOM meetings that have taken place every five years since 1991, this conference featured over 500 talks and 250 poster presentations. The vast majority of research presented at ICOM-5 was firmly anchored in the various fields of mainstream experimental psychology: cognitive, clinical, developmental, social, personality, and comparative psychology, for example, as well as cognitive neuroscience and neuropsychology. But in addition to the small minority of philosophers at the conference, there was a striking level of new interest among experimental psychologists too in the ways that remembering has not just neural but also bodily, social, and technological dimensions. To enrich such cross-disciplinary movements, we suggest that philosophers of memory and cognition need to immerse themselves fully in contemporary psychological research, as well as in the methods and assumptions of other disciplines that can potentially offer empirical data on memory, from anthropology to history and archaeology.

Ideas about distributed memory must both make sense of and, in turn, inform empirical work in the cognitive and social sciences in order to deserve wider acceptance. Recent years have seen an explosion of more specialized philosophical debates about how best to express and then to evaluate metaphysical claims about extended cognition. The primary philosophical debate has pitched the view that cognition can be fully and genuinely “extended” into the world against a more conservative view

that it is merely firmly “embedded” in the environment (Clark 2010; Rupert 2009; Sprevak 2009). A range of technical issues have arisen in these discussions about the best ways to capture just what it means to claim that cognition is “extended”, relating for example to the metaphysics of constitution, to questions about intrinsic and derived content, and to the basic issue of whether there is a single defining “mark of the cognitive”. Others have meanwhile sought to distinguish and critically evaluate a range of related approaches under labels such as “embodied cognition”, “enactive cognition”, and “scaffolded cognition”, with considerable progress being made in teasing apart weaker and stronger claims, and the central issues on which disagreements lie (Rowlands 2010; Sterelny 2010; Wheeler 2010).

While we fully appreciate, and have sought to contribute to, these philosophical debates, we see a potential danger in philosophical work exclusively remaining at this kind of distance from relevant empirical studies. This is why we have pitched the current special issue as relating memory research to “distributed” rather than “extended” cognition, since the labels “distributed cognition” and “distributed memory” arose from the rich interdisciplinary investigations we have described above well before the emergence of specialist philosophical debates about the metaphysics of extended cognition. We suggest that many important scientific questions concerning memory in relation to distributed cognition can be addressed without taking a definitive stand in advance on some of the more technical metaphysical issues about the ontology and location of cognitive states and processes: instead, in naturalistic fashion, philosophers can work directly from existing and independently-motivated projects and problems arising within the sciences of memory themselves. It may then turn out, if excellence of theory is a good measure of ontology, that those metaphysical issues might look a little different after the sciences have moved on. At least we can hope to bring a wider array of relevant empirical evidence and examples to bear in formulating and assessing our metaphysical views. As the papers in this issue attest, there is ample work for philosophers to do in critically evaluating ongoing memory research across the disciplines.

Ideally, of course, the metaphysical and naturalistic projects will be mutually informative. Perhaps, for example, we can interpret recent metaphysical discussions as gradually honing in on the identification of a key set of dimensions on which real cases may vary significantly (Wilson and Clark 2009; Rupert 2010; Sterelny 2010; Sutton et al. 2010; and see Clowes’ paper in this issue). On some occasions, remembering (for example) is a more shielded or isolated individual process, from which social resources and material supports are firmly screened off. On other occasions, both the enduring grounds and the occurrent processes of remembering are more interactively spread across heterogeneous inner and outer resources: this is perhaps more likely to occur among established, long-standing social groups with a rich shared history, in which that shared past is regularly revisited and renegotiated in interactive and perhaps contested processes of collaborative remembering, scaffolded by the array of technological supports to which we are now accustomed. Families reminiscing around the dinner table, friends sharing stories in the pub, work colleagues puzzling over old decisions or conflicts, sports teams revisiting recent or ancient triumphs and defeats, military veterans confronting the pain of shared experience, lovers retelling the precious tale of their first meeting, children learning that

their own perspectives on past events may not quite match that of their parents or siblings, or older adults keenly relying on partners, carers, or established systems of notes and reminders: these are all mundane contexts in which, although our memory systems are changing as our technologies constantly alter, we are accustomed daily to spread and incorporate external resources to varying degrees and at many different intensities. If we are as yet some way from a future multi-method science of memory in which experiment and ethnography would be seamlessly interwoven, we nonetheless hope that this special issue can take us forward a little towards addressing the challenges of relating or integrating analysis of the complementary neural, bodily, environmental, and cultural resources which coordinate in so many cases of remembering.

3 Overview of the Special Issue

The eight articles making up the special issue describe a broad range of approaches—from experimental and ethnographic work, through computer simulations, to general considerations in philosophy of science—to understanding distributed memory systems. Throughout the issue, there is an emphasis on interactions and points of contact between theoretical work on distributed and extended cognition and empirical memory research, with some articles focussing on potential contributions of the theory to empirical memory research, while others focus on implications of empirical memory research for theoretical debates around distributed/extended cognition.

Two natural ways of grouping the articles are (1) in terms of their focus on either agent-technology distributed memory systems or agent-agent distributed memory systems and (2) in terms of their focus on either metaphysical questions or explanatory questions about the mechanisms involved in distributed memory. **Agent-technology vs. agent-agent systems:** Arango–Muñoz, Clowes, Loader, and Rupert focus primarily on agent-technology interactions (though Rupert’s argument is pitched at a high level of generality and may apply to both agent-technology and agent-agent systems, and Arango–Muñoz proposes a framework which should apply equally to agent-agent systems), while Dahlbäck, Kristiansson, and Stjernberg focus more on cases involving both types of interaction at once. In contrast, Fagin, Yamashiro, and Hirst, as well as Theiner and Tollefsen, Dale, and Paxton, are concerned with distributed remembering in small groups of agents. (We return to the need for work looking simultaneously at agent-agent and agent-technology interactions in Section 5.) **Mechanisms vs. metaphysics:** The hypothesis of extended cognition, along with Rupert’s alternative hypothesis of embedded cognition, is a reference point for all of the articles dealing with agent-technology systems, but, while two of these (Rupert and Loader) are concerned with the implications of the psychology of memory for the metaphysics of external memory (though these two papers approach the question in radically different ways), the remainder (Clowes; Arango–Muñoz; and Dahlbäck, Kristiansson, and Stjernberg) focus more directly on explanatory questions about the mechanisms involved in interactions between agents and external memory resources. In the group of articles on agent-agent interactions, on the other hand,

both Theiner and Tollefsen, Dale, and Paxton draw on empirical memory research to respond to ontological arguments against the existence of irreducibly group-level cognitive systems (responding in part to arguments given elsewhere by Rupert), while Fagin, Yamashiro, and Hirst draw on such empirical research to answer the question of why humans engage in distributed group remembering.

3.1 Agent-technology Systems

3.1.1 *Mechanisms of External Memory*

Clowes is concerned with the cognitive implications of our increasing reliance on novel forms of *E-memory*—new computer-based memory technologies (as opposed to *O-memory*: internal, organic or bio-memory (Kalnikaitė and Whittaker 2008; Bell and Gemmell 2009; Smart 2012)). Critics (see Carr 2010 for one prominent example) have voiced a number of worries about this tendency, but the general thrust is that it leads to cognitive diminishment; Clowes responds to this diminishment argument by examining the factors determining the extent to which external memory technologies can achieve cognitive incorporation, and explores the implications of the hypothesis of extended cognition (HEC) (Clark and Chalmers 1998; Clark 2003) and the alternative hypothesis of embedded cognition (HEMC) (Rupert 2004, 2009) for the effects of incorporation. E-memory is characterized by a cluster of distinctive features—*totality*, *practical cognitive incorporability*, *autonomy*, and *entanglement*—which jointly determine the likelihood of its achieving deep cognitive incorporation (see also Heersmink 2012) and thereby transforming our cognitive profiles in important ways. While totality and practical cognitive incorporability encourage incorporation, autonomy and entanglement pull in the opposite direction, suggesting that we are likely to continue to rely heavily on E-memory, even as it is unlikely to achieve deep incorporation. Clowes argues that HEC may be better positioned than HEMC when it comes to avoiding an overly pessimistic interpretation of our increasing reliance on E-memory, since the former can acknowledge that an extended system (agent + E-memory device) may be capable of more than a nonextended system (agent alone), whereas the latter appears to be committed to the view that increased reliance on E-memory involves a diminished level of cognitive activity in the agent, the only cognitive system that it acknowledges.

Arango-Muñoz, like Clowes, adopts an optimistic view of our increasing reliance on novel forms of E-memory, but his optimism is based not on HEC (he works within Sterelny's alternative scaffolded cognition framework (Sterelny 2010), which assumes that, while more modest forms of cognitive scaffolding are commonplace, bona fide cognitive extension is rare) but rather on an investigation of the role of metacognition in agent-E-memory interactions. He describes generalized versions of the *selection problem* and the *endorsement problem* (Michaelian 2012b): When confronted with a given cognitive problem, how does the agent decide whether to solve it using an internal resource or, instead, an external resource? When information has been retrieved from a given internal or external resource, how does the agent determine whether to endorse it? Arango-Muñoz argues that agents solve both the internal and the external versions of these problems by relying on *metacognitive*

feelings—phenomenal feelings produced by subpersonal monitoring and control of cognitive processes (e.g., the feeling of knowing, the tip-of-the-tongue phenomenon) (De Sousa 2008; Dokic 2012; Proust 2007). On this view, agents are able to manage the trade-off between internal and external memory in the apparently adaptive manner described, for example, in Sparrow et al.'s recent study (Sparrow et al. 2011), due to the reliability of metacognitive feelings (Reber and Unkelbach 2010), along with their ability to take into account both the reliability of the internal and external sources available to them and the costs involved in accessing internal vs. external sources.

Arango-Muñoz's focus on the role of the interplay between internal and external memory in achieving adaptive remembering is shared by **Dahlbäck, Kristiansson, and Stjernberg**, who argue, on the basis of an ethnographic study of the mnemonic strategies used by older adults coping with cognitive decline, that remembering often involves a complex web of processes drawing on both internal and external resources, including both external memory stores and other agents. Observing that the distinction between internal and external memory is often blurry (Sutton 2010) (consider the case of the mnemonic for remembering the number of days in each month by counting your knuckles and the spaces between them), they argue that research on distributed remembering should focus more on the *process* of remembering, assigning a greater role to the active, structuring role of the remembering agent than has often been the case. In their view, cognition should be viewed as being distributed across the agent, the activity, and the setting (Lave 1988), and they apply this to focus on how the activity of remembering is shaped and constrained by the changing nature of the task and the environment, with the agent negotiating these changes in order to remember successfully. For example, an older adult may rely on an external resource to remind her to complete a task (a case of external prospective memory), but, anticipating the possibility of failure, she may also set up an additional external resource which will kick in if the primary resource fails to remind her.

3.1.2 *Metaphysics of External Memory*

This approach, with its emphasis on the role of the agent in actively structuring his environment and adaptively responding to changing conditions in order to achieve a *resilient* distributed memory system, would seem to take us well beyond the simple picture of the interaction between the agent and his external memory suggested by the case of Otto and his notebook (along with Inga and her biological memory) in Clark and Chalmers' original article (1998). **Loader** similarly argues that the view of external memory which has informed debates over HEC rests on an objectified conception of internal memory, which would better be replaced by a conception of remembering as an activity. According to Loader, just as Otto's external memory is described by Clark and Chalmers in objectified terms, so is Inga's (normal) biological memory viewed as a sort of portable store that the agent carries around with her. Loader argues against this objectified conception of memory, drawing on a variety of theoretical resources, including enactivism (Noë 2004), a Wittgensteinian/connectionist view of remembering as an activity not requiring stored representations (Wittgenstein 1980; Stern 1991), and a Vygotskian view of

the development of remembering as a process of internalizing previously external strategies (Vygotsky 1978). He suggests that, insofar as Inga's internal memory is notebook-like, this is due to her previous interaction with notebook-like forms of external memory.

Responding to critiques (including Clark 2010; Menary 2006; Rowlands 2009; Sprevak 2009; Adams and Aizawa 2008; Wheeler 2011, 2010) of his earlier arguments (Rupert 2004, 2009) against the hypothesis of *extended* cognition defended by Clark (and Chalmers) (Clark and Chalmers 1998; Clark 2003), **Rupert** provides new arguments in favour of the alternative, less radical hypothesis of *embedded* cognition, according to which genuine cognitive extension is possible but rare, with embedding rather than extension being the typical case. The focus here is on a reading of Clark and Chalmers' argument for HEC on which it maintains that the natural kinds investigated by cognitive science have both internally and externally located instances. Rupert argues that, if natural kinds are individuated in a fine-grained way, the memory-related kinds investigated by cognitive science are unlikely to have externally-located instances. If, alternatively, kinds are individuated in a more coarse-grained way, so that they are more likely to have both internally- and externally-located instances, he argues, the natural kinds argument fares no better: the generic kind "memory" seems to play no explanatory role in cognitive science, and the disunity of human memory may suggest that a coherent science of memory as such is impossible. Rupert also considers the role of the so-called "Martian intuition" in the extended cognition debate, arguing that it cannot provide significant support for HEC, before turning his attention to the existence of a generic kind "cognition"; he argues for scepticism on this front, suggesting that current cognitive science is largely about specifically human cognition.

3.2 Agent-agent Systems

3.2.1 Mechanisms of Group Memory

Fagin, Yamashiro, and Hirst focus on distributed conversational remembering in small groups, drawing on current experimental and simulation-based work to argue that distributed remembering is adaptive not, as the focus of much of the literature would suggest, due primarily to its beneficial effects on individual memory, but rather because it promotes sociality, promoting *mnemonic convergence* and thereby contributing to the formation of collective memory, which is crucial to humans as a social species. As Theiner also emphasizes (see below), distributed remembering has cognitive costs as well as cognitive benefits. Fagin, Yamashiro, and Hirst focus here on two apparent costs of distributed remembering: *social contagion* (Roediger et al. 2001) and *socially shared retrieval-induced forgetting* (SS-RIF) (Cuc et al. 2006; Stone et al. 2009; Hirst and Echterhoff 2012). Social contagion and SS-RIF promote mnemonic convergence, the formation of a shared representation of the past. And results of simulations suggest that, under the right conditions, this may happen on a larger scale, with SS-RIF propagating in groups of up to 50 individuals (Coman et al. 2012). Fagin, Yamashiro, and Hirst thus argue that, though we will

need to invoke additional mechanisms to explain the formation of collective memories in large-scale societies, our tendency to engage in conversational remembering may well be explained to some extent by its promotion of mnemonic convergence, with its resulting benefits for sociality.

3.2.2 *Metaphysics of Group Memory*

The remaining articles draw on empirical work on memory in small groups to respond to reductionist arguments about group cognitive systems. **Theiner** draws on analyses of mechanistic explanation in the philosophy of science (Bechtel 2009; Craver 2007), including Wimsatt's work on reduction and emergence (Wimsatt 1986), to argue that research on *transactive memory systems* (TMS) (Wegner 1987; Ren and Argote 2011) supports nonreductionism about group-level cognition, in opposition to the more popular reductionist view. Pavitt (Pavitt 2003a; 2003b), for example, has argued that, given that the typical finding is *collaborative inhibition*, groups do not remember more than the sums of their parts, so the notion of a TMS can only be metaphorical. While there is also some evidence for *collaborative facilitation* (Meade et al. 2009; Harris et al. 2011), Theiner's case for a non-metaphorical understanding of TMSs does not rest on this evidence. Contemporary approaches to group cognition explicitly adopt an information processing perspective (Larson and Christensen 1993; Hollan et al. 2000) and are thus compatible with standards of mechanistic explanation, but they do not locate the mechanisms responsible for group-level cognition entirely at the individual level, thus allowing for group cognition as an emergent phenomenon. In the case of transactive memory, it appears that the explanatory work is done in part by the organization of the individuals constituting the transactive memory system, which means that a complete mechanistic explanation of transactive memory will refer not only to individual information processing mechanisms but also to the collective information processing that emerges from the organization of the group. If so, Pavitt's argument for reductionism about group memory rests on a mistake: what matters is not whether groups remember more than the aggregated memories of their members, but rather whether group memory is an organization-dependent phenomenon that departs systematically from aggregated individual memories; from this perspective, both collaborative inhibition and collaborative facilitation can provide evidence for the existence of irreducibly group-level cognitive systems.

Like Theiner, **Tollefsen, Dale, and Paxton** appeal to research on TMSs to argue for irreducibly group-level cognition, but, rather than appealing to general theoretical considerations, they rely primarily on empirical work on *alignment* in linguistic and interpersonal interaction (Richardson and Dale 2005; Tollefsen and Dale 2011). Alignment refers to the sort of dynamic matching between the behavioural or cognitive states of two or more people occurring, e.g., in conversation, in which it can be observed in everything from body posture to sentence structure to use of irony, and which may be generated primarily by low-level mechanisms, such as priming and the mirror neuron system. Tollefsen and Dale have argued that this forms the basis for what they refer to as an *alignment system*, "a loosely interconnected set of cognitive processes that have evolved to facilitate social interactions and joint actions,

in particular” (Tollefsen and Dale 2011), and they here argue that further research on alignment in the context of transactive memory may well support the claim that alignment gives rise to group-level cognition, since alignment may be an important part of the mechanism by which people become coupled to form a TMS. They argue that, in certain circumstances, groups may achieve the sort of tight coupling required for them to form cognitive systems—even in Rupert’s demanding sense, according to which cognitive systems must be persistent and display a set of capacities that persists across different contexts (Rupert 2011)—supporting this by appealing both to work in fields such as computer-supported collaborative learning (Dillenbourg et al. 2009) and computer-supported cooperative work (Olson and Olson 2008) and to their own current research examining alignment in naturalistic conversation in dyads (Paxton and Dale 2012).

4 Implications for Research on Distributed Memory

Considered as a whole, the contributions to the special issue have a number of general implications for both empirical and philosophical work on distributed and extended memory.

4.1 Implications for Empirical Memory Research

If Fagin, Yamashiro, and Hirst’s main conclusion—that distributed remembering is adaptive not only due to its impact on individual cognition but also due to its promotion of sociality—is right, this suggests a broadening of focus for research on distributed cognition. Taking an example from a different domain, Mercier and Sperber have recently argued that, while it is natural to view reasoning as a means to improve and extend the agent’s knowledge, viewing it in evolutionary perspective as a means to influence other agents allows us to account for a range of common errors and distortions (for example, the confirmation bias) that would otherwise be difficult to explain (Mercier and Sperber 2011). Analogously, **viewing distributed remembering in evolutionary perspective**, as a means to promoting group cohesion, may enable us to explain why we tend to engage in conversational remembering despite effects (such as social contagion and SS-RIF) which, from the point of view of individual cognition, can only appear as costs.

At this stage, **it remains unclear to what extent differences among different distributed/extended theories can make a difference to empirical memory research** (this point likely generalizes to other areas of empirical research). Coming to grips with the nuances of the different positions that have emerged in the theoretical debates around distributed cognition and extended mind—extended cognition, embedded cognition, scaffolded cognition, and so on—is likely not a prerequisite for effective empirical research on the roles played by external resources (whether technological or social) in remembering. For example, while Arango-Muñoz adopts the scaffolded cognition framework, his explanation of how agents rely on metacognitive feelings to negotiate the trade-off between internal and external memory would

seem to be consistent with an extended cognition framework. Similarly, Dahlbäck, Kristiansson, and Stjernberg's exploration of the active role of the agent in structuring his environment so as to be able to adaptively draw on external resources in remembering is not hampered by their decision not to take a stand in the embedded/extended debate; nor need Fagin, Yamashiro, and Hirst answer the question of whether (and when) cognition is genuinely extended in order to develop their explanation of the adaptivity of conversational remembering. As Clowes remarks, external memory can influence our cognitive profiles regardless of whether it achieves cognitive incorporation sufficiently deep to count as part of our minds, so the investigation of the impact of external memory on remembering need not wait for a resolution of the theoretical debate over cognitive extension. As far as empirical research is concerned, then, the value of distributed/extended theories may lie primarily in their emphasis on the importance of agent-technology and agent-agent couplings for explanations of how memory works in ecological settings.

4.2 Implications for the Philosophical Debate

Going in the other direction, **it is unclear whether empirical evidence can resolve the debates among extended/embedded/etc. theories.** Theiner and Tollefsen, Dale, and Paxton do argue that empirical TMS research supports the existence of irreducible group-level cognitive systems, while Rupert's argument against HEC draws on empirical claims about the functional profiles of biological memory systems. On the other hand, Clowes (though he argues that HEC may prove more useful than HEMC in investigating the effects of our increasing reliance on novel forms of external memory) points out that, given a case of putatively extended cognition, a theorist committed to one or another alternative view can always redescribe it so that it is consistent with his view. Thus it is at this point unclear whether empirical evidence can help to resolve the debate among the varieties of distributed theory.

A number of the contributions point out that much of the theoretical debate over HEC has relied on unrealistically simple examples of external memory, such as Otto's notebook, and argue that **theoretical discussions of distributed/extended cognition should incorporate more realistic examples of external memory.** For example, Arango-Muñoz points out that Otto's case is unusual in that he has only one external memory resource to rely on, whereas real agents typically have to choose between internal and external memory and among multiple available external memory resources. Dahlbäck, Kristiansson, and Stjernberg take a similar position, urging attention to the complexities of the interaction between the agent and the relevant external memory resources—reliance on external memory is rarely simply a matter of automatically accepting externally retrieved information, as it is in the Otto case. And Clowes points out that, when we attend to differences among memory technologies, we see that there may be a spectrum of cases, ranging from full-blown incorporation, in which it is relatively plausible that cognitive processing extends into the external resource, to regular reliance without incorporation, in which, while it is implausible to say that cognitive processing extends into the resource, use of the resource nonetheless significantly shapes internal cognitive

processing. As Loader argues, reliance on unrealistically simple examples of external memory may be due to acceptance of an inadequate conception of biological memory; attention to more realistic examples of external memory, as well as to differences between biological and external memory (Rupert 2004; Michaelian 2012a) may reinforce the move from first-wave, parity-based distributed cognition to second-wave, complementarity-based distributed cognition (Sutton 2010).

5 Directions for Future Research

Where do we go from here? The special issue suggests a number of questions for future research on both the metaphysics and the mechanisms of distributed memory systems.

5.1 Metaphysics of Distributed Memory Systems

What is the relevance of theoretical controversies to empirical memory research? As noted above, it is at present unclear to what extent theoretical debates around HEC, HEMC, and related views have implications for empirical memory research. Future work could explore this question, either showing that there are such implications or arguing that the theoretical debate is worthwhile even if there are not.

What are the implications of distributed memory research for the ontology of group cognition? Assuming that Theiner and Tollefsen, Paxton, and Dale are right that empirical research on alignment and transactive memory may provide evidence of irreducibly group-level cognition (perhaps consistent with Rupert's systems-based approach), this represents a potentially profitable area of research. But even should it turn out that empirical research cannot provide evidence of irreducibly group-level cognition, there is clearly a need for the sort of research described by Tollefsen, Paxton, and Dale on the low-level mechanisms responsible for creating and sustaining TMSs.

5.2 Mechanisms of Distributed Memory Systems

What are the cognitive mechanisms mediating interactions between biological memory and external memory? Arango-Muñoz's metacognitive framework is designed to explain how agents select among available internal and external resources when faced with a given memory task, as well as how they decide whether to endorse information retrieved from a given internal or external resource. While it is empirically grounded, this framework remains speculative, and future research, perhaps drawing inspiration from existing experimental work on metacognitive feelings, could seek to test and refine it. (Such research could also examine whether there is a role for metacognitive feelings in resolving conflicts among information retrieved from internal and external sources.)

What are the implications of human communication research for distributed memory research? The external version of the endorsement problem has both non-social and social forms, and the latter brings distributed memory research into contact with human communication research, which looks at the factors influencing the decision to trust information received from another agent (Sperber et al. 2010). Much of this research is pessimistic with respect to our sensitivity to deception (Michaelian 2012c); a question for future research is thus whether the results of human communication research should temper the optimism of Arango–Muñoz’s view of the reliability of scaffolded memory. Such research should take into account the non-cognitive factors emphasized by Fagin, Yamashiro, and Hirst.

What are the mechanisms of mnemonic convergence at the population scale? As Fagin, Yamashiro, and Hirst acknowledge, the effectiveness of social contagion and SS-RIF as means for producing mnemonic convergence may be limited to relatively small groups, so explanations of mnemonic convergence at the level of entire societies will likely need to invoke additional factors. This suggests a point of contact between research on distributed memory in small groups and research on the mechanisms of collective memory formation at the level of entire societies (for example, Anastasio et al. 2012), an area which so far is underexplored.

What do entanglement and autonomy mean for episodic and autobiographical memory? As Clowes suggests, emerging E-memory technologies correspond most closely to semantic memory. But we can expect them to have important effects not only on semantic but also on episodic and autobiographical memory. Given the reconstructive nature of episodic memory, the availability of massive verbatim records related to personal experiences (for example, through lifelogging (Sellen and Whittaker 2010) and related practices) may strongly shape episodic memory. Given the novel features of E-memory—autonomy and entanglement, in particular—the resulting new forms of episodic memory may themselves have novel features, with greater influence from other agents over the contents of the individual’s personal memories.

How do agent-technology-agent couplings work? Most of the contributions to the special issue focus primarily either on agent-technology interactions in distributed cognition or on agent-agent interactions, without investigating patterns that appear when both types of interaction are operating together. Future research could look for such patterns, perhaps following up on Arango–Muñoz’s suggestion that epistemic feelings may play a mediating role in both cases or employing the sort of ethnographic methods used by Dahlbäck, Kristiansson, and Stjernberg.

5.3 Normative Questions About Distributed Memory

As Arango–Muñoz and Clowes note in their contributions, the recent debates over the implications of novel external memory technologies for our mnemonic ability are broadly reminiscent of much older worries about the implications of writing itself (voiced, for example, in Plato’s *Phaedrus*), and indeed, though normative questions

are beyond the scope of this issue, there is room for future work in this area, in both epistemology and ethics.

What are the epistemological implications of distributed memory research?

There is an emerging debate over the epistemological implications of extended cognition, since there appears to be a tension between virtue/credit theories in epistemology and HEC (see Pritchard 2010; Vaesen 2011; Kelp 2011; Clark et al. 2012). While it is too early to anticipate the outcome of this debate, the central example of external memory will surely play a role here; ideally, the debate will be informed by empirically sound work on the ways in which agents interact with external memories. If it turns out that HEC is incompatible with plausible epistemological theories, some may be prepared to count this as evidence in favour of Rupert's alternative HEMC.

What are the ethical implications of distributed memory research?

What are the ethical implications of our increasing reliance on novel forms of external memory? So far there is little work in this area, but parallels with ethical issues around memory enhancement (Liao and Sandberg 2008) suggest that it represents a potentially fruitful avenue of inquiry. For example, Loader points out that literally extended memory raises the possibility that one agent's memories could be accessed or stolen by another, a scenario which raises interesting ethical questions.

Acknowledgements Thanks to Roberto Casati and Paul Egré for their support during the preparation of the special issue (and to Paul for comments on this editorial). Thanks also to the referees for submissions to the special issue, many of whom delivered detailed reports on very short notice.

References

- Adams, F., & Aizawa, K. (2008). *The bounds of cognition*. Blackwell.
- Agre, P.E., & Chapman, D. (1989). What are plans for? In *Robotics and autonomous systems* (pp. 17–34). MIT Press.
- Anastasio, T.J., Ehrenberger, K.A., Watson, P., Zhang, W. (2012). *Individual and collective memory consolidation: Analogous processes on different levels*. MIT.
- Anderson, J.R. (1976). *Language, memory, and thought*. Lawrence Erlbaum.
- Barnier, A.J. (2010). Memories, memory studies and my iPhone: editorial. *Memory Studies*, 3(4), 293–297.
- Barnier, A.J. (2012). Memory, ecological validity and a barking dog. *Memory Studies*, 5(4), 351–359.
- Barnier, A.J., & Sutton, J. (2008). From individual to collective memory: theoretical and empirical perspectives. *Memory*, 16(3), 177–182.
- Barnier, A.J., Sutton, J., Harris, C.B., Wilson, R.A. (2008). A conceptual and empirical framework for the social distribution of cognition: the case of memory. *Cognitive Systems Research*, 9(1–2), 33–51.
- Bartlett, F.C. (1995/1932). *Remembering: A study in experimental and social psychology*. Cambridge.
- Bechtel, W. (2009). Looking down around, and up: mechanistic explanation in psychology. *Philosophical Psychology*, 22(5), 543–564.
- Bell, C., & Gemmell, J. (2009). *Total recall: How the E-memory revolution will change everything*. Dutton.
- Boden, M. (2006). *Mind as machine: A history of cognitive science*. Oxford, 2 volumes.
- Butler, K. (1998). *Internal affairs: Making room for psychosemantic internalism*. Kluwer.
- Butterfill, S.A., & Sebanz, N. (2011). Editorial: joint action: what is shared? *Review of Philosophy and Psychology*, 2(2), 137–146.

- Carr, N. (2010). *The shallows: How the internet is changing the way we think, read, and remember*. Atlantic.
- Clancey, W.J. (2009). Scientific antecedents of situated cognition. In P. Robbins & M. Aydede (Eds.), *Cambridge handbook of situated cognition* (pp. 11–34). Cambridge.
- Clark, A. (1997). *Being there: Putting brain, body, and world together again*. MIT.
- Clark, A. (1998). Authors response: review symposium on *being there*. *Metascience*, 7, 95–103.
- Clark, A. (2001a). *Mindware: An introduction to the philosophy of cognitive science*. Oxford.
- Clark, A. (2001b). Reasons, robots and the extended mind. *Mind & Language*, 16(2), 121–145.
- Clark, A. (2003). *Natural-born cyborgs: Minds, technologies and the future of human intelligence*. Oxford.
- Clark, A. (2010). *Supersizing the mind: Embodiment, action, and cognitive extension*. Oxford.
- Clark, A. (2011). Finding the mind. *Philosophical Studies*, 152(3), 447–461.
- Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, 58(1), 7–19.
- Clark, A., Pritchard, D., Vaesen, K. (Eds.) (2012). Extended cognition and epistemology. *Special Issue of Philosophical Explorations*.
- Coman, A., Kolling, A., Lewis, M., Hirst, W. (2012). Mnemonic convergence: From empirical data to large-scale dynamics. In S.J. Yang, A.M. Greenberg, M. Endsley (Eds.), *Social computing, behavioral—Cultural modeling and prediction* (Vol. 7227 of Lecture Notes in Computer Science, chapter 31, pp. 256–265). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Conway, M.A. (1990). *Autobiographical memory: An introduction*. Open University Press.
- Craver, C. (2007). *Explaining the brain: Mechanisms and the mosaic unity of neuroscience*. Clarendon.
- Cuc, A., Koppel, J., Hirst, W. (2007). Silence is not golden. *Psychological Science*, 18(8), 727–733.
- Cuc, A., Ozuru, Y., Manier, D., Hirst, W. (2006). On the formation of collective memories: the role of a dominant narrator. *Memory & Cognition*, 34(4), 752–762.
- De Sousa, R. (2008). Epistemic feelings. In G. Brun, U. Doğuoğlu, D. Kuenzle (Eds.), *Epistemology and emotions*. Ashgate.
- Dillenbourg, P., Järvelä, S., Fischer, F. (2009). The evolution of research on computer-supported collaborative learning. In N. Balacheff, S. Ludvigsen, T. Jong, A. Lazonder, S. Barnes (Eds.), *Technology-enhanced learning* (chapter 1, pp. 3–19). Netherlands, Dordrecht: Springer.
- Dokic, J. (2012). Seeds of self-knowledge: Noetic feelings and metacognition. In M.J. Beran, J.L. Brandel, J. Perner, J. Proust (Eds.), *Foundations of metacognition* (pp. 302–321). Oxford.
- Donald, M. (1991). *Origins of the modern mind*. Cambridge, MA: Harvard UP.
- Donald, M. (1998). Hominid enculturation and cognitive evolution. In C. Renfrew & C. Scarre (Eds.), *Cognition and material culture: The archaeology of symbolic storage* (pp. 7–17). McDonald Institute for Archaeological Research.
- Drayson, Z., & Clark, A. (2013). Augmentation, agency, and the spreading of the mental state. *Neuroethics* (forthcoming).
- Greenfield, P. (1984). A theory of the teacher in the learning activities of everyday life. In B. Rogoff & J. Lave (Eds.), *Everyday cognition* (pp. 117–138). Harvard.
- Halbwachs, M. (1992). The social frameworks of memory. In L.A. Coser (Ed.), *On collective memory* (pp. 35–189). University of Chicago Press (Original work published 1925).
- Harris, C.B., Barnier, A.J., Sutton, J. (2011). Consensus collaboration enhances group and individual recall accuracy. *The Quarterly Journal of Experimental Psychology*, 65(1), 179–194.
- Harris, C.B., Paterson, H.M., Kemp, R.I. (2008). Collaborative recall and collective memory: what happens when we remember together? *Memory*, 16(3), 213–230.
- Heersmink, R. (2012). Mind artifact: A multidimensional matrix for exploring cognition-artifact relations. In *Proceedings of AISB/IACAP world congress*.
- Hirst, W., & Echterhoff, G. (2012). Remembering in conversations: the social sharing and reshaping of memories. *Annual Review of Psychology*, 63(1), 55–79.
- Hollan, J., Hutchins, E., Kirsh, D. (2000). Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction*, 7(2), 174–196.
- Hutchins, E. (1996). *Cognition in the wild*. Bradford.
- Hutchins, E. (2005). Material anchors for conceptual blends. *Journal of Pragmatics*, 37(10), 1555–1577.
- Hutchins, E. (2011). Enculturating the supersized mind. *Philosophical Studies*, 152(3), 437–446.
- Kalninkaitė, V., & Whittaker, S. (2008). Cueing digital memory: How and why do digital notes help us remember? In *Proceedings of the 22nd British HCI group annual conference on people and computers: Culture, creativity, interaction* (Vol. 1, BCS-HCI '08, pp. 153–161). Swinton, UK: British Computer Society.

- Kelp, C. (2011). Extended cognition and robust virtue epistemology. *Erkenntnis*, 1–8. doi:10.1007/s10670-011-9301-3.
- Larson, J.R., & Christensen, C. (1993). Groups as problem-solving units: toward a new meaning of social cognition. *British Journal of Social Psychology*, 32(1), 5–30.
- Lave, J. (1988). *Cognition in practice: Mind, mathematics, and culture in everyday life*. Cambridge.
- Liao, S., & Sandberg, A. (2008). The normativity of memory modification. *Neuroethics*, 1(2), 85–99.
- Malafouris, L. (2004). The cognitive basis of material engagement: Where brain, body and culture conflate. In E. DeMarrais, C. Gosden, C. Renfrew (Eds.), *Rethinking materiality: The engagement of mind with the material world* (pp. 53–62). McDonald Institute for Archaeological Research.
- Meade, M.L., Nokes, T.J., Morrow, D.G. (2009). Expertise promotes facilitation on a collaborative memory task. *Memory*, 17(1), 39–48.
- Menary, R. (2006). Attacking the bounds of cognition. *Philosophical Psychology*, 19(3), 329–344.
- Mercery, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences*, 34(2), 57–74.
- Michaelian, K. (2012a). Is external memory memory? Biological memory and extended mind. *Consciousness and Cognition*, 21(3), 1154–1165.
- Michaelian, K. (2012b). Metacognition and endorsement. *Mind & Language*, 27(3), 284–307.
- Michaelian, K. (2012c). (Social) metacognition and (Self-)trust. *Review of Philosophy and Psychology*, 3(4), 481–514.
- Miller, G.A., Galanger, E., Pribham, K.H. (1960). *Plans and the structure of behavior*. Henry Holt.
- Neisser, U. (1997). The ecological study of memory. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 352(1362), 1697–1701.
- Nelson, K., & Fivush, R. (2004). The emergence of autobiographical memory: a social cultural developmental theory. *Psychological Review*, 111(2), 486–511.
- Noë, A. (2004). *Action in perception: The enactive approach to perception*. MIT.
- Norman, D.A. (1993). *Things that make us smart*. Basic Books.
- Olick, J.K., & Joyce, R. (1998). Social memory studies: from “collective memory” to the historical sociology of mnemonic practices. *Annual Review of Sociology*, 24(1), 105–140.
- Olson, G.M., & Olson, J.S. (2008). Computer-supported cooperative work. In F.T. Durso, R.S. Nickerson, S.T. Dumais, S. Lewandowsky, T.J. Perfect (Eds.), *Handbook of applied cognition* (2nd Edn., pp. 497–526). John Wiley & Sons Ltd.
- Pavitt, C. (2003). Colloquy: do interacting groups perform better than aggregates of individuals? *Human Communication Research*, 29(4), 592–599.
- Pavitt, C. (2003). Why we still have to be reductionists about group memory. *Human Communication Research*, 29(4), 624–629.
- Paxton, A., & Dale, R. (2012). Frame-differencing methods for measuring bodily synchrony in conversation. *Behavior Research Methods*. doi:10.3758/s13428-012-0249-2.
- Pritchard, D. (2010). Cognitive ability and the extended cognition thesis. *Synthese*, 175(0), 133–151.
- Proust, J. (2007). Metacognition and metarepresentation: is a self-directed theory of mind a precondition for metacognition? *Synthese*, 159(2), 271–295.
- Rajaram, S., & Pereira-Pasarin, L.P. (2010). Collaborative memory: cognitive research and theory. *Perspectives on Psychological Science*, 5(6), 649–663.
- Reber, R., & Unkelbach, C. (2010). The epistemic status of processing fluency as source for judgments of truth. *Review of Philosophy and Psychology*, 1(4), 563–581.
- Ren, Y., & Argote, L. (2011). Transactive memory systems 19852010: an integrative framework of key dimensions, antecedents, and consequences. *The Academy of Management Annals*, 5(1), 189–229.
- Renfrew, C., & Scarre, C. (Eds.) (1998). *Cognition and material culture: The archaeology of symbolic storage*. MacDonald Institute for Archaeological Research.
- Richardson, D.C., & Dale, R. (2005). Looking to understand: the coupling between speakers’ and listeners’ eye movements and its relationship to discourse comprehension. *Cognitive Science*, 29(6), 1045–1060.
- Roediger, H.L., Meade, M.L., Bergman, E.T. (2001). Social contagion of memory. *Psychonomic Bulletin & Review*, 8(2), 365–371.
- Rogers, Y. (1997). A brief introduction to distributed cognition. Available at: <http://mcs.open.ac.uk/yr258/papers/dcog/dcog-brief-intro.pdf>.
- Rogoff, B. (1991). *Apprenticeship in thinking: Cognitive development in social context*. Oxford.

- Rogoff, B., & Lave, J. (Eds.) (1984). *Everyday cognition*. Harvard.
- Rowlands, M. (1999). *The body in mind: Understanding cognitive processes*. Cambridge.
- Rowlands, M. (2009). Extended cognition and the mark of the cognitive. *Philosophical Psychology*, 22(1), 1–19.
- Rowlands, M. (2010). *The new science of the mind: From extended mind to embodied phenomenology*. MIT.
- Rubin, D.C. (1995). *Memory in oral traditions: The cognitive psychology of epic, ballads, and counting-out rhymes*. Oxford.
- Rumelhart, D.E., Smolensky, P., McClelland, J.L., Hinton, G.E. (1986). Schemata and sequential thought processes in PDP models. In J.L. McClelland & D.E. Rumelhart (Eds.), *Parallel distributed processing* (Vol. 2, pp. 7–57). MIT.
- Rupert, R. (2009). *Cognitive systems and the extended mind*. Oxford.
- Rupert, R.D. (2004). Challenges to the hypothesis of extended cognition. *The Journal of Philosophy*, 101(8).
- Rupert, R.D. (2010). Extended cognition and the priority of cognitive systems. *Cognitive Systems Research*, 11(4), 343–356.
- Rupert, R.D. (2011). Empirical arguments for group minds: a critical appraisal. *Philosophy Compass*, 6(9), 630–639.
- Saito, A. (2000). *Bartlett, culture, and cognition*. Psychology Press.
- Salomon, G. (Ed.) (1993). *Distributed cognitions: Psychological and educational considerations*. Cambridge.
- Sellen, A.J., & Whittaker, S. (2010). Beyond total capture: a constructive critique of lifelogging. *Communications of the ACM*, 53(5), 70–77.
- Smart, P.R. (2012). The Web-extended mind. *Metaphilosophy*, 43(4), 446–463.
- Sparrow, B., Liu, J., Wegner, D.M. (2011). Google effects on memory: cognitive consequences of having information at our fingertips. *Science*, 333(6043), 776–778.
- Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origg, G., Wilson, D. (2010). Epistemic vigilance. *Mind & Language*, 25(4), 359–393.
- Sprevak, M. (2009). Extended cognition and functionalism. *Journal of Philosophy*, 106(9), 503–527.
- Sterelny, K. (2010). Minds: extended or scaffolded? *Phenomenology and the Cognitive Sciences*, 9(4), 465–481.
- Stern, D.G. (1991). Models of memory: Wittgenstein and cognitive science. *Philosophical Psychology*, 4(2), 203–218.
- Stone, C.B., Barnier, A.J., Sutton, J., Hirst, W. (2009). Building consensus about the past: schema consistency and convergence in socially shared retrieval-induced forgetting. *Memory*, 18(2), 170–184.
- Stone, C.B., Barnier, A.J., Sutton, J., Hirst, W. (2012). Forgetting our personal past: socially-shared retrieval-induced forgetting of autobiographical memories. *Journal of Experimental Psychology: General*. doi:10.1037/a0030739.
- Suchman, L. (1987). *Plans and situated actions*. Cambridge.
- Sutton, J. (2008). Material agency, skills, and history: Distributed cognition and the archaeology of memory. In L. Malafouris & C. Knappett (Eds.), *Material agency: Towards a non-anthropocentric approach* (pp. 37–55). Springer.
- Sutton, J. (2009). Remembering. In P. Robbins & M. Aydede (Eds.), *Cambridge handbook of situated cognition* (pp. 217–235). Cambridge.
- Sutton, J. (2006). Introduction: memory, embodied cognition, and the extended mind. *Philosophical Psychology*, 19(3), 281–289.
- Sutton, J. (2010). Exograms and interdisciplinarity: History, the extended mind, and the civilizing process. In R. Menary (Ed.), *The extended mind* (pp. 189–225). MIT.
- Sutton, J., Harris, C., Keil, P., Barnier, A. (2010). The psychology of memory, extended cognition, and socially distributed remembering. *Phenomenology and the Cognitive Sciences*, 9(4), 521–560.
- Thelen, E., & Smith, L. (1994). *A dynamic systems approach to the development of cognition and action*. MIT.
- Tollefsen, D., & Dale, R. (2011). Naturalizing joint action: a process-based approach. *Philosophical Psychology*, 25(3), 385–407.
- Tollefsen, D.P. (2006). From extended mind to collective mind. *Cognitive Systems Research*, 7(2–3), 140–150.
- Vaesens, K. (2011). Knowledge without credit, exhibit 4: extended cognition. *Synthese*, 181(3), 515–529.

- Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. Harvard.
- Wegner, D. (1987). Transactive memory: A contemporary analysis of the group mind. In B. Mullen & G. Goethals (Eds.), *Theories of group behavior* (pp. 185–208). Springer-Verlag.
- Wegner, D.M., Giuliano, T., Hertel, P. (1985). Cognitive interdependence in close relationships. In W.J. Ickes (Ed.), *Compatible and incompatible relationships* (pp. 253–276). Springer.
- Weldon, M.S. (2000). Remembering as a social process. In D.L. Medin (Ed.), *The psychology of learning and motivation* (Vol. 40, pp. 67–120). Elsevier.
- Wheeler, M. (2010). In defense of extended functionalism. In R. Menary (Ed.), *The extended mind*. MIT.
- Wheeler, M. (2011). In search of clarity about parity. *Philosophical Studies*, 152(3), 417–425.
- Wilson, R.A., & Clark, A. (2009). How to situate cognition: Letting nature take its course. In P. Robbins & M. Aydede (Eds.), *Cambridge handbook of situated cognition* (pp. 55–77). Cambridge.
- Wimsatt, W.C. (1986). Forms of aggregativity. In M.G. Grene, A. Donagan, A.N. Petrovich, M.V. Wedin (Eds.), *Human nature and natural knowledge* (pp. 259–291). Reidel.
- Wittgenstein, L. (1980). *Remarks on the philosophy of psychology*. Chicago.
- Wood, D., Bruner, J.S., Ross, G. (1976). The role of tutoring in problem solving*. *Journal of Child Psychology and Psychiatry*, 17(2), 89–100.
- Zhang, J., & Norman, D.A. (1994). Representations in distributed cognitive tasks. *Cognitive Science*, 18(1), 87–122.