Are There Any Situated Cognition Concepts of Memory Functioning as Investigative Kinds in the Sciences of Memory?

Ruth Hibbert

PhD Thesis in Philosophy

June 2015

University of Kent

School of European Culture and Languages

Word count: 98 590

Abstract

This thesis will address the question of whether there are any situated cognition concepts of memory functioning as investigative kinds in the sciences of memory. Situated cognition is an umbrella term, subsuming extended, embedded, embodied, enacted and distributed cognition. I will be looking closely at case studies of investigations into memory where such concepts seem *prima facie* most likely to be found in order to establish a) whether the researchers in question are in fact employing such concepts, and b) whether the concepts are functioning well – functioning as investigative kinds – and should therefore continue to be employed, or whether something has gone wrong in the practice of the science and they should employ a different kind of concept. An historically situated approach to the case studies will allow me to answer part b) here.

Along the way, I will argue for a way of construing scientific research that I call the dynamic framework account, an account of (im)maturity for science, a variety of conceptual role semantics with respect to scientific concepts, and the historically situated case study-based method I will employ in answering the central question. My conclusions, and the way I reach them, constitute contributions to debates about situated cognition particularly, and to philosophy of science more generally, as well as recommendations for scientific practice.

Acknowledgements

I would particularly like to thank my supervisors David Corfield and Kristoffer Ahlstrom-Vij, and other members of the University of Kent Philosophy Department, especially Julia Tanney, and past and present members of the postgraduate community. Their feedback on all or part of this thesis has been invaluable. Thanks also go to conference and reading group participants, students, friends and family for inspiration and support of many and various kinds.

Contents

1. The Question

Introduction

In recent years, a number of perspectives on cognition which could be grouped under the umbrella term *situated cognition* have come to prominence. These perspectives, in one way or another, take into account the world beyond the brain – in which the brain is situated – when studying cognition. Such perspectives, including the theories of *extended cognition*, *embedded cognition*, *embodied cognition*, *enacted cognition* and *distributed cognition*, raise a number of perplexing issues, and one of these will be explored in what follows. I will be asking whether any situated cognition approaches are legitimately being employed in cognitive science, focussing particularly on situated cognition concepts of *memory*. Concepts are legitimately employed if they are functioning as *investigative kinds* (this term will be explained below). I will refer to this question (the title question of the thesis) as *the situated cognition question* for short.

I will be taking a philosophy of science approach, in particular using a historically situated case-study-based conceptual ecology to analyse the concept(s) of memory (MEMORY[1]) that scientists employ in their research. This method will allow me to look for situated cognition concepts of memory that have been in use in recent practice, and assess the extent to which they are legitimate.

This work will make a contribution to the philosophy of particular sciences that study memory, and also to the philosophy of science more generally. I hope that it will also be of use to scientists working on memory, particularly those trying to engage in the tricky but essential practice that is *interdisciplinary* scientific research. One of the biggest problems facing interdisciplinary research is the partial nature of communication between subdisciplines, which I hope can be better understood by the analysis of MEMORY in various case studies that is undertaken here.

In this chapter, I will clarify various aspects of the situated cognition question and its importance. The central aspect of the question is situated cognition theories, so that is where the chapter begins. I will then go on to justify my focus on *science*, justify

[1] I will follow the practice of using small capital letters to denote the concept (MEMORY).

my focus on scientists' *concepts*, and explain the notion of investigative kinds, defending it as the best way to frame my question.

Situated cognition and a question about memory

The best known of the situated cognition perspectives to most philosophers is *extended cognition*. I will therefore begin by explaining this position, before moving on to outline the others with reference to it.

Extended cognition

The term "extended cognition" was introduced in 1998 in a paper by Andy Clark and David Chalmers (Clark and Chalmers, 1998). According to the hypothesis of extended cognition, cognitive processes are extended beyond the head of the individual. For example, a player of Scrabble may rearrange the pieces in front of him to help him to work out possible words he could play next (Kirsh, 1995: 64–65; Clark and Chalmers, 1998: 8–10). On the brainbound view that extended cognition aims to displace, this rearranging of the tiles would be seen as an input to or output from the cognitive process which takes place exclusively in the head of the player. The hypothesis of extended cognition says that the cognitive process includes the rearrangement of the tiles. The process therefore takes place spread across brain, body and world. The tiles and the player's hands as he rearranges them are therefore as much vehicles of cognition as the neurons and synapses in his head.

Extended cognition proponents aim to change the way we see cognition (no longer thinking of it as brainbound), and also to alter the practice of cognitive science. According to the hypothesis of extended cognition, cognition is drastically changed or irreparably damaged in the absence of certain features of the environment, or when their manipulation by the agent is prevented. To go back to the Scrabble example, if scientists were to study the play of Scrabble, players should be allowed to physically rearrange the tiles as they work out what word to play next. If they are not allowed to do this, the usual cognitive processes of Scrabble playing will be damaged. Expecting people to play Scrabble effectively while not rearranging the tiles would be parallel to expecting them to play effectively with partial brain damage. In both cases, part of the substrate of cognition is damaged, so the cognitive process is impeded.

Of course, scientists could still study Scrabble without physical tile rearrangement, for example if they wanted to study the role of such rearrangement in Scrabble play. But this would again be parallel to the brain damage case: If one wanted to study the role of a certain part of the brain in Scrabble play, one could compare play in people with and without brain damage to this area. The point is that "normal" Scrabble play cannot be expected when cognition is impeded by refusing to allow physical tile rearrangement.

Talk about extended cognition refers to the extension of the physical substrate or vehicles underlying cognition. It therefore does not imply that cognition has a location, merely that its substrate does. In the most neutral terms, the idea is that whatever role the brain normally plays in cognition, parts of the external world can also play that role.

My question here concerns memory. Not only is memory studied by a wide range of disciplines, it is a central case in the extended cognition literature. One of the first and probably the most famous example of extended cognition centres on memory. This is the example of Otto and his notebook (Clark and Chalmers, 1998). It is designed to show that the mental state "belief" can be extended, but it also seems to show that the cognitive processes of storing and recalling memories are extended, and it has featured in the literature as an example of extended memory ever since (e.g. Clark, 2008; Sutton et al., 2010; Wilson 2005a; Heersmink, 2013).

The example concerns two people, Otto and Inga. Inga wants to go to an exhibition at the Museum of Modern Art (MoMA). She remembers that MoMA is on 53rd Street and sets off. Otto suffers from memory loss due to Alzheimer's disease, and so writes things down in a notebook, which he carries everywhere with him. He also wishes to see the exhibition at MoMA. He looks through his notebook, finds where he has written down that MoMA is on 53rd Street, and also sets off. We are meant to apply a principle of parity to this case and conclude that both Otto and Inga have the belief that MoMA is on 53rd Street, but that in Otto's case the belief is extended. In terms of memory, they have both remembered the relevant information, but in Otto's case the vehicle for the memory was external so the memory process was extended.

Here, Clark and Chalmers note that the Otto and Inga case is not necessarily about how we use the word "belief", but about how we *should* use it. Seeing the mind as extended is recommended to us as the simplest and most elegant description of these kinds of cases: 'By using the "belief" notion in a wider way, it picks out something more unified, and is more useful in explanation' (Clark and Chalmers, 1998: 14). This is a general point and also applies to memory; this argument in favour of extended cognition says that the sciences of memory would be able to provide simpler and more unified explanations if they accepted the possibility of extended cognition.

Three important points are revealed here. The first is that accepting extended cognition would have an effect on scientific methodology, because it recommends that scientists treat the inner and outer components as parts of a single cognitive system, rather than treating the outer components as mere inputs and outputs to cognitive processes. The extended cognition question is therefore of importance for scientists as well as philosophers.

The second is that much of the extended cognition debate centres on the *science* of memory, rather than our ordinary language notion of memory. I will follow this aspect of the literature and take a philosophy of science approach, the methodology for which will be outlined in chapter 5. The choice to focus on science will be defended further in the subsection "Why study science?" below.

The third point is that Clark and Chalmers' argument is a form of *inference to the best explanation* (Lipton, 1991); they advocate extended cognition because it would result in better explanations for science. My approach is in line with this aspect of the literature in a sense, and this too will be explained and defended further below, in the sub-section on so-called "investigative kinds".

Other situated cognition perspectives

In addition to the extended cognition story, there are other similar frameworks for cognition that could be used to give an account of memory. These include embedded, embodied, enacted, and distributed cognition. Here I am using "situated cognition" as an umbrella term for all of these perspectives that are opposed to the brainbound picture in similar ways.

Different proponents of these theories are not in complete agreement about their details, and attempts to taxonomize them precisely therefore vary. I will not attempt to contribute to this taxonomic project here; I will make use of Sven Walter's (2010: 63–66) working definitions of the different positions to outline the landscape of the debate, with one small alteration, and the addition of Edwin Hutchins' definition of distributed cognition. Walter's first four definitions are:

> *Embodied Cognition I…*: Cognitive processes are partially *dependent* upon *extracranial bodily* processes.
> *Embodied Cognition II…*: Cognitive processes are partially *constituted* by *extracranial bodily* processes.
> *Embedded Cognition…*: Cognitive processes are partially *dependent* upon *extrabodily* processes.
> *Extended Cognition…*: Cognitive processes are partially *constituted* by *extrabodily* processes.

A few points are worthy of note. First, although Walter uses "constituted" here, it is not clear what it could be for a *process* to be constituted. This should be read as explained above, that the *substrate* of the cognitive process is partly made up of extracranial or extrabodily materials, thus the cognitive process is spread over brain and body for embodied cognition II, and brain, body and world for extended cognition.

The use of "dependent" is also worthy of brief exposition: Exactly what it means may be a matter for debate (Walter, 2010: 63), but it is something like *ineliminable causal dependency*, which is weaker than depending for its existence (as this stronger kind of dependence would collapse embodied cognition I into embodied cognition II, and embedded cognition into extended cognition).

Finally, although Walter construes embedded cognition as dependence upon *extrabodily* processes, in practice this position is rarely if ever held without the accompanying belief that cognition also depends on *extracranial bodily* processes.

That is to say that someone who holds embedded cognition also holds embodied cognition I (although not necessarily vice versa). For my purposes here then, I will alter the definition of embedded cognition slightly, to read:

> *Embedded Cognition…*: Cognitive processes are partially *dependent* upon *extrabodily and extracranial bodily* processes.

The last of Walter's definitions is:

> *Enacted Cognition…*: Cognition is the relational process of sense-making that takes place between an autonomous system and its environment.

Enacted (sometimes called "enactive") cognition is part of a broader program of *enactivism* which focusses on the agent's interaction with the world, rather than manipulation of representations (see Noë, 2004; Varela, Thompson and Rosch, 1993).

Walter does not offer a definition of distributed cognition, but the theory can be credited to Edwin Hutchins, who characterises it as follows:

> [C]ognitive processes may be distributed across the members of a social group, cognitive processes may be distributed in the sense that the operation of the cognitive system involves coordination between internal and external (material or environmental) structure, and processes may be distributed through time in such a way that the products of earlier events can transform the nature of later events. (Hutchins, 2000: 1–2).

Distributed cognition differs from extended cognition in its lack of a central controlling locus of cognition. Extended cognition is *organism centred* (Clark, 2008: 139); it is always *someone's* cognition that extends. However, for distributed cognition, there is no such individual controller or owner of the cognizing. In Hutchins' example of the navigation of a ship (Hutchins, 1995), the cognitive process of navigating the vessel is spread across the team of people involved, their instruments, and through time. From an extended cognition perspective, this whole process of navigation would not count as cognitive. The only candidates for cognitive processes would be the contributions of particular individuals, though each may recruit instruments and possibly other people to fulfil their roles.[2]

Situated cognition and memory

One approach to the situated cognition question is to take an example of a state or process that everyone agrees is cognitive. Memory is one such paradigmatic example. If extended memory is possible then extended cognition is possible because, whatever exactly it means to be cognitive (and this is a subject of intense debate), memory is cognitive (and similarly for the other situated cognition perspectives).

The specific question I am interested in here then is whether memory is always brainbound, or whether there are any cases of extended, embedded, embodied, enacted or distributed memory, and what it would mean to claim this. This is an important question because how we conceive of memory is currently changing, while at the same time there is pressure on the sciences of memory to produce results.

New technology has the potential to change how we see memory, and many of the examples in the situated cognition literature trade on this. Modern technology is also giving rise to speculation that memory itself is changing, for example due to our reliance on search engines to find information and the prevalence of portable devices with which to access it (see for example Sparrow, Liu and Wegner, 2011). In part because of this, we are beginning to question how we see our relationship to older more established technologies, and the role of other people.

More specific issues about memory currently shaping the agenda for science include:

- A question of how young people should be educated to use their memories in a world where memory "offloading" to technology is ubiquitous.
- An increasing elderly population, creating a crisis in how to care effectively for large numbers of dementia patients. Changing the environments of these patients is one way of helping them (see Clark, 2003: 140), and how this is

[2] Hutchins (2011) reviews Clark's 2008 book on extended cognition, recommending the theory become more "enculturated". An enculturated version of extended cognition would be a view closer to distributed cognition.

researched is affected by whether the research is carried out in a situated cognition framework.

- How memory should be treated in law, for example research on false memory has had an impact on how eye-witness accounts are treated in court (see Elizabeth Loftus et al.'s "lost in the mall" experiments e.g. Loftus and Pickrell 1995; and discussion of Loftus' work in Heaton-Armstrong, Shepherd and Wolchover (eds) 1999, especially pp. 25–27). Of particular relevance is how memories are affected by communication with other people, and situated cognition perspectives may affect how this is construed.
- Making money from an understanding of memory, for example by making advertisements that are particularly memorable.

All of these challenges and changes create a shifting but high-pressure context in which memory research is currently taking place.

The sciences of memory are often referred to as immature sciences, but the questions they are asking and trying to answer are extremely important. I will develop my own account of what it is for a science to be immature in chapter 3, but for the purposes of this chapter, the notion can be left intuitive. In the terminology of Dudley Shapere, a philosopher whose work I will return to in what follows, the questions in this domain are *important*, but the science is not yet *ready*. Shapere says that

> …in some cases domains which are "important" enough need not be "ready" in order that their investigation be considered reasonable and appropriate at a certain stage…Briefly, if the problem is deemed sufficiently important, and has achieved a certain level of precise formulation, then investigation of the area in question is often considered at least marginally appropriate despite the "unreadiness" of current science in other aspects to deal with the problem. Such conditions may, indeed…serve as incentives to try to *make* the state of science ready to deal with the problem: research will be generated to increase the precision of data about the domain, to develop technology for doing so, and so on. (However, in some circumstances, especially when such technological developments do not seem feasible, investigation even of domains which are recognized as important will often be looked down on as excessively "speculative", sometimes to the point of being "unscientific".) (Shapere, 1977a: 533).

This is the situation in which the sciences of memory currently find themselves. I will explore this situation further in chapter 3, but here I want to note that it is important to understand how the sciences of memory work, (what is distinctive about them, etc.) for answering the situated cognition question.

The key question I am trying to address is whether memory is the sort of phenomenon that can include Otto-like cases, or other kinds of situated cognition cases. One way of putting this question is to simply ask "what is memory?" This seems like a basic question that is at the root of much scientific research into memory. I will argue that the scientific process of asking and answering questions of this form is not as simple as it looks. In fact I will argue that we cannot make sense of the idea of asking about "memory itself", shifting instead to an interpretation of the question which focuses on *scientists' concepts* of memory. The next two sections will begin to explain and justify this step, with a fuller exposition in chapter 2.

Why study science?

Why study the sciences of memory rather than ordinary language or "folk" talk about memory? One answer to this is just that my question is about the sciences of memory; questions about ordinary language are also interesting, but they are not my questions. However, there are some that would block this move by arguing that studying the sciences of memory is not an interesting or worthwhile project. I want to defend my project against this accusation.

There are two main groups of people who might make this accusation. The first is those who claim that the ordinary concept of memory is the object of scientific research on memory; it is what the science is trying to explain. We therefore cannot look at the sciences of memory in isolation from ordinary talk. The second group consists of those who claim that it is *memory itself* that we want to know about, and the ordinary concept is what fixes the reference to this reality, so we should study this ordinary concept. On this latter view, the ordinary concept is studied as a means to an end; on the former it is the object of investigation in its own right.

I will deal with the former view first. This view says that science cannot get away from ordinary mental concepts because they are what the scientists are trying to explain (Tanney, 2013: 291–292). On this view, what a term like “memory” means is normatively constrained by its use in memory-talk, and a science that came up with a concept that moved away from this memory-talk would be changing the subject and telling us, not about memory, but about some other construct of its own devising. A philosophy of science study of these new constructs would therefore have no merit in isolation from a study of ordinary talk.

While I accept the importance of the use of term “memory” for determining its meaning, I reject the conclusion that science is therefore bound by ordinary concepts in its investigations. It may use them as a starting point (where else could one start, given that scientists are ordinary folk too?) but it often diverges from them for good reasons. For example, the ordinary concepts HEAT and FLUID do not coincide with the way physics uses those concepts. Why is this?

The scientific concepts are embedded in networks of theory and practice that constitute fruitful science. What exactly it is for science to be fruitful is a question famously fraught with problems, but roughly it is science that gives good explanations, accurate predictions, and produces practical applications that are useful for society. I will say more about this in the next chapter, but for now it is enough to note that these are not necessarily the aims that ordinary people have in going about their day-to-day lives. In ordinary life we may need to do these things to some extent, but we do not need to make the same predictions as scientists, do not seek to produce the same applications etc. If we did, we would not need science. The uses to which terms like “memory” are put within scientific discourse may therefore be different to those in ordinary discourse.

To give an illustrative example from the sciences of the mind, consider the categorization of mental disorders. Here it may appear that both scientists and ordinary people have the shared aim of providing effective psychotherapy, so they should have the same concepts. However, there are also other aims that are not shared. For example, the scientists may want categories of mental disorder that fit in with theories from other disciplines such as neuroscience. This might result in categorizations based on brain lesions. Ordinary talk does not share this aim, so need not share the same concepts. Ordinary people on the other hand have concepts that have grown up to fulfil the aim of living with psychiatric problems in the absence of effective treatments. This aim may have led to explanations of psychiatric illness that shift the blame away from the patient (e.g. demonic possession), or shift blame onto them in order to justify harsh treatments (e.g. illness as punishment for wrongdoing). A science that does not share these aims need not share the same concepts. The networks of theory and practice produced by science to meet its aims can be different to those used by ordinary people, and so can the concepts that are embedded in them.

The opponent may accept what I have said so far, but insist that problems arise when discoveries about the objects of the new scientific concepts are announced as though they were discoveries about the objects of the concepts we use in everyday talk (Bennett and Hacker, 2003). I accept that, when this happens, it is a problem. However, more often it is not just a discovery that is announced, but also an accompanying conceptual revision. In these cases, ordinary language adopts a scientific term (e.g. "gene") or changes the meaning of an ordinary term to take on board scientific connotations (e.g. "star" in modern ordinary talk usually means something like our Sun, a very great distance from the Earth). Even when the ordinary language concept does not change in line with the scientific one, the concepts are often sufficiently different that there is little or no confusion (for example the cases of HEAT and FLUID).

I now come to the second type of criticism introduced at the beginning of this section. This was the argument that we should study ordinary language concepts because they fix the reference to the phenomena themselves, and it is the phenomena themselves we really want to know about. The idea of reference-fixing is familiar from Kripke (1972/1981) and is based on the idea that our folk concepts pick out kinds based on some underlying essence.

This idea is sometimes supported by arguments that our folk concepts must be accurate because they do not classify things randomly or miraculously (Jackson, 1998: 64[3]); our success in attributing mental states to others (our agreement on what

[3] I am not implying that Jackson would advocate not looking at science. I am only attributing the no-miracle argument in favour of conceptual analysis to him.

states to attribute and the fact that doing so allows us to predict behaviour) would be a miracle if there was not a pattern underlying our classifications.

To answer the situated cognition question on this view, one might think that the best thing to do would be to analyse our folk concept of memory, to find out what kind it fixes the reference to, and then to see whether situated cognition examples would qualify as members of this kind. Looking at science may be important for the second step of this process – the folk do not often know what the essential properties are which their concepts track – but again, science should not be looked at in isolation from ordinary concepts.

For the sake of argument in this part of the chapter, let us grant that science is in the business of carving up the world into kinds so that there is sense to be made of talk about what memory *really is*. (The next chapter will present a different picture of the scientific enterprise). Let us also grant that folk psychology is some kind of theory that also attempts to divide the world up into kinds, and that allows us to successfully interact with the world and one another.

My response, even given these concessions, is to point out that our success in using folk psychological concepts seems only to reduce the likelihood of our being *widely* mistaken in our categories. There is no reason why a few individual categories could not be completely wrong, in the sense that they fail to pick out kinds. It is widely agreed that we have mistaken categorizations in folk physics (for example a belief in a force of impetus that is given to moving objects when they are pushed, and mistaken intuitions about the path of an object in circular motion when it is released). These mistaken categorizations do not prevent us from interacting successfully with the physical world in the vast majority of situations. Similarly we could be completely mistaken in some of our mental categories.

The defender of studying folk categories may reply that there is something special about folk psychology that blocks the comparison with folk physics. Perhaps our first-person access to the mental means that we cannot be mistaken about folk psychology in the same ways we can be mistaken about folk physics. If this was right, it would show that not only is the study of folk concepts important, it is *more important than* the study of scientific concepts.

However, a brief consideration of the history of psychology will show that this response is a non-starter. Introspectionism in psychology was popular in the 19th and early 20th centuries, but the program broke down amid hopeless disagreement as to how it should be carried out and what observations resulted (see Kusch, 1999 for historical detail, and Kusch, 1999 and Varela, Thompson and Rosch, 1993: 32 for some possible explanations for this breakdown). If we had some privileged access to the mental, it seems this story should have been very different. There should have been no need to attempt to develop special techniques of introspection, and there should have been more agreement. The failure of this paradigm alone demonstrates that, even if there is something different about mental phenomena compared to physical phenomena, we do not have any special privileged access to the mental.

The history of folk psychology may also suggest a lack of privileged access to the mental. If Martin Kusch is correct that folk psychology changes over time (Kusch, 1997: 24), either we have no such special access, or mental kinds in the world have changed over time. This latter picture is unappealing to our opponent who needs a picture on which kinds are unchanging, to get the idea of reference-fixing off the ground.

There is no reason to think that ordinary people have any privileged access to how mental phenomena should be divided up that science lacks. There is therefore no reason to think that folk psychology is any less likely to be mistaken than folk physics, which we have seen can be radically mistaken.

Here I do not need to argue that science is *better* at describing the world; it is enough to justify my project that it gives *some* description of the world that is worthy of study in its own right, without an accompanying study of ordinary concepts. My position is therefore open to those who agree with John Dupré (1993) that folk categorisations of the world are no worse than scientific categorisations, they just meet different aims. It is also open to those who think that science is better at describing the world because it has evolved methods best adapted to achieving this.

Here, I only want to argue that there is no reason to think that science is *worse* at describing the world than ordinary language, or that scientific concepts can only be studied in terms of their relationship to ordinary concepts. This is enough to justify my focus on science as a legitimate project.

Why study concepts?

Why think that it is scientists' *concepts* that are important? My view on this will become clearer in the next chapter where I discuss how science works and the role that concepts play, but here I want to give some suggestion that there is a conceptual problem facing the sciences of memory. Given that the problem can be expressed as one about concepts, a study of concepts is a reasonable place to look for a solution.

The disciplines investigating memory are very diverse and include at least the following: neuroscience, psychology (including cognitive psychology, neuropsychology, social psychology, discursive psychology), artificial intelligence, human-machine interaction, sociology, philosophy, anthropology. Some of these disciplines overlap (e.g. neuroscience and neuropsychology), some may contain multiple subgroups studying memory that take very different approaches (e.g. bio-anthropologists and social anthropologists), and their diversity indicates a great breadth of study. When I refer to "the sciences of memory" I mean loosely all of the above disciplines, or at least those of their subgroups which have an interest in memory.

This breadth of study and wide range of approaches leads to plurality in the study of memory. Different concepts of memory and different sets of theories about it are in operation for different groups of researchers. Some divide memory into sub-categories (e.g. long and short term memory, semantic and episodic memory etc.) while others divide it into more or different sub-categories, and still others treat it as a unified concept.

In addition to this, "what is memory?" (or "what is long term memory?", "what is semantic memory?" etc.) are questions that could be answered with a kind of brain process, a specifically located brain state, a functional specification, a list of necessary and sufficient conditions (functional or otherwise), a disjunction (again functional or based on substance), a cluster of features (none of which is individually necessary), etc.. The interesting thing about the case is that it is currently very open which of these could be the case, because different scientific subdisciplines favour different options from this (incomplete) list. In later chapters, I will put this in terms of the sciences of memory being *immature sciences* at an early stage of the process of *conceptual development*.

This disagreement and disunity over concepts is something that troubles those working in the sciences. For example, there is a large amount of anxiety in psychology about the fragmentation of the discipline. The Journal of Clinical Psychology article "Psychology Defined" asks:

> What is psychology? Is it a single, coherent scientific discipline awaiting transformation from the current preparadigmatic state into a more mature unified one? Or, is it a heterogeneous federation of subdisciplines that will ultimately fragment into a multitude of smaller, more specialized fields? This is, in essence, the "to be or not to be" question of the field. Currently, psychology exists as an uneasy compromise between unification and fragmentation. On the one hand, the existence of numerous societal institutions suggests that psychology is a singular entity at some level. Academic courses, degrees, and departments, as well as organizations like the American Psychological Association (APA) suggest that the concept of psychology is a specifiable, coherent entity (Matarazzo, 1987). On the other hand, a more detailed inquiry reveals a remarkable degree of confusion, fragmentation, and chaos at the theoretical level. So formidable is the problem of conceptual incoherence that several prominent authors have flatly stated that it is insurmountable (e.g., Koch, 1993). (Henriques, 2004).

The only things holding the discipline together according to this quotation seem to be institutions and organisations, not concepts of the phenomena or the way they are investigated. Different scientists studying the same phenomenon have different concepts of that phenomenon embedded in different networks of theory and practice. This variation is particularly pronounced between different subdisciplines (hence the anxiety about fragmentation).

In the first chapter in a collection of papers entitled *Science of Memory: Concepts*, Dudai, Roediger and Tulving acknowledge that the problem of interdisciplinary communication about memory is a problem with concepts specifically. They say:

> We believe that with proper attention, communication and interdisciplinary understanding can be improved. We also believe that a direct confrontation of the issue at what we regard as the most fundamental level of knowledge and analysis – the conceptual level – is the best approach. The discussion of concepts in contemporary science of memory is underdeveloped. Some exceptions notwithstanding (e.g. Tulving 2000; Dudai 2002…), most practicing students of memory seem to shy away from spelling out and debating the concepts that form, or should form, the foundations of their own science. (Roediger et al., 2007: 1).

They then go on to briefly review some arguments for and against the focus on concepts, concluding that the arguments *for* outweigh those *against* (Roediger et al., 2007: 3–5).

The aim of the book is to lay the foundations for 'a new science of memory' on the assumption that '[f]or the practitioners of the science of memory to be able properly to exploit, and benefit from, the rich multidisciplinarity of methods and findings, they must understand the language and *modus operandi* of their colleagues in other subdisciplines. Such understanding is a *sine qua non* of the success of the venture' (Roediger et al., 2007: 1).

Much of the literature on memory refers to the problem of variation in MEMORY between different subdisciplines (e.g. Sutton, 2004: 188; Wilson, 2005a; Welze and Markowitsch, 2005: 64–65; Figdor, 2013). Sutton (2004) gives a good overview of the scale of the problem: 'How could the concepts, models, or practices of such glaringly incompatible activities as clinical neuropsychology and media theory, or developmental psychology and Holocaust studies ever be imported into neighbouring discursive universes? More to the point why would anyone bother?' (Sutton, 2004: 187). He goes on to discuss why and how we might bother, concluding 'I hope that this provides sufficient motivation for trying to show how such different memory researchers—from neurobiology to narrative theory, from the developmental to the postcolonial, from the computational to the cross-cultural, might one day be able to talk to each other.' (Sutton, 2004: 211).

But even if you are less optimistic than Sutton for the prospects of communication about memory, the extent of the conceptual diversity here should still be a cause for concern. Even for more specific subtypes of memory discussed by more closely related subdisciplines, there is variation. (For example, concerning the variation in COLLECTIVE MEMORY specifically, see Hirst and Manier, 2008: 183; Wessel and Moulds, 2008: 289; Wertsch and Roediger, 2008: 318; Barnier et al., 2008: 36–37). Even within a subdiscipline, there is polysemy; in a paper about the Cognitive Atlas Project which I will consider in detail in chapter 2, Poldrack et al. give three different definitions of working memory, all of which are found within neuroscience (Poldrack et al., 2011:1). All of this conceptual disunity indicates that the problem is a conceptual one.

It is important to note at this point that when I talk about scientists having "different concepts", I do not mean completely different non-overlapping concepts. There is still enough in common for communication to take place, albeit sometimes partial and with difficulty. I return to the issue of concept individuation and communication in chapter 4.

The situation of conceptual disunity makes it appear that there might be no unified answer to the situated cognition question for memory. Perhaps some subdisciplines use a concept of memory that would be amenable to an extended reading, others a distributed reading, and others a brainbound one. Here I will be interested in finding out whether this is the case and, if it is, whether that is a situation that should be allowed to continue, or whether conceptual unity should be sought. One powerful reason to look for such unity would be to facilitate interdisciplinary communication. This is important for the sciences of memory, and also for the cognitive sciences, since cognitive science is an interdisciplinary enterprise.

One possible way of tackling this question which may look promising is to make use of the idea of natural kinds. The variation in MEMORY discussed in this section may make it look impossible that memory could be a natural kind, but there is a pluralist account of natural kinds available which looks like it might be a good fit. I will argue that in fact this is not a helpful way to construe the question, opting instead to put it in terms of Brigandt's notion of *investigative kinds* (Brigandt, 2003). This is the subject of the next section.

From natural kinds to investigative kinds

There are two prominent strands in the extended cognition literature, one I will describe as metaphysical, and one epistemic. The two strands are intertwined, but I distinguish them here because I want to argue that the best strategy is to focus on the epistemic strand, while remaining neutral on the metaphysical issue.

What I am calling the metaphysical strand is a line of enquiry about what cognition (or particular cognitive states or processes like memory) actually *are*. In this debate, the problem of whether cognition can extend is often put in terms of finding demarcation criteria for the cognitive (Kaplan, 2012) or “marks of the cognitive” (Adams and Aizawa, 2001, 2008). The idea is that if we know what criteria have to be met in order for something to count as cognitive, we can see whether putative examples of extended cognition meet these criteria. Such criteria have proved elusive, and there is debate over whether there even are such criteria (for arguments that there are not, see Hurley 2010, Ross and Ladyman, 2010).

In places, Clark himself seems to subscribe to something like this form of argument. For example, he claims that genuine cognitive processes are individuated according to their proper function of providing information, that is, in virtue of their selection history. Alternatively, they may be a side-effect occurring within a mechanism that was selected for this function (Clark, 2008; see Allen-Hermanson, 2013 for discussion). In a sense, he is here giving a teleological mark of the cognitive.

What I am calling the epistemic strand has already been mentioned above; this is the view that construes the debate in terms of inference to the best explanation. The argument is that particular situated cognition approaches result in better explanations for science than do brainbound alternatives.

In places, Clark seems to endorse this means of arguing. For example, in the discussion of Otto’s notebook quoted earlier in this chapter, Clark and Chalmers say ‘[b]y using the “belief” notion in a wider way, it picks out something more unified, and is more useful in explanation’ (Clark and Chalmers, 1998: 14; see also Clark, 2008: 80 for a similar style of argument in response to a common objection to the Otto story). In places he sounds almost entirely instrumentalist, for example in his discussion of the debate between the Hypothesis of Extended Cognition (HEC), the Hypothesis of EMbedded Cognition (HEMC), and his solution, the Hypothesis of Organism Centred cognition (HOC), he says:

> HEC, HEMC, HOC? We should not feel locked into some pale zero-sum game. As philosophers and as cognitive scientists, we can and should practice the art of flipping among these different perspectives, treating each as a lens apt to draw attention to certain features, regularities, and contributions while making it harder to spot others or to give them their problem-solving due. (Clark, 2008: 139).

The relationship between the metaphysical and epistemic ways of construing the debate is complex, but many believe that in order to be useful in explanations and inductions in science, a kind must track a real grouping in the world. Here I argue that there is no need to presuppose this. It is the epistemic issue that matters – the issue of finding out which concepts result in successful science – not what you assume must be true about the world to cause that success.

A natural way to put the metaphysical question is in terms of whether memory is a natural kind. It may seem that it is an important question for scientists working on memory whether or not it is a natural kind. As Kourken Michaelian says:

> If memory is a natural kind, then it will often be profitable to investigate the various memory phenomena together, as if they constitute a coherent whole, for then investigation of one kind of memory will often tell us something about features of other kinds of memory. But if memory is not a natural kind, then it will not in general be profitable to investigate the various memory phenomena together, for then investigation of one kind of memory will not typically tell us anything about the other kinds of memory; in other words, if the answer to the question whether memory is a natural kind is negative, then we should not be aiming for a general theory of memory. (Michaelian, 2010: 171).

Not only is there a question about the overarching kind "memory", there are also questions about whether various types of memory (e.g. semantic memory, episodic memory) are natural kinds, even if the overarching kind is not.[4] At this stage however, there is not even agreement on how to taxonomise these putative sub-

[4] For the idea that the types of memory may be natural kinds even if memory is not, see Rupert (2013: 39).

kinds. Even text books note that there is some disagreement over the taxonomies they give (see e.g. Toates, 2007: 282; Martin, Carlson and Buskist, 2009: 292, 305, 332); philosophers also recommend changes of taxonomy (e.g. Michaelian, 2010). Answers to the natural kinds question about sub-kinds, like the question about memory as a whole, would seem to have an effect on how the science should proceed – which phenomena should be investigated together, etc.

However, this is mistaken; in fact it is not an answer to the metaphysical question that is important here, but an answer to the epistemic question. What will decide this issue is how well the sub-kind concepts play their roles in science – how they function in explanation etc. – and whether there is a role for an overarching kind to play. Whether you think that the success of a concept in allowing fruitful science to proceed is down to its picking out a real division in the world is irrelevant to whether that concept should continue to be used; the success alone is enough to recommend this continuation. In this way, the metaphysical question is reduced to the epistemic one in practice.

Some accounts of natural kinds point this way, with a distinctly epistemic slant. One example of this is probably the most common and most plausible account of natural kinds applied to the special sciences: Richard Boyd's Homeostatic Property Cluster (HPC) theory (e.g. Boyd 1991). It is this kind of account that Michaelian, quoted above, is working with, as are others cited in this dissertation, including Paul Griffiths and Ingo Brigandt. Due to its epistemic slant, and its popularity with many people I cite here, Boyd's theory is worthy of further discussion.

An HPC kind, as its name suggests, is defined by a cluster of properties. To be a member of the kind, an instance does not have to share all of the properties in the cluster: '[m]ost of the kind members possess most of these properties, but none of the properties in the cluster has to be shared by all kind members, permitting variation among the members of an HPC natural kind' (Brigandt, 2011: 3). The "homeostatic" refers to a mechanism that makes it such that the kind members share the properties in the cluster. For example for the kind species, the mechanism is whatever makes it such that cats all share cat-features (interbreeding according to one version of the modern species concept (Boyd, 1991: 142)).

The definitions of HPC kinds are determined *a posteriori*, not by social convention. This is because it is the underlying causal structure of the world that makes it such that the kind members share the properties which define them. When we group kind-members, we are therefore tracking something in the world, not merely agreeing on a convention. What holds the kind together is causal, not conceptual or conventional (Boyd, 1991: 129, 141). This makes it appear that the theory of HPC kinds is an answer to the metaphysical question. However, HPC natural kinds can be artificial or social kinds; *any kinds involved in induction and explanation are included* (Boyd, 1991). This is because Boyd takes science to require kinds that are projectable in the sense of Goodman (1955). These are the kinds that can support induction from one case to another, and this also grounds explanation. Again, we see that answering the epistemic question is primary, despite the extra metaphysical commitment made by Boyd and his followers. I therefore advocate focussing solely on this epistemic question and avoiding the extra commitment. This is perhaps more an alternative reading of Boyd's theory than an alternative theory, but I think it is an important one.

There are a few reasons to prefer this epistemic approach to the issue:

- Making fewer presuppositions is a virtue in itself, particularly at the beginning of an investigation.

- The sciences of memory are immature sciences; as noted above, it is too early to be sure whether memory will come to be defined by a kind of brain process, a specifically located brain state, a functional specification, etc., or several of these. It is not clear that all of the options would be amenable to an HPC reading, or any other natural kinds reading for that matter. This makes the metaphysical question a very hard one to answer (possibly impossible without further science). Given this, avoiding making metaphysical presuppositions in advance is a virtue in this specific case.

- The HPC theory of natural kinds is pluralist, and the epistemic question about which concepts would allow the best explanations etc. is therefore going to be needed anyway in order to decide which (one or more) of the several possible paths it leaves open should be followed.

According to the HPC theory, a natural kind is a group whose members share properties due to an underlying mechanism, and this can be interpreted in a pluralist manner: 'Pluralism about natural kinds would seem the most prudent option: there are as many kinds as there are distinct and dissociable mechanistic entanglements (Boyd, 1999[a], and Wilson, 2005[b], both embrace this kind of idea). What counts as a maximally refined scientific kind depends on which mechanistic entanglement is most relevant to the practical or scientific project in which one is engaged.' (Craver, 2009: 584).

We could agree on a pluralist answer to the metaphysical question (that extended, brainbound, distributed etc. concepts of memory could all be HPC natural kinds), and this would still not tell scientists which of these concepts they should work with given the demands and resources they have in their particular research. Helping them with this would involve answering the epistemic question; which concepts play a role in good explanations is relative to what you are trying to explain, what you have available to explain it in terms of, what kind of explanation would meet the aims of your research, etc. (see chapter 2 for more on this way of looking at the scientific enterprise).

It seems that the question of whether memory is a natural kind in the metaphysical sense is not the important one here. The important question is about what a science using a situated cognition concept of memory would be like – whether it would allow us to make good explanations and predictions and produce useful practical applications. This is how the question is often put in the extended cognition literature (e.g. Clark and Chalmers, 1998: 9–10, Clark, 2010: 49–52) and no metaphysical account of natural kinds is required to address it.[5]

The issue is about the role the concept of memory plays in scientific theory and practice. In fact, both Griffiths and Brigandt (Grittiths, 2004; Brigandt, 2011)

[5] In fact Clark, 2010 talks instead about "scientific kinds", which, like the term "investigative kinds" that I favour, is a more neutral term that does not carry a metaphysical account of kindhood.

interpret HPC theory primarily in this way – as an epistemic issue about concepts – although they do also make the assumptions about underlying mechanisms that I am remaining neutral about. I think it is worth ceasing talk about natural kinds and following these philosophers in a change in terminology:

> A more neutral substitute for 'natural kind' that carries many of the right connotations is 'investigative kind' (Brigandt 2003). This term highlights the fact that the emphasis in the model of natural kinds outlined above [Boyd's HPC theory] is an open-ended investigation. A natural-kind *concept* is a concept that it makes sense to seek to clarify through empirical inquiry. Such concepts are ongoing projects of inquiry in which extension and intension are altered to preserve inductive and explanatory power. (Griffiths, 2004: 907).

This way of putting things has a better fit with my description of how science works in the next chapter, where I will talk in particular about conceptual change over time. I will also return to it in chapter four, where I discuss concept individuation.

Brigandt says of investigative kinds, '[t]he crucial feature of my account…is to stress the fact that the species concept as an investigative-kind concept may be subject to *conceptual change* based on empirical findings. Investigative-kind concepts exhibit the *open-endedness of scientific search*.' (Brigandt, 2003: 1310, note 1, my emphasis). Concepts are particularly important given my characterisation above of the problem as a conceptual one; the problem is an epistemic one about which concepts will best fulfil the roles scientists need them to fill. Conceptual *change* is particularly important for my question, given the shifting context of memory research in the modern world, as I said when I introduced the question. I will say more about the dynamic nature of science in chapter 2, and conceptual change in chapters 4 and 5.

Given the aims and focus of my project, the concepts that are legitimate for the sciences of memory are those functioning as investigative kinds in the sense set out in this section.

Conclusion

The situated cognition question for memory is particularly difficult given conceptual disunity about memory in the sciences. The best way to tackle the question without presupposing a particular outcome to this conceptual disunity is to phrase it as an epistemic question in terms of investigative kinds, instead of as a metaphysical question.

Whether or not MEMORY or its sub-categories are investigative kinds is the question of whether the kinds can support fruitful science, an epistemological question about how the concepts function in scientific theory and practice (e.g. in giving explanations). By focussing on the epistemological question we can get away from the need to speculate about metaphysics, and move towards making concrete recommendations for scientific practice. I will therefore sometimes refer to concepts of memory being "legitimate", but never to them being "correct". The concepts that are legitimate are those functioning as investigative kinds in the sense set out here.

My question for this project is "are there any situated cognition concepts of memory functioning as investigative kinds in the sciences of memory?" In this chapter, I have clarified what is meant by "situated cognition", "investigative kinds" and "the sciences of memory", and justified my focus on science, and my focus on concepts.

2. The Dynamic Framework Account of Science: How Science Investigates

Introduction

This chapter will set out an account of scientific investigation that I call "the dynamic framework account". It construes scientific investigation as an ever-changing and interlinked network of theory and practice, and is inspired in particular by the work of Dudley Shapere. The section on concepts also makes use of the work of Paul Griffiths. Later chapters will analyse the sciences of memory in terms of the dynamic framework account, hopefully confirming the usefulness of the picture.

One of the main things the account will enable me to do is to talk about the development of concepts (in particular MEMORY) over time. This will be important for assessing the development of any situated cognition concepts of memory that I find in my investigation, as I will explain in the methods chapter.

In addition to presenting a useful picture of scientific investigation, part of the purpose of this chapter is to clarify certain presuppositions I make about science. The first section of the chapter will be concerned with these presuppositions. Here I cannot consider in detail all the arguments for and against what I say, but I want to make clear what is assumed in what follows so that it is clear which aspects of my final conclusion are outcomes of my investigation, and which are assumptions that were built in from the start.

The main part of the chapter consists of a schematic outline of the dynamic framework account, followed by an example experiment that illustrates the account ("Effects of caffeine on learning and memory in rats tested in the Morris water maze", Angelucci et al., 2002), in particular how it handles interdisciplinary communication. The final part of the chapter is a discussion of concepts within the framework, applying the account to some example scientific concepts (WATER, GENE, and mental concepts in cognitive neuroscience). This will pave the way for my discussion of MEMORY in later chapters.

The aims of science

In very broad terms, science aims to predict and explain phenomena, and to produce useful practical applications (a form of control of the environment). These ends are achieved by a variety of practices including designing and carrying out experiments, and producing scientific theories. I take it that these aims are widely agreed upon, although details of how to cash them out are much disputed.

One important aspect of this dispute is whether science aims at truth or not. This is the debate between scientific realism and anti-realism. There are many varieties of both of these doctrines and the debate is too complex and involved to get into here for reasons of space and focus. However, there is one important presupposition I make: that we have no reason to think that there is only one true description of the world that science is converging on. In other words, it may well be that there are multiple equally correct ways of describing the world.

Some realists may take my position to be a form of anti-realism, and here I am not going to dispute that claim. I refer to my position as realist because I assume that there is a mind-independent world that constrains which theories and practices will produce successful explanations, predictions and applications. Therefore we do not have unrestricted proliferation of theories and practices, or proliferation restricted only by social and political whims. If this is too minimal a claim to constitute realism in the eyes of some readers, they may refer to my position as anti-realist without effect on what follows.

The position does run counter to certain kinds of realism. The view it opposes is explicit in positions such as convergent realism (e.g. Hardin and Rosenberg, 1982; Putnam, 1982), but is implicit in much of the tradition (e.g. in talk about verisimilitude or truthlikeness). However, it is not an essential part of many of these accounts that there is only one account on which science converges; this was simply an assumption in much of the early debate. My assumption is therefore not necessarily incompatible with much of what is said by philosophers subscribing to various kinds of realism.

My view is derived from the position that John Dupré calls *Promiscuous Realism* (see especially Dupré, 1993). Promiscuous realism is *epistemically pluralist* in the sense that it allows that there could be multiple equally correct accounts of the world. I leave it open at this stage whether science is in practice converging on a single account (even though others are conceptually possible) or whether different parts of current science are making use of different accounts (compare Cartwright, 1999). Although later parts of my investigation will bear on this issue, I do not presuppose a position.

Even if different parts of science make use of different accounts, it may be that it turns out to be most fruitful to work out the relationships between these accounts and have them interact with one another. As Rick Dale says '[the pluralist approach] may also recommend integrating these competing theories in meta-theoretical frameworks that would sustain their co-existence' (Dale, 2008: 156). Much of the literature on pluralism discusses potential ways of doing this (e.g. Sandra Mitchell's Integrative Pluralism (Mitchell, 2002, 2003); from a *New Ideas in Psychology* special issue on pluralism (Lamiell et al., 2010), Goertzen's Dialectical Pluralism, and Smythe and McKenzie's Dialogical Pluralism are attempts at a similar thing). We do not know what a successful pluralist picture of the cognitive and social sciences would look like—what degree of integration between different accounts of the world there should be etc. so this is also something about which I make no presuppositions, although my later investigation will bear on the issue.

In summary, I assume that science is aiming to find *one or more full or partial accounts of the world that allow prediction, explanation and control*, rather than aiming to find a single complete and true account of the world. I will now go on to discuss the process by which science works to meet its aims in terms of the dynamic framework account.

Subdisciplines, domains and frameworks

Domains

Modern science splits the world into domains for investigation. This is what Shapere calls the *piecemeal approach*. Science did not always take this approach. As Shapere says,

> This *piecemeal approach to the knowledge-seeking enterprise* replaced an older *holistic approach* wherein all phenomena were to be dealt with at a stroke, so to speak - attempting,

> for instance, to explain the nature of "change" or of "substance" *in general*, or to determine the necessary conditions of knowledge *in general*, an approach which had been widespread in the middle ages. (Shapere, 1984a: 641).

It is unimaginable that science today would tackle any topic so broadly, and specialisation has if anything increased since Shapere wrote that paper.

Domains then are a very important feature of how science works. What exactly is a domain? Shapere tells us the following:

> A domain is an association of "items" of putative information into *areas for investigation*, having the following characteristics: (1) the association is based on some relationship (or putative relationship) between the items; (2) there is something problematic about the body of information so related; (3) that problem is an important one. Contrariwise, the domain constitutes a *domain of responsibility* for a theory of it: the theory is expected to account for the items of the domain fully and well. (Usually also, in order to count as a *scientific* domain, (4) science must be "ready" to deal with the problem.) In sophisticated stages of science, domains often have names such as "solid-state physics", "rare-earth chemistry", galactic astronomy"; but the every-day work of science is done with subject-matters - domains - still more specialized than these. (Shapere, 1984a: 641).

So a domain is a collection of interrelated pieces of information about which there are scientific problems. The phenomena of interest to a subdiscipline constitute a domain, as can narrower subjects of interest. “Memory science” is not a unified discipline of any kind and memory is too broad and heterogeneous to be a domain in Shapere’s sense. It is made up of bits of other domains which overlap and interrelate in complex ways which have yet to be fully worked out (and may not be). However, the subjects of many of the subdisciplines studying memory would count as domains, and certain types of memory would count as narrower domains within these subdisciplines (e.g. semantic memory in cognitive psychology would be a domain). More specific subjects of particular research programmes would also be domains (e.g. the spatial memory of rats solving a particular type of maze under certain conditions). It therefore seems that domains can be construed at different levels, and I will return to this idea below.

Domain formation is itself an achievement; as Shapere says, 'nature does not come to us neatly and obviously packaged into unalterable areas for investigation.' (Shapere, 1984a: 649). However, this is not an achievement that, once reached, necessarily remains accomplished. Domain formation and reformation can be an ongoing process:

> Old domains were split ("salts"; arguments about the unity of the subject of "electricity" in the eighteenth and early nineteenth centuries), items of experience and belief rejoined into new domains (electricity and magnetism, and the subsequent history of the interpretation of the electromagnetic spectrum), entirely new ones formed (halogens; M-type stars - and note the later split of that domain into "giants" and "dwarfs"; leptons and hadrons). What were formerly considered bases of classification (shininess, color, crystalline form) came to be dismissed as superficial, and other bases, previously considered superficial or not noticed or known at all, became fundamental (valency, spin, strangeness). (Shapere, 1984a: 650).

It can already be seen from these examples of Shapere's that this lumping and splitting of domains has an effect on scientific concepts. For example, when the domain of "salts" was split, the concept SALT changed. There is currently some dispute over how MEMORY is to be split and what the relationships between the different subtypes are. This is closely related to the problem of how domains in the sciences of memory (individually and collectively), should be organised.

Frameworks

Here I want to delineate, not just problems and concepts, but all the associated theories and practices that make up the piece of scientific investigation in a particular domain. This includes methods, experimental apparatus, equations, laws or generalisations etc. Here I will refer to this network of theory and practice as a *framework*. All of the elements in a framework are related to one another, and changes in one aspect will change other aspects. I will discuss this dynamic aspect of the framework below.

Framework is a notion somewhat similar to a Kuhnian paradigm (Kuhn, 1970, 1996), or disciplinary matrix (Kuhn, 1977), but without Kuhn's revolutionary account of scientific change, as I will explain in the next sub-section. It is also similar to the notion of a research program for Lakatos (1968), a research tradition for Laudan (1977), or a research framework for Von Eckardt (1993). The account is not necessarily incompatible with all of these (determining this for certain would be another project), but it allows me to talk in useful ways that are not obviously opened up by other accounts. In particular, my framework account is designed to allow me to focus on concepts, and to be construed at different levels, as will be made clear in the remainder of this chapter.

The framework can be pictured as a network, where the concepts, methods, goals, theories etc. are interconnected nodes, each of which constrains or applies pressure to the surrounding nodes. I am focusing on MEMORY so I am interested in concepts as nodes. However, it is important to remember that concepts are not the only kinds of things that can be nodes; the picture I am advocating is therefore distinct from conceptual role semantics as it is usually construed, although that theory also pictures an interrelated network of concepts. I will discuss the idea of my account as a variety of conceptual role semantics in chapter 4.

The questions a science chooses to ask are one kind of thing that can be nodes, and they are shaped by surrounding nodes in the framework, or in other words they are not independent of the scientific investigation itself (its theories and practices). As Hasok Chang points out, ‘it was one of Kuhn’s main points regarding incommensurability that different paradigms have different lists of problems they consider legitimate and important.’ (Chang, 2012: 19). Chang goes on to show that some different problems were considered important by phlogistonists and oxygenists during the Chemical Revolution (see summary table on p.20). This example nicely illustrates how having a different theory about the part of the world you are interested in leads you to focus on certain questions and neglect others. How the domain under investigation is construed is therefore part of the framework, because which problems are investigated is a result of pressure from surrounding theories and practices, as well as applying pressure to them.

Some current problems of interest in memory sciences were mentioned in chapter one, including how memory-use should be taught in schools and universities, and an increasing elderly population and consequent larger numbers of dementia patients. As discussed in chapter one, these questions arise in part because of increased availability and use of technology, and may be best solved using one of the situated cognition concepts of memory. The questions can thus be seen to arise because of pressures from social goals and other theories within the framework, and to have an effect on the concept of memory employed in answering them.

As well as the questions asked, the *standards* for what counts as good answers to those questions are not independent of the other aspects of the framework. *What counts as an observation* and *what counts as a good explanation* depend on the surrounding theory and practices.

The idea that observation is theory-laden is familiar (Hanson, 1958; Kuhn, 1970, 1996) and this can be put in terms of the framework account: What counts as an observation is shaped by pressure from relevant theories and practices in the framework. We can say the same about what counts as a good explanation. Other parts of the framework particularly affecting standards for a good explanation include theories already accepted by the scientists in question, the aims of the science within society, and whether fit with neighbouring subdisciplines is required (which often depends on whether collaboration is needed).

This may look problematic, because it lacks any mention of a good explanation as one that gets the world right. However on my view, any such standard must be implicit in criteria such as meeting the aims of the science, and fit with other disciplines. The world constrains what could meet these criteria, so it is already part of the picture; it is not the case that any standards would do. However, there is no such thing as fitting the world *simpliciter*, just as there is no such thing as a good explanation *simpliciter*. These things are relative to a framework.

What Shapere says about success is helpful here. According to Shapere, success for a theory is about accounting for its own domain.[6] This is the kind of local success that I am interested in here. For Shapere, theories that are both successful and free from specific and compelling doubts – doubts which arise from beyond the domain but

[6] Although Shapere says that to be successful is not the same as to be correct, and he still seems to be operating with a domain-neutral notion of “correct”, even if not of “success” (Shapere, 1984a: 643–645).

which are not general like radical sceptical doubts – are admissible as ‘legitimately usable background beliefs’ or ‘background information’ for the science (Shapere, 1984a: 645). These are the things the science can rely on in future research.[7]

The scientists working in a particular subdiscipline or research cluster constitute a social community with a shared framework. Their shared interests partly constrain what it is rational for them to ask questions about as well as the standards for what constitutes good answers to those questions. But the community of the subdiscipline is also embedded in and related to the wider scientific community, and society more broadly, and these wider communities’ needs and interests also constrain what it is rational to investigate. I noted above that domains can be thought of at different levels, depending on the level of community being considered, and we can now see that the same is true of frameworks. Here I am focussing primarily on the level of the subdiscipline, and the shared framework for investigation at that level. However, it would also be possible to investigate at the level of bigger disciplinary groups, or smaller research clusters etc. As I said above, these different groups usually investigate domains at different levels, e.g. semantic memory for a subdiscipline, or spatial memory of rats solving a particular type of maze under certain conditions for a particular research group.

The framework taken to be shared between the participants would be construed at a coarser level of grain at the subdiscipline level compared to a finer level of grain for the research group.[8] Frameworks that are taken to be shared at higher levels will therefore no longer necessarily be shared at lower levels because they are being construed at a finer level of grain. This allows us to explain why interdisciplinary communication can appear successful, yet problematic. It is successful at the level of the overarching discipline, where frameworks are shared. However, described at the level of the subdiscipline, communication is much more partial because the more fine-grained frameworks at this level are not completely shared. The example below will illustrate this.

[7] For Shapere, sciences are rational to the extent that they rely on this kind of background. This aspect of the picture will become important later when it comes to assessing science, but for the moment I hope just to present the framework account as a useful picture.

[8] These levels are not intended to imply any kind of reductionism, and it is of course possible to be members of the same research group while coming from different disciplines.

The framework as dynamic

One of the most important aspects of science to account for is *scientific change*. The framework should therefore be thought of as dynamic. Its change is non-teleological because, according to the version of epistemic pluralism explained above, there is no reason to presuppose a single “correct” endpoint.

One of the most important accounts of scientific change is Kuhn’s notion of paradigm change in scientific revolutions (Kuhn, 1970, 1996). However, it has since been argued that change in science is often not as revolutionary as Kuhn described. A major problem is that if Kuhn’s account was correct, it would be difficult to explain consensus among scientists. This is because paradigms, like my frameworks, consist of the goals and standards of the science, amongst other things. Paradigms are therefore each successful according to their own standards of success, so it seems that paradigm shift could only be a matter of taste.

One solution to this is to make change less holistic. For example, Laudan says that

> [W]e solve the problem of consensus once we realize that *the various components of a world view are individually negotiable and individually replaceable in a piecemeal fashion* (that is, in such a manner that replacement of one element need not require wholesale repudiation of all the other components) (Laudan, 1984: 73).

Laudan offers what he calls the “reticulated model”, consisting of methods, theories and aims, and the relations between them (Laudan, 1984: 63). The dynamic framework account takes this a step further by having more than three components. One of its advantages is that other things can be construed as nodes, depending on the focus of your investigation, with concepts being of particular interest here.

This makes the interrelations between the nodes even more complex than the interrelations between Laudan’s three components. As Laudan says, components can be replaced or negotiated in a piecemeal fashion rather than in a full-scale revolution. However, on my view, this replacement or negotiation will be constrained by the surrounding nodes in the framework and some of them may have to be changed also.

The specific questions asked by a particular science (what exactly it seeks to explain, predict and control) is the most obvious thing that will change over time. This is partly because the social and political context (e.g. funding availability, the needs of society), are always changing. For memory, the questions currently concern the issues discussed above, e.g. increasing numbers of dementia patients, but these have not always been the main problems faced by memory research, and there is no reason to think that they always will be.

Differing focuses for memory research leads to the domain being conceived of in different terms and carved up differently between different subdisciplines. Because domains and their theories change, what it is to be successful changes too (Shapere, 1984a: 651). Discussing the history of chemistry, Shapere says that

> [t]he very goals of science altered: seventeenth and eighteenth century chemistry passed gradually from a goal of perfecting matter to one of understanding material substances in terms of their constituent parts, the arrangements of those parts, and whatever it is that holds those parts together - that "compositionalist" goal itself having been altered in profound ways in the succeeding centuries. (Shapere, 1984a: 650).

An explanation in terms of composition became the type of explanation that was considered successful where it would not have been before. This criterion for a good explanation was something that emerged over time from the science itself and developed from the body of theory and the concepts employed in it.

Chang also discusses the turn to compositionalism (compositionism in his terminology) and claims that it was the most important factor in the triumph of oxygen over phlogiston (Chang, 2012: 37). This example shows how the type of explanation a science looks for can in turn affect the theories and concepts employed in that science. The concept OXYGEN changed as theories changed, from the essential component of acids, or “acidifying principle”, to the concept we know today. I will discuss the role of concepts in the dynamic framework, and consider further example concepts, in the final section of this chapter.

In summary, changes in one part of the framework will have an effect elsewhere. For example, once a particular question has been asked and investigated, the results of this investigation and associated techniques etc. that may have been developed also affect the possibilities for future science. They become part of the background, alongside the pressures from society and science in other domains, which constrains which questions are seen as important to tackle next, and what would count as good answers to those questions. This replaces the idea of a revolutionary wholescale change of framework with something that makes better sense of actual change in science.

Using my account, we can look at change in science with a particular focus, for example on a concept. The ability to focus on different things as nodes is one important kind of flexibility my account offers, and the ability to focus at different levels is another. I now return to that feature of the account with an example.

Frameworks and levels: An example

As an illustrative example, consider the scientists working on the experiments detailed in the 2002 paper "Effects of caffeine on learning and memory in rats tested in the Morris water maze" from the *Brazilian Journal of Medical and Biological Research* (Angelucci et al., 2002).

This example will illustrate how various elements of the framework for a particular piece of research can be construed on different levels, and how this affects interdisciplinary communication.

In this experiment, rats were released into the maze in a training session, then after an interval of 48 hours, released again for a test session to assess their memory of the previous training. The rats had been given varying amounts of caffeine at varying times, administered dissolved in saline by intraperitoneal injection. Their performance was tested by measuring latency to escape to the platform (solving the maze), distance travelled from the starting point to the platform, and swimming speed.

This paper concludes:

> The present investigation shows that caffeine improves memory consolidation and suggests that it can also improve memory retrieval in a task specific for the spatial/relational memory system that models the human hippocampal memory system. (Angelucci et al., 2002).

Consider two hypothetical situations in which one of the authors of this paper summarises the results by stating "we have discovered that a rat's memory for a maze is better after consuming caffeine". Compare case a) in which she says this to a colleague in the same research group with case b) in which she says the same sentence to a friend who is a physicist.

In case a), the scientists share a framework at quite a fine-grained level. Working in the same research group, they have shared goals, knowledge of methods and practices, etc. In case b), although both parties are scientists, they share a much coarser-grained framework. Consider the following aspects of the framework in question:

- *What it means to "discover" something*:
  A "discovery" here is something that is consistent with observations in certain types of experiment, to a certain level of accuracy in a certain kind of statistical analysis. In Angelucci et al., '[d]ata from the training and test sessions were analyzed separately by two-way ANOVA taking the number of the trial as a repeated measure. Differences between groups were evaluated by the post hoc Duncan test.' P values resulting from these analyses are given in the paper, with $P \leq 0.05$ being deemed a significant result. In case a), the scientists share a knowledge of the types of experiment, the analyses that were carried out on the results, and the significance level that would count as a discovery. In case b), the scientists probably share the basic idea, but the physicist is likely to not be aware of what counts as a good enough level of accuracy to declare a discovery in this field (since the standards in physics are often different, for example the 5 sigma level of significance required for a "discovery" in particle physics is equivalent to $P \leq 0.0000003$). He is also unlikely to have specific knowledge of the experiments and statistical analyses carried out. Of course, they can discuss the details of the experiment

etc. and the physicist can learn these things. In this case, he comes to share a finer-grained framework with the psychologist in this narrow and specific area. He gains something like the kind of "interactional expertise" talked about by Harry Collins and Gareth Evans (Collins and Evans, 2002) that a sociologist of science can attain in the area of her research. This kind of expertise is opposed to the "contributory expertise" of a scientist working in the relevant area. Interactional expertise is not enough to allow full participation in the research, but it is enough to understand the research produced. Even if the physicist gains this interactional kind of expertise, this still does not constitute sharing a finer-grained framework in the same way as the colleagues, because they still do not share aims, history of researching together, fine-grained expertise in carrying out all of the relevant techniques, or the broader framework of concepts and methods outside the specific piece of research in question.

- *What it is to remember:*
  This is the most important point for my purposes. In the case of a rat solving a maze, "to remember" is something like "to reproduce behaviour". The type of memory or aspect of memory involved in the experiment is also implicit in "remember" here. For Angelucci et al., the type of memory is spatial/relational memory, which they say 'models the human hippocampal memory system.' Psychologists distinguish between memory encoding (which can be further subdivided into acquisition and consolidation), storage/retention, and retrieval. In the experiment, the rats were given caffeine in various doses either 30 minutes before training (to test the effect on acquisition), immediately after training (to test the effect on consolidation), or 30 minutes before the test session (to test the effect on retrieval). Their results provide evidence that caffeine aids in memory consolidation and retention (it isn't clear how it could distinguish these two things, and this issue is not discussed in the paper), has no effect on memory acquisition, and may improve memory retrieval. These specifics would be part of the shared framework in case a), but not in case b) because they would not be known by the physicist; the physicist may be more inclined to generalise the result to everything that we call "memory" in ordinary

language. Again, in case b) they could discuss these things and the physicist could come to know much of the above, but they would not share the broader framework, including innumerable assumptions about memory that have been built up over a long period of education and practice.

- *What counts as a maze:*
  This is an example of a relatively fine-grained piece of knowledge about the kinds of experiments that tend to be carried out in this field; it would be shared within the field, but not by the physicist. In the example case I am using here, the maze is described as follows:

  > The water maze consisted of a round tank 170 cm in diameter and 70 cm deep, filled with water. The water temperature was maintained at 25ºC. A platform (11 x 14 cm) submersed 2 cm under the water surface was placed on the center of one of the four imaginary quadrants of the tank and maintained in the same position during all trials.

  To "solve" the maze was to find the platform. This is a variant of the Morris water maze (Morris, 1984) which is widely used, so would be known about within the field, but probably not outside it. As with the previous two points, the physicist could come to know about it by further discussion in case b), and this would make the sharing of frameworks somewhat more fine-grained, but not as fine-grained as in case a).

- *What is meant by "after consuming caffeine":*
  This is another point about the specific details of the experiment (how the caffeine was administered, how long before attempting the maze, etc.), but it is also linked to broader knowledge shared by those who work with animals in laboratory experiments. For example, there will be assumptions about what constitutes a high dose of caffeine versus a low dose for a rat, and this may have certain assumed correlations to what constitutes a high or low dose in a human. For Angelucci et al., doses varied from 0.3–30 mg/kg, in 0.1 ml/100 g body weight, administered by intraperitoneal injection, either 30 minutes before training, immediately after training, or 30 minutes before the test session.

In case a), there is relatively fine-grained communication between the scientists. In case b), on hearing “we have discovered that a rat’s memory for a maze is better after consuming caffeine”, the physicist will probably hear something like “caffeine improves memory in rats”. This is correct, but what has been communicated is relatively coarse-grained compared to what is communicated in case a), because a coarser-grained framework is shared.[9]

My interest is at the level of subdisciplines, and at this level, communication has succeeded in case a) because the framework is shared, but not in case b). Although this sounds counterintuitive, I suggest that this is because we would normally analyse such a conversation at a coarser level of grain at which communication does succeed because the framework is shared. The scientist from the rat experiment research group and the physicist cannot be said to communicate when considering the framework at the subdiscipline level because they do not share enough to achieve the aims that would be part of the shared framework at that level. In other words, they could not work together like members of the same subdiscipline to meet the goals of the subdiscipline.

This account is not meant to give any kind of ontological priority to subdisciplines because they too can be changed, typically forming around domains for investigation. The point is just that methodologically, it is useful to think about frameworks at these levels, since the concepts, theories, goals and practices are those had by that group of people.

In terms of my project specifically, the *variation* in MEMORY *between* subdisciplines can be explained in terms of frameworks. Neighbouring subdisciplines only partially share frameworks at the level of grain we are interested in. They have different (although overlapping) goals, theories and methods. They also therefore have different (although overlapping) concepts.[10]

[9] A layman may hear something like “coffee will improve my memory” which is more coarse-grained still. I will not say more about this here, because I am focussing specifically on frameworks and concepts in science.
[10] I will say more about the issue of concept individuation that this raises, as well as related issues about the reference of the concepts, in chapter 4.

Using the dynamic framework account

We have already seen some of the flexibility my account can offer, through allowing both focus on different nodes, and focus on different levels. My focus is on MEMORY at the level of subdisciplines, but many other possibilities are opened up by the dynamic framework account.

What sorts of things does thinking in terms of frameworks allow us to do? I will answer this question in more detail in chapter 5, but it should already be clear that frameworks can be better or worse at achieving the aims of explanation, prediction and control (although the standards for what counts as a good explanation or prediction or a successful intervention in the world are not framework-independent). Normative judgements about frameworks are therefore possible and desirable.

There is also therefore a notion of progress to be had for the science – the framework getting better at meeting the aims of prediction, explanation and control/intervention. In line with the openness to epistemic pluralism discussed above, this idea of progress retains something important from the Kuhnian picture: it is not a teleological notion. In other words, there is no assumption of a final best framework that science is working towards. Other attempts to provide a notion of progress while respecting this insight often lapse in to Whiggishness. For example Laudan, accepts that '[a]ll this [notion of progress] sounds rather "whiggish," and so it should, for when we ask whether science has progressed we are typically asking whether the diachronic development of science has furthered cognitive ends that we deem to be worthy or desirable' (Laudan, 1984: 65). My account, by contrast, allows us to judge research programmes by their own lights without lapsing into the kind of relativism on which "anything goes", thus steering between Laudan and the radical interpretation of Kuhn.

The anxiety about fragmentation found in the psychology literature can be described as an anxiety about fragmentation and reformation of domains and their associated frameworks. But on the view outlined here, changing and reformation of domains is a normal process, so psychology need not necessarily be worried about the possible fragmentation of the discipline; perhaps the domains need rearranging and new frameworks would be an advantage for investigation. However, there is still reason to worry about exactly *how* the domains should change if they should do so.

How a situated cognition concept of memory can fit into the picture is one small part of this problem. If memory can extend, could we have a unified science of extended memory, or must different types be dealt with as part of different domains with different frameworks? If so, what happens to interdisciplinary collaboration? I hope that answers to such questions will be one consequence of the work I begin here.

In the next section I will focus on changes over time in the *concepts* of target phenomena. This is closely related to changes in the *questions* asked about a domain, because here I am interested in the question "what is memory?" and the concept of memory involved in asking and answering it.

Conceptual development

According to the dynamic framework account, what question a science chooses to investigate at a particular time depends on the needs of the society at the time, and on current theory and practice in the science. However, choosing which questions to ask does not by itself determine the investigation that science embarks on. Importantly, not only do scientists need to decide on a question, they need to have some idea of what sort of answer to the question would be acceptable. I have already discussed one aspect of this, namely standards for what counts as an observation, or what counts as a good explanation. I now want to address another aspect, namely the *concept* of the phenomenon under investigation. I will therefore now move on to talk about the process I call *conceptual development*. This is the most important part of the scientific process for my purposes, and will further justify my focus on the *concept* of memory.

According to the dynamic framework account, as scientists go through the process of choosing to ask a particular question and seeking an answer to it, the meaning of the question itself changes. This is because the scientists' concept of what they are investigating develops as the investigation proceeds. As well as choosing to ask the question "what is x?"; scientists must develop a concept of x that shapes the meaning of the question, and therefore what will count as an acceptable answer at any point during the enquiry. This means that scientists already have an embryonic idea of what x is when they ask the question. This embryonic idea is defined by the constraints or pressures applied to the concept by other elements of the framework. To push the analogy further, there is a "space" left in the framework for the concept to fill.

To describe this "space", I will borrow some terminology from Paul Griffiths and Karola Stotz, and call it the *epistemic niche* of the concept. Although Griffiths and Stotz do not talk in terms of frameworks, they define the epistemic niche as the needs the group of scientists have in their investigation. Over time, changing needs lead to diversification in the concept. 'As a result of such conceptual evolution, what was originally a shared concept between two or more communities of researchers can become a range of related but distinct concepts.' (Griffiths and Stotz, 2008: 508). This is like the variation we see in MEMORY. Different groups of scientists with different concepts of memory are working with different (although often substantially overlapping) frameworks, creating different epistemic niches for MEMORY to fill.

It is changes in the epistemic niche that guide conceptual development, and it is the process of the development of MEMORY in response to its epistemic niches in various subdisciplines that I will be interested in investigating because I believe it is at a particularly interesting stage that is of crucial importance for the situated cognition question.

The idea that there are constraints on what would count as an answer to a question like "what is memory?" is of course not a completely novel one. For example, Ingo Brigandt and Alan Love talk about *criteria of adequacy* associated with a *problem agenda*, (or a *complex explanandum* or *epistemic goal* in Brigandt's preferred terminology, 2010a: 299) in their research on interdisciplinary work in biology (Love, 2008; Brigandt, 2010a; Brigandt and Love, 2010). A problem agenda is 'a "list" of interrelated questions (both empirical and conceptual) that are united by some connection to natural phenomena' (Love, 2008: 877). The criteria of adequacy 'set standards for what counts as an adequate solution' (Brigandt 2010a: 299).

Shapere expresses a view even closer to mine in the following quotation:

> Thus far I have been speaking of theories as answers to questions. While there is a point to this, it should be remembered that those questions themselves, in the cases considered, involve a general idea of what their answer will be like. In this sense – and it is a sense which seems *prima facie* to fit a great many cases in the history of science – a theory is *gradually developed* by a process of increasingly precise and detailed statement of the initial vague idea; there is then no single point in time at which one can say unambiguously that the theory has been *arrived at*…Theory development would then be more appropriately describable, in some cases at least, as a *process of convergence from generality to (relative) precision* than as a precisely datable event like answering a question. (Shapere, 1977a: 553, note 54).

For Shapere, something like Brigandt and Love's criteria of adequacy are contained *in the question* that the science asks, because asking a question like "what is memory?" already involves certain expectations about what an answer to the question would look like. I locate these expectations in the epistemic niche for the concept, in this case for MEMORY. The epistemic niche is the aspects of the framework that apply pressure to the concept, or constrain it. As research progresses, these constraints may become tighter, and this is the process whereby the range of answers that would be accepted to the question become narrower, which often causes the concept to become more precise.

According to Shapere, it is important to analyse the patterns of reasoning leading to our expectations, whether or not they are met by the answer given to the question (Shapere, 1977a: 522). Unsurprisingly, since my view of asking questions in science is similar to Shapere's, I share a similar conception of the work that we should do in analysing the situation. I will therefore return to Shapere's work in chapter 5 when I outline my methods for investigating the sciences of memory.

When I talk about conceptual development, it is important to make clear that developing a concept of x is more than just ruling certain entities, processes, events etc. into the concept and ruling others out. That picture would imply that the question at issue is purely terminological; a question about to which things in the world we attach the label "x". This would of course change what answer to the question is appropriate, but according to the framework account, there is something more substantive at issue in conceptual development. The world isn't divided up into entities, events etc. that can be ruled in or out of a concept until we so divide it by developing our concepts. If a portion of the world comes to be seen as x, that portion becomes a separate entity or event, where before it may not have been so separated. This is not to say that anything goes and we could make up whatever theories and categories we like, but once a part of the world is delineated by a concept, investigation of it will proceed in a certain way that it would not have done had the concept been drawn differently or not at all. Our concepts are tied to other aspects of the frameworks of which they are a part, and changes in concepts will affect other aspects of the framework, changing the theories and practices of the science in question and perhaps redrawing the boundaries of the domain of investigation. This will affect what is seen as significant in the future, and change the future course of investigation (compare Kitcher's significance graphs, Kitcher 2003).

Conceptual development doesn't only affect which theories or practices or explanations of a phenomenon are accepted or engaged in, but also *what sort of thing counts* as an explanation or observation or as a good theory or an appropriate experiment. As discussed above, these things are also important aspects of the relevant frameworks.

Conceptual development is therefore a very important process. The concepts scientists have affect what they take themselves to be trying to investigate, the kind of thing that would count as an explanation or an appropriate investigative practice, the specific goals they pursue, and the domains of investigation. The aspects of investigation discussed so far all affect one another, and the concept of the phenomenon being investigated is a central part of this.

I will now briefly illustrate the process of conceptual development for the concepts of water and of the gene, then discuss a project (the Cognitive Atlas project) which seems to be attempting to map conceptual development for concepts in cognitive neuroscience.

Water

The answer to the question “what is water?” seems to be “water is H_2O”. However, investigation has not been as simple as asking the question and giving this answer, and in fact it is not even a strictly correct answer in very modern science.

Hasok Chang’s (2012) book “Is water H_2O?” details the history of the scientific investigation of water. Water was initially seen as an element and only gradually came to be seen as a compound of hydrogen and oxygen. At that point many chemists gave it the formula HO, only later settling on H_2O. Since then, water has come to be seen, not as a pure molecular substance, but as a ‘complex and dynamic congeries of different molecular species, in which there is a constant dissociation of individual molecules, re-association of ions, and formation, growth and dissociation of oligomers.’ (Hendry, 2008: 523, quoted in Chang, 2012: 248). As Chang says, ‘[i]f we had a simple heap of H_2O molecules, it would not be recognizable as water’ (2012: xvi).

According to the dynamic framework account, this process is more than just a search for the answer to a fixed question: “what is water?” As research takes place, what would constitute an acceptable answer changes because the epistemic niche for WATER, and therefore the concept responding to that niche, change. When water was considered to be one of four elements, “water is H_2O” would have been meaningless. It would not have been an acceptable answer to the question, because the concept of water did not describe a molecular substance before the development of modern atomic theory, and did not describe a compound until even later.

There is no reason to think the process is complete; the meaning of the question and our best answer to it are likely to continue to change. Chang’s discussion brings out this point through his account of the Chemical Revolution and the demise of phlogiston, via the electrolysis of water and different interpretations of it, and various systems of atomic theory and counting atoms, leading to the adoption of the H_2O formula. In terms of what I have been saying here, each of the different theories and practices he talks about had a different epistemic niche for WATER to fill. At any point in this process of conceptual development, the concept was embedded in a framework of theories, practices and other concepts that helped to shape the kind of thing scientists were looking for when they asked the question “what is water?” Factors external to the science also had their part to play, as famously discussed for the Chemical Revolution in the play Oxygen (Djerassi and Hoffmann, 2001).[11]

In the case of water, conceptual development can be glossed as a move from disagreement towards consensus (although the real story is much less neat and tidy). As a pluralist, Chang argues that this should not be the case; more avenues of research should be kept open rather than trying to force agreement on a single one. This is compatible with my view, although whether multiple avenues are productive enough to be kept open is something which must be judged on a case-by-case basis. I will not judge this issue here for the case of WATER, but it will be an important part of my analysis of MEMORY. For the gene concept there seems to be more divergence than consensus, and this is one of the reasons it is illuminating.

The gene

The idea of the gene is usually taken to have been proposed by Gregor Mendel in the 1860s. Mendel did not use the term “gene”, but proposed a hypothetical unit of heredity to account for the results of his experiments on inheritance. The word “gene” was coined by the Danish botanist Wilhelm Johannsen in 1909 to describe these Mendelian units of heredity (National Human Genome Research Institute, 2013, Online Education Kit). Looking for something to fill this hypothetical role, scientists eventually settled on DNA. This is the standard story of classical genetics. A Nature article on the subject of genes says:

> In classical genetics, a gene was an abstract concept — a unit of inheritance that ferried a characteristic from parent to child. As biochemistry came into its own, those characteristics were associated with enzymes or proteins, one for each gene. And with the advent of molecular biology, genes became real, physical things — sequences of DNA which when converted into strands of so-called messenger RNA could be used as the basis for building their associated protein piece by piece. The great coiled DNA molecules of the chromosomes were seen as long strings on which gene sequences sat like discrete beads. (Pearson, 2006: 399).

[11] Another interesting example of the effect of factors from outside the science is Chang’s analysis of the rejection of J. W. Ritter’s views in terms of his style and strategy of communication (Chang, 2012: 127).

Genes began as *functionally* defined, moving towards a definition in terms of *substance*, first proteins, then DNA. In other words, not only the answer but the *type of answer* changed over time as surrounding theory and practices changed.

Increasingly, the classical picture of the gene is breaking down. Results, for example those showing the involvement of RNA, and extragenomic modes of inheritance (epigenetics), are making it increasingly untenable. The problems this is causing are worth quoting at some length because they are to some extent paralleled in research on memory:

> [D]oes it matter that many scientists not directly concerned with molecular mechanisms continue to think of genetics in simpler terms? [the classical picture]. Some geneticists say yes. They worry that researchers working with an oversimplistic idea of the gene could discard important results that don't fit. A medical researcher, for example, might gloss over the many different transcripts generated by a sequence at one location. And the lack of a clear idea of what a gene is might also hinder collaboration. "I find it sometimes very difficult to tell what someone means when they talk about genes because we don't share the same definition," says developmental geneticist William Gelbert of Harvard University in Cambridge, Massachusetts.
>
> Without a clear definition of a gene, life is also difficult for bioinformaticians who want to use computer programs to spot landmark sequences in DNA that signal where one gene ends and the next begins. But reaching a consensus over the definition is virtually impossible, as Karen Eilbeck can attest. Eilbeck, who works at the University of California in Berkeley, is a coordinator of the Sequence Ontology consortium. This defines labels for landmarks within genetic-sequence databases of organisms, such as the mouse and fly, so that the databases can be more easily compared. The consortium tries, for example, to decide whether a protein-coding sequence should always include the triplet of DNA bases that mark its end.
>
> Eilbeck says that it took 25 scientists the better part of two days to reach a definition of a gene that they could all work with. "We had several meetings that went on for hours and everyone screamed at each other," she says. The group finally settled on a loose definition that could accommodate everyone's demands. (Since you ask: "A locatable region of genomic sequence, corresponding to a unit of inheritance, which is associated with regulatory regions, transcribed regions and/or other functional sequence regions.") (Pearson, 2006: 401).

This quotation shows the problems encountered when scientists with different but overlapping frameworks try to find a concept that will fit all of their differing epistemic niches.

One attempt by scientists to begin to deal with this sort of problem (not for the gene concept itself but for gene products) is the Gene Ontology project (Gene Ontology Consortium, 2000). To quote from their website,

> Biologists currently waste a lot of time and effort in searching for all of the available information about each small area of research. This is hampered further by the wide variations in terminology that may be common usage at any given time, which inhibit effective searching by both computers and people…The Gene Ontology (GO) project is a collaborative effort to address the need for consistent descriptions of gene products in different databases. (Gene Ontology, 1999, documentation).

The large size of this project relative to its narrow scope (restricted to gene products in a cellular context) indicates the magnitude of the problem in the biological sciences. The ontology project itself does not determine which concepts should be adopted however. This is a feature of such big data approaches that I will return to in the next sub-section.

One possible solution to the variation in definitions is splitting the concept:

> Rather than striving to reach a single definition — and coming to blows in the process — most geneticists are instead incorporating less ambiguous words into their vocabulary such as transcripts and exons. When it is used, the word 'gene' is frequently preceded by 'proteincoding' or another descriptor. "We almost have to add an adjective every time we use that noun," says Francis Collins, director of the National Human Genome Research Institute at the National Institutes of Health in Bethesda, Maryland. (Pearson, 2006: 401).

The splitting solution is also suggested by the work of Paul Griffiths and Karola Stotz. It is from Griffiths and Stotz that I take the term "epistemic niche". They identify three different concepts of the gene – "instrumental", "nominal", and "postgenomic" – developed in response to three different epistemic niches (Griffiths and Stotz, 2006, 2008[12]).

The instrumental gene is similar to the original concept of the gene as a factor in a model of heredity. Genes in this sense are described as being akin to centres of mass in terms of their instrumental role in a model (Griffiths and Stotz, 2006: 499). Nominal genes are specific named DNA sequences. 'Many but not all instrumental genes correspond to nominal molecular genes (and lend them their names), and many but not all nominal molecular genes correspond to instrumental genes.' (Griffiths and Stotz, 2006: 500). Griffiths and Stotz use the term 'postgenomic gene' to refer to 'collections of DNA elements that play the role of the gene as envisaged in early molecular biology – acting as templates for the synthesis of gene products – but which are not "nominal" genes, because the way in which DNA is used in the production of the relevant gene products does not fit the traditional stereotype.' (Griffiths and Stotz, 2006: 500).

Griffiths and Stotz (2008) discuss how these different concepts best fit the epistemic niches for different research. For fields like medical genetics and population genetics, the instrumental Mendelian gene concept is the best fit. It allows scientists to talk about the "gene for" a particular disorder, meaning the 'sections of chromosome whose pattern of inheritance explains the phenotypic differences observed in patients' (Griffiths and Stotz, 2008: 516). This may not play a role in the development of the abnormality, i.e. it may not be a molecular gene (Griffiths and Stotz, 2008: 516–517). When the development of the abnormality is researched, the epistemic niche is different and the molecular gene concept may be of more use. In my terms, these different epistemic niches consist of differences in the surrounding frameworks for the subdisciplines carrying out the different kinds of research.

The interesting thing about the divergence here is that Griffiths and Stoz's three gene concepts, or three answers to the question "what is a gene?", are not even similar in kind. They do not, for example, all refer to substances. There has therefore been a

[12] Griffiths and Stotz do not talk in terms of epistemic niches in their 2006 where the results of their work are discussed in detail, but they do reanalyse it in these terms in their more methodological 2008 paper.

divergence in what would count as an answer to the question (substance, functional, etc.), not purely in the answers themselves. It is an open possibility that a similar conceptual splitting could happen for the concept of memory, given that there is no consensus on whether science is looking to answer the question “what is memory?” in terms of a substance, a functional role etc. I am not assuming that there must be a *single* correct answer here, but there are still criteria for what counts as (an) adequate answer(s).

An example from cognitive science

This example is more speculative because these sciences are less mature (see next chapter), but I include it because it concerns a subdiscipline that investigates memory, namely cognitive neuroscience. The Cognitive Atlas project (Poldrack, n.d.) seems to be investigating something like conceptual development for cognitive neuroscience (albeit in a way very different from my approach).

The Cognitive Atlas project is an open collaborative project to map the current ontology of cognitive neuroscience, where an ontology is ‘an “explicit specification of a conceptualization,” [Gruber, 1993] or a structured knowledge base meant to support the sharing of knowledge as well as automated reasoning about that knowledge.’ (Poldrack et al., 2011: 2). In this way, it is similar to the Gene Ontology project referred to in the section on genes above, and Poldrack et al. explicitly compare the two projects.

The Atlas is divided into *concepts* and *tasks*. ‘A mental concept is a latent unobservable construct postulated by a psychological theory… Some potential kinds of mental concepts include (but are not limited to) mental representations and mental processes.’ (Poldrack et al., 2011: 3). ‘A mental task is a prescribed activity meant to engage or manipulate mental function in an effort to gain insight into the underlying mental processes.’ (Poldrack et al., 2011: 3). Various relations between these terms (e.g. is-a, part-of, measured-by) are also included. A page for each concept gives a definition of the concept, its relations to other concepts, tasks used to measure the concept, links to databases containing any associated fMRI images (e.g. NeuroSynth, see Yarkoni et al., 2011), a section for discussion, and a bibliography of relevant papers. Note the centrality of concepts here, showing further consensus that problems in science can be addressed at this level. There are also *collections* of concepts and tasks. One type of collection is *task batteries*; these are collections of tasks. Another type of collection is *theories*. These theories are lists of assertions, the assertions being made up of concepts, tasks and their relations.

When the project began, '[a]n initial vocabulary of more than 800 terms was identified manually through analysis of a broad set of publications on cognitive psychology and cognitive neuroscience and curated by three of the authors (Russell A. Poldrack, Robert M. Bilder, Fred W. Sabb).' (Poldrack et al., 2011: 3). Subsequently, anyone approved by the authors as a contributor can make changes and additions. Discussion of these changes is encouraged first, and there is space incorporated for this.

This project collects scientists' concepts as its data, but wants to get at an ontology that captures the world. The fMRI data is seen as a way to do this. One of the problems the project was set up to overcome is the failure of current neural data to match current psychological categories. That is to say that, for most psychological categories, multiple brain areas seem to be involved; also each brain area is involved in several psychological states or processes. Neuroscience and psychology have different but overlapping frameworks creating different niches for their concepts. As they attempt to collaborate, tensions emerge because of these different frameworks.

There are also other tensions, including disagreement over the mental concepts themselves. For example Poldrack et al. give three different definitions of working memory, all of which are found within cognitive neuroscience (Poldrack et al., 2011: 1). These definitions apply to distinct processes that occupy different roles in investigation, or places in different frameworks.

There is also a problem of equating tasks with mental constructs (Poldrack et al., 2011: 2). Poldrack et al. detail several problems this causes, one of which being the assumption that each task measures a specific construct. They give the example of 'the "Sternberg item recognition task"…often referred to as the "Sternberg working memory task", which implies that it measures a specific mental construct ("working memory")' (Poldrack et al., 2011: 2). Not only is there ambiguity in what "working memory" means, but also

> any link between tasks and constructs reflects a particular theory about how the task is performed; thus, equating tasks with constructs makes theoretical assumptions that may not be shared throughout the community (and further, those community assumptions may be incorrect). (Poldrack et al., 2011: 2).

In the terms I have been using here, the task and mental construct are part of different frameworks for different members of the community. Task and construct are related to one another and to other aspects of theory and practice in different ways in these different frameworks.

These tensions have arisen as the aims of the sciences involved have changed, for example by becoming more collaborative. One important change is that the frameworks of psychology and neuroscience used to attempt to account for different data, but now both want to account for the fMRI data. There might be several possible ways to achieve this, but one is to change the psychological categories.[13] This process would have to take place by mutual readjustment between the concepts and the rest of the framework, because the current concepts are embedded in the other aspects of that framework and cannot be changed without profoundly affecting it.

The Cognitive Atlas project maps the current situation in cognitive neuroscience and will continue to map its changes over time, i.e. to map what I have called conceptual development, as the sciences involved attempt to resolve the tensions I have mentioned. However, the Atlas project, like the Gene Ontology project, does not in itself determine what concepts should be adopted. Whether the discussion pages involved in compiling the Atlas are helpful in doing this remains to be seen. This kind of big data approach is a new method that is being tried out in a fairly speculative manner. Relatively neat examples of conceptual development like the story of water are not available for the cognitive sciences, largely because they are

[13] At least change them within cognitive neuroscience. Whether this should affect our folk psychological categories is an interesting question, but one I set aside here.

less mature than the physical sciences. What this means will be the subject of the next chapter.

Conclusion

In this project, I assume that science is aiming to find one or more full or partial accounts of the world that allow prediction, explanation, and useful practical applications, rather than aiming to find a single true account of the world.

I have outlined a picture of scientific investigation that I call the *dynamic framework account*. The framework is a network of theories and practices that mutually influence and constrain one another. Such aspects of theory and practice include goals, methods, concepts, criteria for what counts as an observation, criteria for what counts as a good explanation, etc. Scientific communities can be construed at different levels, which share frameworks at different levels of grain. For example those in the same research group share a more fine-grained framework than those in different research groups, but the same subdiscipline.

Frameworks change over time. I am particularly interested in the change in concepts (conceptual development) but this is constrained by surrounding aspects of the framework, so cannot be looked at in isolation. I am calling these aspects of the framework that constrain or apply pressure to a particular concept the *epistemic niche* of that concept. When a question such as “what is memory?” is asked, the epistemic niche for the concept of memory affects what kinds of answer will be acceptable.

Different subdisciplines studying memory have different but overlapping frameworks, which may result in different concepts of memory. On a pluralist picture, more than one of these concepts may function equally well as investigative kinds (see chapter 1). My project here is to investigate whether any situated cognition concepts of memory are functioning in this way.

I have illustrated the dynamic framework account and the notion of the epistemic niche by discussing the conceptual development of WATER, GENE, and mental concepts in cognitive neuroscience. This final example begins to bring out the difficulties of working with immature disciplines like many of the sciences of memory, and the next chapter will discuss this further.

3. The Sciences of Memory as Immature Sciences

Introduction

Many of the sciences of memory are often described as “immature sciences”, both by philosophers of science and by the scientists themselves. This applies in particular to each of the cognitive and social sciences (e.g. psychology, sociology), but also to the interdisciplinary venture of cognitive science. While some of the disciplines falling under the umbrella of cognitive science might not typically be considered immature (for example philosophy), the collaborative multi-discipline is relatively new and is often so described. In terms of what I said in the previous chapter, we should expect many of the central concepts in these sciences to be at an early stage of conceptual development. This seems to be borne out by the disunified nature of these sciences. What exactly is the link between disunity and immaturity, and what role does conceptual development play? These are the questions that this chapter attempts to address.

I will offer my own account of what it is for a science to be immature to help clarify the situation in which many of the sciences of memory find themselves, and explain why it is particularly difficult but particularly interesting to study them in the way I intend to at this point in their development.

The received view

Whatever immaturity is, it doesn’t seem that it can just be a matter of time, given that the mind and social behaviour have been studied since the ancient Greeks if not before. Even taking the dates of the foundation of scientific disciplines in the modern sense, and for example dating the founding of psychology to the time of the establishment of Wundt’s lab in 1879, some newer disciplines such as genetics aren’t usually taken to be as immature as psychology (Rand & Ilardi, 2005: 9).

Talk about immaturity often cites Kuhn, in particular the *Structure of Scientific Revolutions* (1970, 1996). According to the Kuhnian account, the cognitive and social sciences are immature because they are pre-paradigmatic. For Kuhn ‘[a]cquisition of a paradigm is a sign of maturity in the development of any given scientific field' (Kuhn, 1996: 11).[14] The cognitive and social sciences are seen as preparadigmatic in the sense that their research is still largely based around the accumulation of data and observations without the guidance of a unifying theory, as Rand and Ilardi say about psychology here:

> From its inception as a distinct discipline, psychology has been characterized by conceptual disarray (Henriques, 2004) and relatively slow scientific progress (Meehl, 1978). This is not to suggest any shortage of psychological research, as the field generates a massive empirical literature each year. Rather, we note that psychology's myriad and diverse programs of research are, as a rule, neither coherently connected to one another nor meaningfully linked to relevant lines of investigation in related scientific disciplines (Staats, 1999). The field has instead witnessed the relentless accumulation of assorted facts, findings, and theories that typically fail to find integration across rival research enclaves and theoretical factions (Ilardi & Feldman, 2001[…]; see also Miller, 1992; Staats, 1983). In short, psychology functions as an immature science (Kuhn, 1970). (Rand and Ilardi, 2005: 7).

Poldrack et al. begin their paper about the Cognitive Atlas project in cognitive neuroscience (discussed in chapter 2) with a quote from Rutherford B. Rogers saying that "[w]e're drowning in information and starving for knowledge", expressing a similar view.

As the quote from Rand and Ilardi reveals, it is "*conceptual* disarray" and the failure to integrate "*programs of research…facts, findings, and theories*" that are at issue. In other words, sciences like psychology have a plurality of disunified *frameworks* in play, in the sense of framework introduced in the last chapter (recall that frameworks consist of concepts, theories, practices, methods, etc.). Because a framework is similar to a Kuhnian paradigm, my terminology captures the received view, but without carrying any implications about the rest of the Kuhnian machinery of the scientific process. This seems to provide the main account of what it means for a science to be immature on the received view. In a slogan: immaturity is disunity.

On this view, in order for a science to mature, it must become more unified. Arthur Staats is one figure who holds this view about psychology. Staats says:

[14] Kuhn's position may not be this simple (see von Eckhardt, 1993) but this is how his view is usually presented in discussions of immaturity.

> [M]y own philosophy of science states as a fundamental principle that all sciences begin in disunity and only advance toward unification by dint of hard and lengthy scientific achievement…Psychology is very much a science, but it is a science early in its career. Psychology is what I call a modern disunified science, with a plethora of diverse and unrelated scientific products but with little investment in unifying those products. The resulting disorganization of knowledge leads people such as Toulmin (1972) to consider psychology a "would-be science." (Staats, 2004: 273).

In Kuhnian terms, until a science is coherent and unified, it cannot be said to have a paradigm or to practice normal science.

Whether or how this unification should happen is of great importance for my project. In the psychology literature in particular, talk about immaturity is coupled with the anxiety about the fragmentation of the discipline discussed in my introduction of the problem in chapter 1. At the root of this problem, I suggested, is the fact that different scientists studying ostensibly the same phenomenon have different concepts of that phenomenon embedded in different frameworks of theory and practice. The conceptual disunity at the root of my problem in answering the situated cognition question is the same disunity that identifies many of the sciences of memory as immature on the received view.

In chapter 1, I suggested that the conceptual disunity in the sciences of memory makes it difficult to answer the situated cognition question. This is because, if there is variation in how memory is conceived, there could well be variation in whether a situated cognition perspective is acceptable or not. Taking this alongside the received view of immaturity, it might appear that we should wait for the relevant sciences to mature and become more unified about their concepts before questions like the situated cognition question can be decided.

However, in chapter 1 I also stated that I take an open-minded approach to a kind of epistemic pluralism, assuming that we have no reason to think that there is only one true description of the world that science is converging on. This suggests that a plurality of disunified frameworks may be no bad thing. In the next subsection, this assumption will be explored further, leading me to abandon altogether the idea that immaturity is disunity.

Pluralism in the cognitive and social sciences

In chapter 1, I said that we have no reason to think that there is only one true description of the world that science is converging on. In other words, it may be that there are multiple equally correct ways of describing the world. I left it open whether science is in practice converging on a single account (even though others are conceptually possible) or whether different parts of current science are making use of different accounts. I also noted that it may turn out to be most fruitful to work out the relationships between the different accounts and have them interact with one another.[15] However, this interaction may not amount to unification of the kind required for maturity on the received view.

What I referred to there as "accounts" can now be put more precisely as "frameworks". On my view then, the plurality of disunified frameworks that signifies immaturity on the received view is not necessarily a bad thing. In this subsection I want to briefly review some literature that suggests it might in fact be a good thing. This will only be suggestive, but the rest of the chapter will illustrate the benefits of following this suggestion and developing a new account of immaturity. This will open up a new potential route to maturity, as well as explaining the prevalence of the received view. It will also have implications for how to go about answering the situated cognition question.

Kellert, Longino and Waters (2006) distinguish usefully between plurality and pluralism, and I will make use of their distinction here. Plurality is a descriptive term, while pluralism is normative; it is a program one would advocate. The current situation in many of the sciences of memory is one of plurality; in this section I am looking at literature that advocates pluralism. In other words, literature that suggests that the current situation is no bad thing and that these sciences *should*, for one reason or another, be plural in nature.

[15] If this kind of integrative approach turns out to be the right way to approach pluralism, Rand and Ilardi's notion of non-reductive consilience with the natural sciences may not be too far wide of the mark in the sense that finding out relationships between theories in psychology and in other sciences could be important. However, their picture still seeks a unified science of psychology (Rand and Ilardi, 2005: 14) where the discipline is a "harmonious whole" (Rand and Ilardi, 2005: 17).

There is no single account of pluralism in the literature, and perhaps this shouldn't surprise us. As Goertzen and Smythe point out, 'any attempt at a unified conceptualization of pluralism would be inimical to the very spirit of pluralism' (Goertzen and Smythe, 2010)! Here I want to focus on the idea of plurality as disunity in the sense discussed in the last subsection, and pluralism as a program that advocates keeping this disunity.

Pluralism is becoming an increasingly popular outlook in the philosophy of cognitive science (e.g. Dale, 2008) with several journal special issues being devoted to the topic (e.g. American Psychologist's September 1991 issue containing a paper by Staats on "Unified Positivism and Unification Psychology", and replies in a later issue (Staats, 1991; McNally, 1992; Kukla, 1992; Schneider, 1992; Green, 1992; Kunkel, 1992); and a New Ideas in Psychology special issue on "Theorizing Pluralism" (Lamiell et al., 2010)). Dale talks about diversity of theories (Dale, 2008: 156), explanatory schemes (p.156), and theoretical frameworks (p.157). Staats talks about 'many unrelated methods, findings, problems, theoretical languages, schismatic issues, and philosophical positions' and 'many unrelated elements of knowledge' (Staats, 1991: 899), '[d]ifferences in method, theory, and phenomena studied' (p. 900) and different theoretical languages built around separately treated phenomena (p. 900). Looking at a selection of authors from the *New Ideas in Psychology* special issue on pluralism, Goertzen mentions conceptual frameworks (Goertzen, 2010: 202), Watanabe talks about paradigms (Watanabe, 2010: 254), and Smythe and McKenzie about diversity of methodologies and theoretical orientations (Smythe and McKenzie, 2010: 227). It seems from these quotations that pluralism here is about a diverse range of frameworks in the sense introduced above.

Pluralism about science in general is also an increasingly popular position (e.g. Chang, 2012).[16] There are some reasons to think that the situation may be different in the so-called "special sciences" from how it is in physics. Metaphysics based on the special sciences is more likely to advocate pluralism (e.g. based on Dupré's promiscuous realism (Dupré, 1993)). Accounts of natural kinds used in the philosophy of these sciences are often based on Boyd's Homeostatic Property

[16] Although it is not a new position, see e.g. Suppes, 1978.

Cluster (HPC) account (e.g. Boyd, 1991). As discussed in the section on natural kinds in chapter 1, this can be interpreted in a pluralist manner and is interpreted this way in the literature (Boyd, 1999a; Wilson, 2005b; Craver, 2009). One feature which might distinguish the special sciences is causal complexity. According to this view, the cognitive and social sciences study so many causal chains with such complex interconnections (what Wimsatt, 1994 refers to as "causal thickets") that there can be no unified "theory of everything" in their domains. Alternatively, it may be that all sciences are amenable to pluralism (Cartwright's "dappled world" where laws apply only locally (Cartwright, 1999)) is a metaphysics based on pluralism that takes many of its examples from physics).

Pluralism is not just an increasingly popular position; it is also a fruitful one. One reason to advocate keeping plurality is that pluralist theorizing about science has beneficial implications. For example, there is an argument in the feminist philosophy of science that theories and models are partial and goal-directed according to the interests of particular groups. According to pluralism, many of these theories and models should be pursued, thus allowing traditionally marginalised voices to be heard (Longino, 1996: 275–277). This suggests that pluralism is not only increasingly popular (as I argued above), but also leads to advantages.

Although some degree of interaction or integration between frameworks may be advantageous according to some pluralist accounts, this falls short of unification. If these accounts are correct and pluralism is the best approach for the cognitive and social sciences to take, either these sciences should remain immature, or immaturity cannot be disunity. The former option is unappealing, so it seems we need a new account of immaturity.

A new account of immaturity

Introducing Shapere's internal/external distinction

My account of what it is for a science to be immature will focus on ideas developed by Dudley Shapere, particularly in his 1986 paper "External and Internal Factors in the Development of Science". The account will build on Shapere's claim quoted in chapter 1 that some domains are *important* but not *ready*. I discussed the *importance* of the sciences of memory in chapter 1, and the present chapter can be seen as my elaboration of what it means to not be *ready*. My discussion will be in terms of the notion of a framework introduced in chapter 2, my interpretation of which was inspired by Shapere's work.

As quoted in the previous chapter, Shapere sees theory development as a gradual process of 'convergence from generality to (relative) precision [rather] than as a precisely datable event like answering a question.' (Shapere, 1977a: 553, note 54). Shapere talks about internal and external considerations which shape various aspects of science. To put this in terms of my dynamic framework account, particular theories, methods, goals, criteria for what counts as an observation etc. (in other words all the aspects of frameworks) can be internalised into the science, or can be external to it. Any particular one of these nodes in the framework (e.g. a theory, or a concept) is shaped by surrounding nodes, some of which are internal and some external to the science in question.

The debate over internal and external factors in theory choice is an old one in the philosophy and sociology of science, linked to the debate over realism and relativism. While much of the literature has moved on from this distinction (see e.g. Shapin, 1992), I think there is still a lot to be gained from Shapere's discussion. If Shapin is right, the distinction was largely abandoned because the internalist and externalist positions were not drawn or argued for coherently, not because the existence of a distinction was shown to be worthless. A position based on Shapere's more subtle version of the distinction may well therefore still be worth considering, and I think the work it allows me to do here demonstrates one reason to do so. I will first roughly and very briefly outline the traditional debate, so we can see what Shapere's version is not.

External considerations (e.g. social and political context) are sometimes seen as the territory of sociologists of science, while internal considerations (e.g. epistemic factors) are the territory of philosophers. Sociologists of scientific knowledge draw a distinction between the strong and weak programs. On the weak program, scientific theories are shaped by external factors only when something goes wrong; good theories are shaped only by internal factors. According to the strong program, external factors shape all science. The strong program appears to lead to relativism about science because the results depend only on the social and political context of the time. Shapere argues that neither the strong nor the weak view is correct. While he does not deny the theory-ladenness of observation or the importance of context, he argues that ‘internal factors are generally sufficient to guide science in its inquiries…and that that sufficiency has, as a matter of contingent fact rather than a necessity of logic, tended to increase over the history of science’ (Shapere, 1986a: 2). In order to draw this conclusion, he reinterprets the boundary between internal and external considerations in a way that goes beyond the outdated view of the debate just outlined, and will be useful for my methods here.

On Shapere’s view, there is no distinction between the internal and external that can be laid down from a meta-scientific perspective, for example by philosophy or sociology. The distinction must emerge over time from the practice of the science itself; ‘it is a distinction which has been forged in the very process of investigation of nature, not laid down in some edict from heaven or philosophy which determines what counts as scientific and what does not.’ (Shapere, 1986a: 6). Which considerations are internal and which are external is therefore subject to gradual change over time as the science develops. In Shapere’s words,

> [c]larification with regard to these four aspects of inquiry—what to study, what was relevant to the study, the appropriate methods for that study, and the character of an explanatory conclusion to the study—required *learning how to learn* about nature. (Shapere, 1986a: 3, emphasis in original).

Here we see a move beyond the outdated focus on *theory* choice to talk about many different aspects of science (different parts of frameworks in my terms).

To say that considerations are “internalised” means that they become part of the background that the science can rely on. This builds on Shapere’s (1984) work on objectivity and rationality of reasons in science. ‘[…T]he problem is to show how the employment of some “background beliefs” (rather than others) can be described convincingly as the use of *background information*, serving as *reasons*’ (Shapere, 1984a: 640). Beliefs which have proved themselves become internalised – become part of the background information – and this background information is the basis of what science counts as reasons in its deliberations (Shapere, 1984a: 645).[17] External considerations don't count as reasons (Shapere, 1984a: 648). This use of the word "reasons" refers not just to the causes of particular beliefs being accepted into the body of background information, but to *rationality*. A science is rational to the extent that it has internalised the considerations on which it relies (Shapere, 1984a: 654). What does it mean for beliefs to "prove themselves"? Something will come to be taken for granted as background if reliance on it results in success. As Shapere puts it in his 1986 paper, '[t]hose considerations become internal, scientific, which have been found, as a matter of contingent fact, to be doubt-free (successful and coherent) and relevant to the domain under investigation. All other considerations become external, non-scientific' (Shapere, 1986a: 6–7). It is through this process that the internal/external distinction emerges from the investigation of nature, i.e. from the practice of science. This account gives a notion of progress for a science, in terms of the increasing sufficiency of internal factors to guide science as the distinction emerges (Shapere, 1986a: 2).

A few points about Shapere's account and my interpretation of it are important by way of clarification.

Firstly, what is meant by "success" here is something like providing accurate predictions, good explanations, and useful practical applications. It is impossible to be more precise than this while talking in abstract terms, because the criteria something must meet in order to be accurate, good, or useful vary between different frameworks, and change over time. In line with what I have said about realism and epistemic pluralism, although what it means to be successful is relative to a framework, we do not have unrestricted proliferation of theories and practices, or proliferation restricted only by social and political whims.[18]

---

[17] This notion of reasons is a little restrictive. Shapere does say that ideas, methods etc. which have not yet been fully accepted as background knowledge but were constructed based on such knowledge and function in the same ways can be called "reasons" in a derivative sense (Shapere, 1984a: 649).

[18] The meaning of "success" being relative to the details of the particular case is a familiar notion from ordinary language. As Chang says, '…I think it is futile to attempt to define "success" in any one-dimensional way – we don't try to do that with life in general, and it's not clear to me that we should try it in science. The "success of science" can only really mean the achievement of whatever we value in science – Kuhn (1977, 322), van Fraassen (1980, 87), Lycan (1998, 341) and others give a long and diverse list of epistemic desiderara' (Chang, 2012: 230). Since what we value varies, any such notion of success is bound to be relative to the framework (similar to Chang's notion of a

Secondly, it is important to emphasise that relying more on internal considerations and less on external ones does not mean that science should become increasingly isolated from society. Part of the motivation for tackling a particular research question rather than another will still often come from outside the science; there will still be wars, epidemics, technological fashions and so on, and they will still bring with them a demand for certain types of research to be prioritised over others. Part of what it is to be successful is to provide the kinds of predictions, explanations and practical applications that are useful to the society at the time. There is nothing irrational about this context-embeddedness of science. It does not amount to dependence on external considerations in Shapere's sense unless it provides criteria which a science relies on as reasons that determine its theories and practices in a way that does not lead to success (religious considerations today are one example of this, see below).

Thirdly, one might wonder what is distinctive about Shapere's account. Most philosophers of science would recommend retaining successful theories and practices. What more does Shapere give us? Most philosophers of science would recommend retaining successful theories *because they are likely to be true*. Shapere's account gives us more than this straightforward demand for success in terms of accurately describing the world. It respects the fact that what would count as doing this is variable. It tells us that by looking at the history of a scientific framework, we can see what has been internalised, and thus get concrete recommendations for how the science should or should not proceed. This is of more use to both scientists and philosophers than a general instruction to keep successful theories because they are likely to be true. The project I am engaged in in the following chapters will demonstrate one kind of concrete recommendation that can be made.

A fourth important point is the contingency of what is internalised; we cannot predict or decide in advance of doing the scientific research which features will be

---

"system of practice"): 'Any real-life success is a limited, relative, and provisional thing. Even if the truth of a statement within a system of practice is quite precise and assured, our affirmation of that truth should be only as definitive as our acceptance of the system itself, which is in turn only warranted if the system continues to be successful.' (Chang, 2012: 214).

internalised as it progresses. This is part of a more general thesis of Shapere's which he calls the *Principle of Rejections of Anticipations of Nature*, according to which '[t]he results of scientific investigation could not have been anticipated by common sense, by the suggestions of everyday experience, or by pure reason' (Shapere, 1987: 1). This is also an important part of his rejection of a meta-scientific perspective from which science can be judged.

Finally, another important note is that it is not the case that once internalised, methods, concepts, theories etc. are no longer open to question (Shapere, 1986b: 22–23). This is a familiar picture of what it means to accept something in science; it becomes part of the background that is relied upon for further research, but it is still possible to overturn even the most entrenched parts of the science under the right circumstances.

I will now give some more concrete examples to flesh out this theoretical skeleton. Important examples discussed by Shapere include religious considerations, the compositional approach, and unification. I will briefly outline these cases here, largely by way of illustration of the view, rather than as arguments for it. To present the cases as arguments would involve too much historical detail, taking me too far from my aims in this chapter. The usefulness of the view for my project as a whole is the only argument I will give for it.

Turning first to religious considerations, we can see that these used to be an important part of the background according to which scientific theories, results etc. were judged. Newtonian theory needed God to intervene every so often in the motions of the planets, and this was not seen as dependence on something external to science at the time, while it would be today (Shapere, 1986a: 4–5). Shapere argues that religious considerations came to be seen as external because purely internal considerations became sufficient. In his words, 'such considerations were external to science precisely because the laws of science had been shown (even if as yet imperfectly and incompletely) to be sufficient to account for certain phenomena which had previously seemed to require divine intervention' (Shapere, 1986a: 5). Religious considerations did not establish the track record of success required for internalisation into science, so it would now be considered irrational to treat them as reasons.

Turning to my second example, the compositional approach in chemistry[19] – thinking about material substances in terms of their constituents and the forces between them – is at the foundation of the modern discipline. However, we did not always look for compositionalist explanations; the alchemists sought “perfectionist” explanations, based on the idea that material substances were elements in varying degrees of perfection. The compositional standard for what constitutes a good explanation in chemistry came to be internalised because it resulted in success in various ways. For example, it allowed order to be imposed on the large numbers of new substances discovered in the later middle ages and early modern period, resulting ultimately in the periodic table (see Shapere, 1984b for a more detailed analysis). To provide an explanation in compositionalist terms is now considered rational and this standard is internal to the science, whereas to credit something as a good explanation according to the perfectionist standards used by the alchemists would be to use an external standard, and is therefore irrational.[20]

The other example I will give here, unification, is of particular relevance for my project. Unification between different domains was not always seen as important. However, the unification of electricity and magnetism and other similar cases were successful, so compatibility with theories in other domains was internalised as a

[19] I made reference to the rise of the compositional approach in chapter 2, when discussing the dynamic nature of the framework, and how its different aspects or nodes are interrelated.

[20] Shapere focusses more on the chains of reasons for adopting new aspects of the compositional approach (e.g. positive weight as central to explanations, the idea that fire can break substances into their constituents) according to whether they are successful; he talks less about the reasons for dismissing their predecessors/competitors. There is perhaps an assumption of monism inherent in this view – a new approach takes over from an old one. Because of the openness to pluralism of my approach, careful attention should also be paid to reasons for abandoning a current approach and its becoming external; it might be that two approaches should be maintained at the same time, even if they are incompatible. In other words, some space is opened up between reasons for the compositional approach being internalised, and the perfectionist approach becoming external. For the view that both oxygenist and phlogistonist chemistry should have been maintained, based on their success, see Chang (2012). According to this view, ‘[s]uccess is a dynamic criterion, and judging relative to success is a game of ruling-in, not ruling-out; provisional success is a matter of being “good enough to stay in”.’ (Chang, 2012: 214). The issue is not so much with Shapere’s methods (internalisation based on success) but with the way they are applied in particular cases (with an implicit assumption of monism).

criterion for a good explanation. Such compatibility is now a reason for adopting a particular scientific theory. As Shapere puts it,

> [i]n addition to doubts based on its failures to account for its domain of responsibility, a theory can also be doubted on the ground that it fails to conform to a type of theory with which we believe it ought to conform—for example, because that type of theory has been successful in several other domains. (Shapere, 1986a: 6).

I will now go on use this distinction between internal and external factors to construct a new account of what it means to be an immature science.

Making use of the distinction

Although Shapere does not use the phrase "immature science," he does talk about the internal/external distinction with respect to sciences at an early point in their development. He says that at first, a science has internalised very little, or more precisely it does not have a firm internal/external distinction. This amounts to the same thing, because the distinction only emerges via the internalisation of considerations over time. An immature science therefore does not have much that it can rely on as reasons, because there has been no time for anything to establish a track record of success, so the science is forced to try out the unproven, or what Shapere in the following quotation calls "hypotheses":

> In all these respects, what are naturally called "hypotheses" played a role; and there was, in earlier phases of science, little to go on in selecting these hypotheses. Or more exactly, the motivating considerations in selecting explanatory approaches might come from just about anywhere. Antagonism to Aristotelian forms, natures, and final causes, rather than the dictates of nature, entered into adoption of the mechanistic and atomistic approaches of the middle and late seventeenth century; Newton developed his theories of motion (and thus of space and time) at least partly in the light of theological considerations, objecting to Cartesian physics on such grounds just as his own views were deemed atheistic by Leibniz and his followers. And in general, the large gap between scientific ambition and scientific conclusion had to be filled, under such circumstances, by considerations which we today would consider non-scientific, external, though at the time there was little or no ground to so distinguish them. Indeed, even the ambitions of science at such stages were dictated, at least partly and perhaps largely, by considerations which would today be called external. For the distinction between the external and the internal to science was at best only rudimentary and in many cases did not exist at all. (Shapere, 1986a: 4).

Things are different for a more mature science like modern physics, and this gives a useful sense in which physics is more mature than psychology: It has established a firmer internal/external distinction, and has a greater body of internal factors that it can rely on as background knowledge. Psychology and the other cognitive and social sciences on the other hand still have relatively little to rationally go on. In my terms, much of their frameworks can be expected to be hypotheses, including not just their theories, but goals, criteria for what counts as an observation, standards for what constitutes a good explanation etc.

For example, with reference to the Cognitive Atlas project discussed in chapter 2, it is not yet clear whether a good explanation in cognitive neuroscience should respect our folk mental categories, their refinements in use in cognitive psychology, the categories suggested by the fMRI data, or whether it must find some way to integrate these different categorizations. In general, where there are different frameworks in the cognitive and social sciences, it is not yet clear whether we should be aiming for reduction of higher level frameworks to lower level frameworks, establishing some non-reductionist relationship (for example Rand and Ilardi, 2005 talk about "consilience"), or whether the different frameworks concern sufficiently different subject matter that the relationship between them does not need to be (and perhaps cannot be) established. The lack of unifying theories in the cognitive and social sciences is thus partly explained by the fact that it is not clear what such theories should look like because there is a lack of internal criteria to measure them by.

Does the account developed here therefore suggest that the cognitive and social sciences are irrational? It may be a consequence of Shapere's view that psychology is *less* rational than physics, although this is not a derogatory description of psychology on this view. Inevitably a less mature science will have a less well-developed internal/external distinction, so there is less of a body of internal considerations to rationally use as reasons. This is still somewhat unsatisfactory however, and in chapter 5, I will draw the distinction for my methods between relying on hypotheses when nothing else is available, and relying on external factors, such that only the latter is irrational.

On the account presented here, maturity does not necessarily correlate with age, because some sciences may continue to work with little or no internal/external distinction for a long time, while others may internalise criteria and develop this distinction more rapidly. Here, I hope to point out one criterion that the cognitive and social sciences are currently relying on as though it has been internalised, when in fact it has not gone through this process, and is being imposed from outside the science. By moving beyond doing this, a new route to maturing further is opened up. This illustrates the benefit of the new view, and also explains why the received view has been so prevalent. The criterion in question is unification.

Specifically, I claim that unification has not been internalised in the cognitive and social sciences, although it has been in physics and chemistry.[21] Defence of the success of unification leans heavily on examples from physics, chiefly the unification of the theories of electricity and of magnetism (Shapere, 1986a: 5). But there is no reason to assume that what works in one area of science should work in another, so there is no reason to assume that this strategy's success in physics indicates that it will be successful in the cognitive and social sciences. In fact to assume that success in one part of science generalises in this way is to presuppose that unification is a consideration that will lead to success. Making such presuppositions is exactly what we shouldn't do according to Shapere, given that which features will be internalised cannot be predicted in advance.

I am not denying that cognitive and social scientists use unification with other domains as a criterion for shaping their theories, but when they do so, it is an external criterion. It is imposed from outside—from the physical sciences. It has not undergone the process of internalisation into the relevant subdisciplines by proving a track record of success in their domains. In fact it has been in use for some considerable time in the cognitive and social sciences without proving such a track record, indicating that it might be time it was abandoned.

[21] It might be significant that some very modern physical science shows increasing specialisation and disunity. This could suggest that the run of success had by unification in physics and chemistry is coming to an end. If this is right, and the change is rational, it is compatible with Shapere's view that even the most entrenched parts of a science can be overturned. If those who advocate pluralism across the board are correct, it may be that unification should become an external consideration at a coarser level of grain. I take no view on this here, as doing so would involve detailed examination of cases in the physical sciences.

The use of unification as a criterion is the reason the received view of immaturity has been so prevalent; it has provided a standard that the cognitive and social sciences have been expected to live up to. Following the suggestion of the pluralist literature in the philosophy of science, it seems that these sciences might be unable to live up to this standard. Trying to do so thus constitutes a barrier preventing them from maturing further. Dismissing this standard as an external criterion removes the barrier by giving them the freedom to internalise pluralism instead, if doing so results in success. In Kuhnian terms, the disunity that seemed characteristic of the pre-paradigmatic phase may become the dominant paradigm in the mature science.

What about instances where use of unification has proved successful? One might think that, for example, unifying psychology with neuroscience is proving to be so. It has allowed the development of various types of brain scanning techniques, and mental states can be “read off” these brain scans with some degree of accuracy, resulting in the possibility of various practical applications (e.g. see Shirer et al., 2012). This seems to be fruitful science in that it allows some degree of prediction and explanation, and has practical uses.

I acknowledge that there has been some degree of success using unification, but continued anxiety about disunity shows that the scientists themselves believe that it is not enough. The lack of neat unification with neuroscience is something that many psychologists find troubling (e.g. see Uttal, 2001). It is not only lack of unification with other subdisciplines that is an issue; within psychology, there is a high level of anxiety about the fragmentation of the discipline, and this anxiety is often coupled to immaturity talk (on the received view of immaturity). This anxiety about fragmentation was discussed in chapter 1.

In fact, the very observations of disunity that led to the cognitive and social sciences being characterised as immature on the received view can now be seen to count against the internalisation of unification. Any degree of success that has been obtained through unification does not seem to be enough by the scientists’ own standards, so it certainly doesn’t yet seem enough to constitute the proven track record required for internalisation.

Of course, philosophy is also external to most of the sciences of memory, so another worry might be that I am trying to impose a pluralist agenda on these sciences from outside. However, this is to misunderstand my conclusion. I am not arguing that a science like psychology should be pluralist, merely that it is too early to decide that it should not be, in line with the *Principle of Rejections of Anticipations of Nature*. All I assumed at the outset was that we have no reason to believe that there is only one best description of nature that science is converging on. My discussion of pluralism in this chapter goes further only in claiming that the current situation is one of plurality and that there are suggestive reasons to think it may stay that way. Following those suggestions cautions against imposing unification as a criterion at this stage; it does not insist on pluralism.

It may be that unification will in the future prove itself in some or all of the cognitive and social sciences, and come to be internalised, or it may not. Perhaps the successes of the method so far are the beginnings of this process, but perhaps they are just isolated instances. Treating unity as a reason in these sciences at present is to prejudge the issue.[22]

Even if unification is internalised, it may be that unity is to be reached in a different way than it has been in physics. In the physical sciences, mathematics has played a big part in unification. For example Maxwell's equations provide a mathematical unification of electricity and magnetism. However, this may not be the appropriate route in the cognitive sciences. Mathematization looks like another criterion which has proved itself in physics but not everywhere else. There are those that advocate mathematization of psychology, but given that there have also been successful qualitative methods in the history of the discipline, to insist on mathematization as a

[22] Would Shapere agree with my view of unification? He acknowledges that the internalisation of unification could have developed otherwise. He says that '[t]hings need not have turned out that way. Connections between things in nature might have been so tight that a piecemeal approach to inquiry would have failed; theories of different domains might not have been coherent with one another; and so on. The achievement of internalization is a contingent matter, not one of logical necessity or of the nature of science.' (Shapere, 1986a: 7). However, he does seem to believe that unity has proved itself in general so, although we could not have predicted in advance that this would happen, science as a whole will approach unity over time. This is something I am denying here. On my view, it may well be the case that the connections between things in nature are different in more causally complex (or higher level) areas of science. We at least have no reason to suppose otherwise.

standard is to impose a criterion external to psychology; one which comes from the physical sciences, much like unification.

A different possibility for unification in the cognitive and social sciences is so-called "big data" approaches. Rather than mathematization via laws in the form of equations, these approaches aim to collect data on the plurality of concepts, methods etc. into databases which can then be used in scientific research. The Cognitive Atlas project discussed in chapter 2 is one example of this.

An alternative to seeing the big data approach used in The Cognitive Atlas project as a different kind of unification from the sort found in the physical sciences is to see the project as a way to capture pluralism. Currently, disagreements over concepts are built in to the project with space for discussion of disagreements over definitions. If there is sufficient disagreement a concept may fork, i.e. be split into several different definitions with a disambiguation page (Poldrack et al., 2011: 7–8). If this disagreement remains built in to the project, the science has internalised a means of coping with pluralism; if it is just a step towards finding consensus, it has internalised a new method for reaching unity. At this stage this project has not proved successful or unsuccessful, so it is too early to say whether big data approaches will be internalised; this is just one possibility. It is not yet clear where the internal/external distinction should fall.

Other possible routes to unity mentioned in the literature may be something like Darden and Maull's inter-field theories (Darden and Maull, 1977), or Bechtel and Hamilton's mechanistic reduction (Bechtel and Hamilton, 2007), rather than the theory reduction often found in examples from physics. Pluralism is not the only possibility that would be opened up by avoiding importing unity as it has been used in physics.

Even if unification eventually came to be internalised by all the sciences, there would be no reason to privilege it as a criterion of maturity, particularly not in advance of this internalisation happening. The cognitive and social sciences must find their own route to maturity, and this may be different to that taken by physics. Avoiding imposing unification as an external criterion could accelerate that process.

The final part of this chapter will concern the sciences of memory and the situated cognition question in the light of this new account of immaturity, and its recommendation against imposing unification.

Immature sciences and the situated cognition question

On my view, most of the sciences of memory are immature in the sense that they have not yet developed a firm internal/external distinction. Many aspects of their frameworks – criteria for what counts as a good theory or method, what counts as an observation, etc. – can be expected to be hypotheses, rather than considerations internal to the science. How does this affect answering the situated cognition question for memory? If we consider MEMORY within the framework for an immature subdiscipline, we would expect to find that its epistemic niche contains few internal considerations. The subdiscipline is therefore at an early stage in conceptual development, i.e. it does not yet have sufficient internal considerations to constrain its concept of memory sufficiently to have decided on an extended, distributed, brainbound etc. concept. Worse still, we cannot anticipate in advance that it will ever do so. It is therefore not yet clear whether situated cognition concepts of memory are adequate in these sciences, or whether there is even a determinate answer to this question, because of their immaturity.

Should we therefore just wait and see what science does? This does seem to be suggested by some in the extended cognition literature (e.g. Hurley, 2010). I think we can do better than that, even given what has been said in this chapter. It is important to try to do so because of the high pressure context in which the immature sciences of memory find themselves, as discussed in chapter 1.

Although it is true that we cannot yet answer big questions like whether the cognitive and social sciences will internalise unity or whether they should remain pluralist, I think we can begin to answer more local questions, including the situated cognition question within particular subdisciplines. I will explain fully how I plan to do this in chapter 5, but it is clear from the work of this chapter that it must involve looking at the criteria that *have* been internalised, and whether they constrain MEMORY sufficiently to determine whether it should be a situated cognition concept, and if so, which one. It is also clear that unification is not one of these internal criteria, at least for the cognitive and social sciences, so the answer may be different in different branches of memory research.

Conclusion

According to the received view of immaturity, immaturity is disunity. However, not only do many of the sciences of memory currently exhibit such disunity, there are reasons to think that they ought to stay that way. If that is right, either these sciences should remain immature, or immaturity is not disunity. The former option is unappealing, so immaturity cannot be disunity.

I have therefore given a new account of what it is to be an immature science, based on Dudley Shapere's distinction between internal and external considerations. According to this account, an immature science is one which does not yet have a clearly delineated internal/external distinction. In other words, not many aspects of its framework have been internalised so that they can rationally be relied on as reasons to shape other aspects of the framework. The cognitive and social sciences are less mature than physics in this sense. In Shapere's words, they have not yet "learned how to learn" (Shapere, 1986a: 7).

Applying the new account, we have seen that unification is one criterion which has not proved itself successful enough to be internalised into the cognitive and social sciences as a criterion for a good theory or explanation. Where unification with a neighbouring domain is applied as a criterion in these disciplines, it is as an external consideration in Shapere's sense, based on its success elsewhere (in this case in physics). The inability of the cognitive and social sciences to live up to this criterion explains the prevalence of the received view of immaturity. Ceasing to rely on this external consideration by being open to the possibility of pluralism may accelerate the process of maturation for the cognitive and social sciences.

Many of the sciences of memory are immature, so their concepts of memory are weakly constrained by internal factors. In other words, they are at an early stage in the process of conceptual development, so the epistemic niches for MEMORY found in the various subdisciplines can be expected to contain few internal considerations.

However it is important to try to begin to answer the situated cognition question because of the importance of memory research. Doing this must involve looking at the criteria that *have* been internalised, and whether they constrain MEMORY sufficiently to determine whether it should be a situated cognition concept, and if so, of which type. Unification is not one of these internal criteria, so the answer may be different in different branches of memory research.

I will outline the methods that emerge from this idea fully in chapter 5, but first I will cover one more piece of preliminary background. This concerns the issue of what concepts are and how they are individuated.

4. A Theory of Scientific Concepts

Introduction

In this chapter, I will talk more about how I construe scientific concepts.[23] The account is not a metaphysical one about what concepts *are*, and it is not meant to be an account of anything consciously going on in the minds of scientists when they use concepts. Instead, my intention is to draw out the way of talking about concepts that is implicit in the dynamic framework account introduced in chapter 2, and justify it as useful for philosophers of science. The main justification is that the account allows the kind of investigation I intend to pursue in subsequent chapters to be carried out effectively. The non-metaphysical status of my account means that it does not necessarily preclude other accounts of concepts drawn from other perspectives or for other purposes.

I will make use of the work of Ingo Brigandt (2004a, 2004b, 2004c, 2010b, 2011, 2012) on scientific concepts. His account has a similar non-metaphysical status to mine; like me, he defends his account by pointing to its fruitfulness in the philosophical work it allows him to do (Brigandt, 2011; 2012: 98). I will build on Brigandt's account – a variety of conceptual role semantics applied to philosophy of science – developing my own version which ties in with the dynamic framework account. My version offers certain advantages, some specific to the kind of project I am engaged in, and others based on features that Brigandt and I share, and I will be drawing out these advantages in support of my version.

This chapter will further explain the role that concepts play in the framework for a subdiscipline, in particular the way that a concept is shaped by its epistemic niche in the surrounding framework. Once this has been explained, it will be clearer how scientific concepts should be studied, and this will pave the way for a detailed discussion of my methods in the next chapter.

I will also compare the account I develop to the Theory Theory of concepts in psychology. Both my account and the Theory Theory appear to suffer from a similar

[23] It might be that a similar kind of account generalises to everyday concepts, but I will not discuss this here.

problem with concept individuation, and the main critical work of this chapter will be in dissolving this problem, making use of a solution offered by Muhammad Ali Khalidi to the problem as it confronts Theory Theory.

In summary, this chapter will present an account of concepts in science that is derived from the dynamic framework account, and that is useful for research in philosophy of science like mine, and show why it does not suffer from the main problem that accounts of its kind are usually taken to suffer from.

Conceptual role semantics and the dynamic framework account

Conceptual role semantics is not a single theory, but a family of approaches according to which a concept's content is defined by its functional role (the approach is also sometimes called functional role semantics). According to one prominent type of conceptual role semantics, a concept is defined by its *inferential* role, that is to say the inferences in which it appears. This *inferential role semantics* or *inferentialism* (e.g. Brandom, 1994, 2000, 2007) is the kind of approach on which Brigandt bases his theory.

There are two major differences between Brigandt and the inferentialists. The first is that Brigandt talks about a concept's *explanatory* role as well as its inferential role because he takes explanation to be central to science, and '[i]t is not obvious how explanation relates to standard models of inference making, so the inferential role of concepts need not encompass their explanatory role' (Brigandt, 2004a: 3). The second difference is that Brigandt takes inferential role to be just one of three components of conceptual content, the other components being reference, and epistemic goal (2010b, 2011: 6) (The nature of these components will be explained below).

I differ from Brigandt in both of these respects. With respect to the scope of the inferential role, I want to expand it beyond both inferences and explanations to include all aspects of the concept's role in the framework. As I have said, the aspects of the framework that apply pressure to the concept constitute its epistemic niche (what the concept is needed to do for the science) and thus shape it. As a natural consequence of this, all of these aspects of the framework constitute the conceptual role. Therefore, according to my version of conceptual role semantics, a concept is defined by its place in the dynamic framework, or its role within that framework, at the relevant level of grain (I will return to levels of grain below). I will therefore talk about *conceptual role,* rather than *inferential role* to reflect this breadth.

With respect to reference and epistemic goal, I do not need these two extra separate components of conceptual content because of the way I have broadened conceptual role. The purposes for which Brigandt recruits reference and epistemic goal are already performed by the dynamic framework, in ways I will elaborate on below. My two differences from Brigandt are thus tightly connected: it is because I construe conceptual/inferential role broadly that I do not separate reference and epistemic goal as distinct components. Or to put this another way, I think it is because Brigandt construes inferential role narrowly that he is forced to invoke the distinct components. I will argue that reference and epistemic goal are already part of the conceptual role properly (i.e. most helpfully) construed, and that there are some disadvantages to separating them.

First I want to clarify Brigandt's view a little further. Then, in two subsections, I will address each of the two differences between my account and Brigandt's introduced above. In each subsection, I will be looking at why Brigandt's theory takes the form it does, or what those features enable him to do. These useful aspects of Brigandt's account will set up desiderata that mine should meet. I will show how some of these can be met in the subsections themselves, but each of the two subsections will also give rise to a desideratum that I will measure my account against later on (in one case in the final section of this chapter, and in the other case in the next chapter). The final subsection will set out some advantages of my account over Brigandt's.

Brigandt can be seen as working in the tradition of theories following Frege (1948), which separate sense and reference. For Frege, reference is denotation, such that the term "a" (or concept A for our purposes) denotes or refers to the particular a. Sense is the "mode of presentation" of the referent. For example "the morning star" and "the evening star" have the same referent (the planet Venus) but different senses.

Brigandt retains the Fregean notion of reference, and cashes out sense as inferential role (Brigandt, 2011: 12). Brigandt defines the inferential role as 'the set of inferences and explanations in which the term figures and which it supports in virtue of its specific content. The inferential role broadly aligns with the definition of a scientific term.' (Brigandt, 2011: 6).

To reference and inferential role, Brigandt also adds a third component: epistemic goal. The epistemic goal is 'the kinds of inferences and explanations that the concept is intended to support' (Brigandt, 2010b: 8). Contrasting epistemic goal with inferential role, Brigandt says:

> In a nutshell, the epistemic goal pursued by a scientific concept's use is the type of knowledge (certain kinds of inferences, explanations, discoveries) the concept is *intended* to deliver, given its usage by a research community. (The inferential role, in contrast, is the set of inferences and explanations that the concept currently actually supports.) (Brigandt, 2011: 7).[24]

The rationality of changes in reference and inferential role are judged according to how well they meet the epistemic goal(s) (which can also rationally change over time, provided it does so in a gradual fashion, see Brigandt, 2010b, section 2, esp. pp.17–19).

The epistemic goal comes along with criteria for what would count as meeting it. These are not radically different from my notion of criteria for what counts as a good explanation, what counts as an observation, etc. Brigandt says '[a]ssociated with epistemic goals are standards of adequacy that specify what would count as meeting the epistemic goal—what method is suitable for an investigative goal, what evidential standards obtain for an inferential or inductive aim, or what criteria of explanatory adequacy underlie an explanatory goal.' (Brigandt, 2012: 99).

According to my account, these criteria depend on other aspects of the dynamic framework, such as background theories, aims, etc., and Brigandt also seems to accept something like this:

[24] He adds in a footnote (footnote 2) that not all concepts need have epistemic goals.

> Any epistemic goal in science comes—given empirical background knowledge—with implicit criteria specifying what it involves to adequately (or more adequately) satisfy the goal. For instance, if the goal is to explain a certain phenomenon, then the particular scientific context specifies what makes one explanatory attempt better then another one. (Brigandt, 2010b: 9).

Where we differ is that on my view, all of these things can be part of the conceptual role because they are all parts of the dynamic framework and thus can all figure in the epistemic niche for a concept. On my account, the conceptual role can include practice, goals (choices of what to explain, what practical applications to work towards), criteria for what counts as a good explanation or theory, etc. On the other hand, Brigandt's definition of inferential role privileges the kinds of things that can be part of inferences and explanations, which would be semantic or linguistic, and therefore seem *prima facie* not to include practice. [25] He relegates goals, aims, and criteria for meeting them to a separate component (epistemic goal).

I will now consider Brigandt's reasons for developing his account in the way that he does, giving rise to desiderata that my account should meet.

Why does Brigandt construe conceptual role narrowly?

Why construe conceptual/inferential role narrowly, so that it only includes the kinds of things that can be part of inferences and explanations? One reason may be the Fregean tradition in which Brigandt is working. As I said above, for Brigandt, inferential role is effectively sense; sense is a semantic entity, not something that would include practice. However, it is a little ungenerous to Brigandt to suggest that he would be such a slavish follower of tradition against the best interests of his own philosophical approach to science.

A more plausible candidate for the reason is that he wants to limit himself to talking about the public language of scientists: 'I will define the notion of conceptual role based on public language, which fits with the fact that in the study of the historical

[25] It is not completely clear that inferences cannot include practice. It seems that they do in Brandom's inferentialism for example (2000: 28–29; 2007: 657–658). Brigandt refers to practice throughout his work (see especially his 2012) but it is not clear that his focus on public language (see below) makes sufficient room for it.

episodes in science we have to rely on the verbal and written reports of scientists.' (Brigandt, 2004c: 20). This approach lends itself to a view of conceptual role that focuses on linguistic inferences and explanations, rather than practice, aims, goals, etc.

I will also be looking closely at the written reports of scientists. However, I want to emphasise that these reports are a way of finding out about not only inferences and explanations, but also experiments carried out, apparatus used, and practical applications that result, or are aimed at. These things are of just as much importance as, and play a similar role in shaping concepts to, inferences and explanations in which the concepts play a role. It is not clear how features such as the choice of apparatus or method *could* fit into Brigandt's notion of inferential role.

Why does Brigandt limit himself to public language? It seems to be to escape an objection from Fodor (1998) and Becker (1998) (see Brigandt, 2004a: 5, footnote 8). According to this objection, inferences cannot be used to individuate concepts because when two people with different concepts of x draw inferences involving x, they are thereby drawing different inferences. For example, consider the concept DOG. No two people will draw an identical set of inferences involving this concept – for example inferences involving idiosyncratic experiences or particular dogs will obviously differ – so it seems that no two people have an identical concept. This seems to make communication impossible, and it also raises a circularity worry. The concern is that, if everyone has a different concept DOG, there is nothing to tie together all the dog-inferences. Conversely, if we *can* group together all the dog-inferences as those involving the concept DOG, the concept cannot be defined by those inferences on pain of circularity. In general terms, if inferences are individuated by the concepts involved in them, the concepts cannot be individuated by the inferences.

Brigandt replies to this objection by saying that his 'definition of conceptual role is about inferential and explanatory relations between *terms*, words, and utterances' (Brigandt, 2004a: 5, footnote 8) so that two people drawing an inference between the same *terms* are drawing the same inference, even if they have different concepts attached to those terms.

This reply to Fodor and Becker is not open to me, given that I do not want to limit myself to public language, so finding a non-circular way to individuate concepts is a challenge for my view, as is the related communication worry. Answering this challenge is the first desideratum for my account. It is a version of the main problem I will discuss in the last subsection of this chapter, so I will return to this issue there, showing how my account can address the problem.

Why does Brigandt separate reference, inferential role, and epistemic goal as distinct components of conceptual content?

One thing that follows from Brigandt construing conceptual role so narrowly is a need to find a place elsewhere for important features such as practice and the goals of science. This is where his three-part notion of concepts (reference, inferential role, and epistemic goal) comes in. He gives two reasons for separating the three components of conceptual content:

> There are two reasons for recognizing these three components. First, the different components of content (or different semantic properties of scientific terms) are ascribed for and fulfil different philosophical functions. Second, in the course of history a scientific concept may change in any of these components (and one component can change without the others). (Brigandt, 2011: 6).

The second of these reasons does not give Brigandt's account any advantage over mine. For me, different parts of the framework can change over time without other parts changing, and I can talk about that process without ruling some of them out of the conceptual role. Being able to talk about these independently changing features is therefore a desideratum that is easily met by my account. The first point is more interesting. I will discuss first reference, then epistemic goal, with respect to the distinctive philosophical roles they can play, and how my account can provide alternative means to fulfil those roles.

Traditionally, the notion of reference has been part of conceptual content in order to provide a link between concepts and the world, and to make sense of conceptual change without incommensurability. The threat of incommensurability comes about because, if a concept is defined purely by theories about it (or its inferential role, or sense), when those theories change over time, we seem to be unable to say that they were theories about the same thing at all. This leaves us unable to say that, for example, successive theories about electrons were successive theories *about electrons* (see e.g. Brigandt, 2012: 91). This makes comparing theories across time and talking of progress (or regress) problematic. Continuity of reference looks as though it would solve this problem.

However, Brigandt recognises that some continuity can be preserved and incommensurability avoided even when reference changes (see his example of GENE, 2011: 10–12; 2010b). Reference therefore cannot be what performs the function of maintaining continuity, since we can have conceptual continuity without referential continuity. It is this observation that leads Brigandt to invoke the notion of epistemic goal, as I will explain below. The only role left for reference on his account therefore seems to be to provide a link between concepts and the world. This seems to be what he means when he says that 'the traditional conceptual components of reference and inferential role are needed…to account for how concepts make successful practice (verbal behaviour and interaction with the world) possible' (Brigandt, 2011: 7).

While he is right to invoke successful behaviour to emphasise the connection between scientific research and the world, invoking reference as a component separate from conceptual role in order to do this just seems to be mistaken. A connection to the way things are in the world is already built in to any adequate form of conceptual role semantics. In my version, it comes in via Shapere's notion of success. In order to be internalised into a particular science, a component of the framework must be successful, as discussed in chapter 3. Any conceptual role semantics that did not have some notion like this would not be viable because it would allow anything to be part of any framework, with no requirement that the framework be responsive to the world. This link to the world permeates all aspects of the framework on my account, rather than being tacked on to concepts only, as they refer to token things. This is a better reflection of what goes on in science, because it allows me to talk about the link between practice (methods, apparatus, etc.) and the world; although these things are not obviously conceptual, they can be adopted due to a track record of success, and they are important in shaping concepts.

My account can therefore meet the desideratum of showing how concepts are linked to the world, and arguably does so better than Brigandt's. I will return to this point when discussing the advantages of my account. I will now move on from reference to talk about epistemic goal.

I have already mentioned that Brigandt invokes epistemic goal a) to bring in the idea that the goals and aims of science, and criteria for what would count as meeting those goals, are important for shaping concepts and b) to make sense of continuity and avoid incommensurability. I can obviously meet the first of these desiderata within the dynamic framework account, so I will now look at the second in more detail.

Brigandt says 'I introduce the novel notion of the epistemic goal of a concept precisely because it accounts for the *rationality of semantic change and variation*' (Brigandt, 2011: 6, emphasis in original).

Variation in a concept, for example between members of different subdisciplines, is judged as rational if it can be explained in terms of the different groups having different epistemic goals. However, unity can be provided by a more general epistemic goal that is shared (see the example of EVOLUTIONARY NOVELTY, Brigandt, 2012: 79–84). This is reminiscent of my notion of varying levels of grain at which frameworks can be viewed.[26] In general, my account can talk about what is shared at various levels, be they goals, theory, methods etc. I can therefore understand what it is about the epistemic niche of a concept construed at a particular level of grain that unifies it into a single concept, while recognising more fine-grained ways of talking that would explain rational conceptual variation. In fact I have already used the language of the dynamic framework account to talk about conceptual variation when I set up the problem of MEMORY. The levels approach means that, on both my account and Brigandt's, there is no fact of the matter whether two concepts are "the same concept" because there is no unique way of individuating concepts (Brigandt, 2012: 87–88, footnote 6).

[26] Brigandt also refers to variation between individuals and for the same individuals in different contexts (Brigandt, 2012: 96–97).

So much for *variation*. According to Brigandt, the rationality of conceptual *change* is judged according to how well the epistemic goal is met; it should be better met by the new concept than the old in order for change to be rational. Having epistemic goal as distinct from conceptual role allows the rationality of the concept to be judged according to something other than the role which defines it.[27]

Assessing the rationality of conceptual change is also important for my account, and this is another desideratum my account must meet. I will do so by looking at prior changes in the history of the framework, and this strategy will be explained in full in the next chapter. This will establish that my account does not lose the advantage of accounting for the rationality of conceptual change, despite not separating epistemic goal and conceptual role in the way that Brigandt does.

I have so far examined the reasons for which Brigandt introduces his tripartite theory of conceptual content, and his narrowly construed conceptual role. This has brought out certain desiderata for an account of concepts, and I have shown that my account meets some of these as well as Brigandt's does. These include:

- Recognizing that different aspects of a concept can change independently over time.
- Providing a link between concepts and the world.
- Having a role for the goals and aims of science.
- Accounting for the rationality of variation in a concept (e.g. that between different subdisciplines).

There are also two desiderata I have not yet measured my account against: the problem of finding a non-circular way to individuate concepts, which I will address in the last section of this chapter, and providing a method for assessing the rationality of conceptual change, which I will address in the next chapter.

Now I want to claim that my account also offers some distinct advantages. In particular, they are advantages according to the pragmatist terms that Brigandt and I

[27] In the case of GENE mentioned above, the epistemic goal also changes over time (so all three components change) but this is still rational because it takes place gradually (see Brigandt, 2010b).

share – pragmatist in the sense that we are both looking for an account that enables interesting philosophical work to be done.

Advantages of my account

The first advantage of my account is the fact that it distributes the link between concepts and the world throughout the framework, rather than invoking reference as a separate component. I already noted above that this means that my account brings in elements of practice such as methods and apparatus as things that shape concepts, in line with actual science. I now want to argue that I can also use this notion to provide a philosophy of science that could be of use to scientists themselves.

Although Brigandt dispenses with the idea that sameness of reference is necessary for sameness of concept over time, he still talks about sameness of reference as a way to say that scientists are talking about the same thing. But if we think about concepts from the scientists' own perspective, it is not clear how they are to know that they are talking about the same thing, if nothing else other than the referent is shared. When looking at the concepts from the scientists' own perspective, reference cannot be accessed in isolation from role. There is therefore no motivation for saying that reference has stayed the same, other than as something that is inferred based on a similarity in role. From our perspective as philosophers, when we say that reference has remained the same, we are basing this claim on similarity in conceptual role from our own perspective, probably dependent on the role of the current scientific concept. This might be useful for some purposes, but we should be clear that this is what we are doing.

My account says that some aspects of the framework in the epistemic niche of the concept must be the same in order for us to say the scientists are talking about the same thing – the concept must fit into theory and practice in a way that is similar in some respects, or we would have nothing on which to base our claim that the referent is the same. What this similarity consists in could be aims, goals, practices, etc., so could include the things that Brigandt groups under epistemic goal. I base claims about conceptual similarity directly on these similarities in theory and practice – similarities in what I am calling conceptual role – rather than inferring sameness of reference and using that as a criterion. Construing conceptual role broadly in this way makes better sense of scientific practice than invoking reference as a separate component, because it is a construal that could be used by the scientists themselves, and takes their perspective seriously.

Another advantage is that my talk of different levels of grain allows all aspects of the framework to be construed at different levels, rather than just the epistemic goal, as in Brigandt's account. This point is important because it allows us to recognize that scientists sharing a more coarse-grained goal will also share coarser-grained theories, etc., which is an important part of accounting for their apparent sharing of concepts in some contexts and not others (see my example of sharing a framework at different levels in chapter 2). This is another way in which my account of variation in a concept between different subgroups makes more sense of scientific practice than Brigandt's.

If simplicity is taken to be an advantage for a philosophical theory, it is an also an advantage that on my account, all the aspects of the framework are given a definitive part in shaping the concept. All the aspects, including goals and criteria for what would count as meeting those goals, can be part of the epistemic niche which applies pressure to the concept. In particular, the framework already includes practice and a link to the world via the notion of success, so there is no need to bring in reference as a separate component. It also already includes goals, aims and criteria for what would count as meeting them, so there is no need to bring in epistemic goal as a separate component. This gives a simpler and more unified account than Brigandt's tripartite approach.[28]

My account also allows fruitful and diverse work to be done. All the aspects of the framework which apply pressure to the concept also apply pressure to one another, so goals and criteria are shaped by concepts, theories etc. as much as they shape them; I am only focussing on concepts because MEMORY is the subject of my investigation. This means that my methods would be amenable to studying different

[28] I would not want to place too much weight on this, as it would take further argument to show that this kind of simplicity is a virtue of a philosophical account (compare the work in chapter 3 on unification in the cognitive and social sciences).

aspects of the framework than concepts, and such a study should be commensurable with work undertaken on concepts, such as this thesis.

In summary, my account has the same advantages as Brigandt's plus some distinctive advantages of its own.

In other respects, our approaches are quite similar; in particular they both seem to suffer from the same problem, namely that neither Brigandt nor I seem to have any way to demarcate that which is definitive of the concept from that which is not, leading to a sort of radical holism. This is a problem we both inherit from inferentialism, according to which all the correct inferences in which the concept plays a role seem to be definitive. The apparent virtue of expanding the account beyond inferential role if anything only makes the problem more acute for Brigandt and for me. Holism is problematic because no two scientists have quite the same inferential (and explanatory) network, or the same framework. Therefore it seems that no two scientists have the same concept, and it is not clear how they can communicate (Fodor and Lepore, 2007). This, and the related circularity worry, were introduced above: if concepts are used to individuate conceptual roles, it does not seem that conceptual roles can be used to individuate concepts.

This problem of holism and concept individuation is the main one I want to address here, but first I will make a brief detour into the psychology of concepts, suggesting parallels between the account given in this section and the Theory Theory of concepts in psychology. I do this because the Theory Theory seems to suffer from a similar holism problem, and my solution will make use of Khalidi's solution to the problem as it confronts Theory Theory.

Concepts in psychology

Brigandt (2004) discusses the close similarities between conceptual role semantics in philosophy and Theory Theory in psychology, and my version, like his, has much in common with the psychological theory.

Theory Theory arose in the 1980s out of a dissatisfaction with the previously popular Prototype Theory. Prototype theory is concerned with the features or characteristics something must have in order to fall under a particular concept, but instead of being necessary and sufficient conditions, the features in Prototype Theory form a probabilistic list. On this view, there may be no features which all members of a category possess (see e.g. Rosch and Mervis, 1975). This is similar to the Wittgensteinian family resemblance notion of a concept (1953) and Rosch and Mervis specifically acknowledge a debt to Wittgenstein (Rosch and Mervis, 1975: 574–575), even describing their study as ‘an empirical confirmation of Wittgenstein’s (1953) argument that formal criteria are neither a logical nor psychological necessity’ (Rosch and Mervis, 1975: 603).

The probabilistic list results in a graded model of categorization where some instances possess more features from the list than others, i.e. they are more prototypical of the category. For example, a robin is a more prototypical bird than a penguin because it has more of the typical bird features, such as “sings”, “has feathers”, “can fly” etc. (or possesses them to a higher degree for the kind of features that admit of degree).

In the 1980s, psychologists began to suggest that Prototype Theory was inadequate because concepts cannot be made coherent without use of our background theories; similarity between instances of a concept alone is not enough to provide coherence (Murphey and Medin, 1985). There are also various features of our categorization behaviours, such as conceptual combination and conceptual change, that do not fit with the model of a self-contained feature list (see Keil 1989: 44 for a good brief summary). It may seem obvious to us that we use theories in categorization, but we overestimate ourselves in this regard (Keil 2005: 315–316) so our intuitions are not reliable here. The emergence of evidence to this effect was therefore crucial.

In contrast to Prototype Theory, in Theory Theory each concept is not self-contained, but linked to other concepts and our relevant theories about them. These theories need not be (and often are not) fully worked out, rigorous theories. Murphey and Medin say:

> When we argue that concepts are organized by theories, we use *theory* to mean any of a host of mental “explanations”, rather than a complete, organized scientific account. For example,

> causal knowledge certainly embodies a theory of certain phenomena; scripts may contain an implicit theory of the entailment relations between mundane events; knowledge of rules embodies a theory of the relations between rule constituents; and book-learned, scientific knowledge certainly contains theories. Although it may seem to be glorifying some of these cases to call them theories, the term connotes a complex set of relations between concepts, usually with a causal basis. Furthermore, these examples are similar to theories used in scientific explanation (Achstein, 1968). (Murphey and Medin, 1985: 290).

Machery separates two different types of Theory Theory in the literature (Machery, 2009: 101). According to one of these, concepts *are* theories. According to the other, concepts are elements in theories; as Khalidi puts it, concepts are enmeshed in a theoretical network (Khalidi, 1995: 402). Machery notes that some psychologists, including Murphey and Medin quoted above, slide between these two versions (Murphy and Medin, 1985: 298, cited in Machery, 2009: 101). It is primarily the second of these varieties of Theory Theory that I am interested in here, where the "theoretical network" in which concepts are enmeshed is the framework.

As I have said, all the elements of the framework surrounding the concept – that make up its epistemic niche – shape the concept. The concept is defined by its role in the framework. This seems to be very similar to Theory Theory, but it is often assumed that the psychology of concepts and the philosophy of concepts are engaged in different projects. For example, Machery claims that psychologists 'are interested in the properties of the bodies of knowledge that are used by default in the processes underlying the higher cognitive competences' (Machery, 2009: 34, see also Peacocke 1992 cited here). By contrast, philosophers 'are typically interested in what conditions have to be fulfilled for having attitudes about the objects of our attitudes.' (Machery, 2009: 34–35). Khalidi says more succinctly that '[p]hilosophers discuss meaning, psychologists concepts. Psychologists experiment with subjects, philosophers speculate about agents.' (Khalidi, 1995: 402).

Philosophers are taken to be interested in the question of how meaning attaches to the world, and the correctness conditions for applying particular concepts, while psychologists are interested in the mechanisms behind actual instances of concept application. The philosophical project seems normative, where the psychological one is descriptive. In what sense, then, can they be compared?

Brigandt argues that both approaches agree on the structure of concepts and some basic goals of a theory of concepts—'the explanation of behavior and conceptual performance, and the explanation of the change and development of concepts...' (Brigandt, 2004a: 1). This seems right. It may be that philosophers are interested in an additional normative question about concepts that does not usually concern psychologists, but this does not mean they are not talking about the same thing.

More than this, it seems that the two types of question are importantly related. I am interested in the normative question of concept application (when it is right or rational to apply a situated cognition concept of memory) but in order to address this question, we need to know something about where those norms come from. This involves knowing something about how concept application works. We need at least a plausible way of construing a mechanism before we can assess how well it is functioning. The descriptive question is therefore important for answering the normative one.

In addition, it is not completely clear that psychologists are purely interested in mechanistic and not agential questions. Khalidi suggests that the Prototype Theory involves viewing the human cognizer from the *design stance* and treating them as designed to behave in certain ways in certain circumstances, while the Theory Theory involves taking the *intentional stance* and treating the human cognizer as an agent with rational beliefs (Khalidi, 1995). If he is right, psychologists employing the Theory Theory are not so different from philosophers in how they view the objects of their study.[29]

Khalidi also mentions the explicit Quinean influence on Theory Theory (Khalidi, 1995: 411), and discusses links to holism. He uses a Quinean/Davidsonian strategy to get out of the holism and communication problem, further strengthening the link between Theory Theory and the philosophy of concepts. I will return to his strategy in the next section.

[29] If Khalidi is right, Theory Theory and Prototype Theory may be compatible. It is the main purpose of his (1995) paper to claim that they operate at different levels of explanation, so need not be in conflict. I will not take a position on this issue here.

A further apparent difference between my work and that of most psychologists is that psychologists are typically interested in an individual's concepts, whereas I am interested in the concepts of a scientific community. I think that this is the same problem, considered at different levels of grain. As I said in chapter 2, I am looking at the level of the scientific subdiscipline, and at the variation between these subdisciplines, but I could equally have chosen a different level of grain if my aims were different, including the level of the individual scientist. On the account I develop here, it makes sense to treat all these cases as examples of the same phenomenon. Therefore psychologists' work on individuals' concepts is very much relevant to my work on the concepts held by a scientific subdiscipline. I will say more about the comparison between individuals and groups in the next section.

The important similarity between Theory Theory and conceptual role semantics for my purposes is that they both suffer from the problem of concept individuation. I put the problem of concept individuation for conceptual role semantics in terms of a kind of holism where no two scientists have quite the same framework. This seems to mean that conceptual roles, and therefore concepts, cannot be shared. Therefore it is not clear how the scientists can communicate. I also introduced the related circularity worry that if concepts are used to individuate conceptual roles, it does not seem that conceptual roles can be used to individuate concepts.

Theory Theory also has a problem with specifying what constitutes the concept in a non-circular way. Keil says of Theory Theory that 'concepts may only be understood in terms of the theories they are embedded in and theories only in terms of the concepts they embed' (Keil 1989: 49). So many of our theories can potentially be involved in categorization, it becomes difficult to individuate the concept, particularly since there is no clear way of individuating theories (Fodor 1994: 110–111). This also seems to threaten a kind of radical holism according to which concepts can never be shared because the whole network of theories can never be shared.

I will consider both the literature on Theory Theory and on conceptual role semantics in the next section, as I try to dissolve this problem.

The problem of concept individuation and communication

It is usually thought that people need to share a concept of x in order to communicate about x. Therefore we seem to need a criterion for being-the-same-concept-as, i.e. for concept individuation, in order to say that two people have the same concept, and thus explain their successful communication.

In terms of my theory, because of the interconnectedness of the framework, it seems that the whole framework must be shared for successful communication. This radical holism seems to make communication impossible. Worse still, even if it were somehow possible, it does not seem actual, because part of the problem I am interested in is the *variation* in MEMORY. As I have said, this variation does cause communication problems, but an account that says those problems are so severe that each subdiscipline (or even each individual) has a completely different concept of memory cannot be right. Communication, however problematic, does take place within and between the sciences of memory.

There are three possible solutions to this problem:

1) Holism: The concept is individuated by the whole framework. Bite the bullet and admit that no two scientists ever have the same concept, but argue that communication can still take place.
2) Share part of the framework: Some particular part of the framework, or a sufficient amount of it, must be shared in order to say that two scientists have the same concept.
3) There is no single answer as to how concepts are individuated or what it means to share a concept; it depends on the level of analysis.

Solution 1 is strictly speaking not a solution, but a denial that the problem is a problem at all. It does raise the important issue that there is always something partial about communication – some level of grain at which important aspects of the framework are not shared – so there is always room for miscommunication. However, the purpose of concept individuation conditions is to allow us to say that two scientists debating the nature of some phenomenon such as memory are actually having a debate and not talking past one another. Solution 1 does not help us to do that. If it is not sharing concepts that allows communication, what is it?

Solution 2 is the solution taken by some proponents of Inferentialism. A certain sub-set of the inferences are taken to be the meaning-constitutive ones – the ones that must be shared in order to say that the concept is shared. In the literature, it is widely agreed that this could work if there was an analytic-synthetic distinction; the meaning-constitutive inferences would be the analytic ones. However, it is also widely agreed that Quine (1976) has shown that there is no such distinction (Fodor and Lepore, 2007).

In the absence of an analytic-synthetic distinction, Brandom says that concepts don't need to be completely shared; *similarity* of inferences involved for each concept user is enough (Brandom, 2007: 665–666). This seems to be on the right track, and is applicable to my account. In order to share a concept, scientists' frameworks must be *similar enough.* But by itself this is not very illuminating; we have just replaced one vague notion with another.

In particular, the notion of similarity must be cashed out in a non-circular way. The problem is deciding which part of the framework, or how much of it, must be shared. We cannot just say that for two people to share the concept DOG, they must share the dog-inferences, or the theories or aspects of the framework relating to dogs, because we cannot define these independently of the concept. We also cannot say that what must be shared is the particular parts or the amount of the framework that needs to be shared to allow successful communication, because successful communication is what we are trying to explain.

Solution 3 is my favoured solution, but it is not immediately obvious that it is a solution, so I will spend the rest of this section unpacking it. I will begin by considering Brigandt's answer to the holism problem, which is also a variety of solution 3. My solution can be seen as an improved version of his work, and one which better fits my purposes in this thesis. My improvement on Brigandt's approach will make use of Khalidi's solution to the concept individuation problem for Theory Theory that I referred to above.

Brigandt posits two levels of content for concepts. At the individual level, he embraces holism, but at the level of communities, he says that similarity of concepts is sufficient to say that concepts are shared. The major advantage of this strategy is that differing concepts at the individual level allow us to account for differing behaviours between individual scientists, whereas sharing at the level of communities allows us to account for communication between them.[30] Whether we focus on the level of the individual or the community depends on whether we want to explain differing behaviours or successful communication.

On the individual level, Brigandt says: 'If two scientists have a different conception of genes and thus on my account associate a different meaning with the term "gene", then due to their different conceptions they may make different theoretical claims and conduct different experiments.' (Brigandt, 2004a: 4). In a sense, this is accepting a role for radical holism (solution 1 from my list).

Despite embracing holism at this level, some aspects of conceptual role will be shared between two individuals, so similarities in behaviour can be explained as well as differences:

> [O]nly a certain part of the total conceptual role is important for a particular situation. A layman and a Drosophila geneticist have very different conceptions or 'concepts' of a fly. But when we explain how they succeed in catching a fly, we just need to make recourse to a few shared beliefs about flies that are sufficient to explain their behavior, such as the assumption that flies can fly. So holism and variation between the content of individuals does not prevent us from giving intentional explanations. The total conceptual role is an important resource for a whole range of different explanations. Any difference in individual content may feed into some explanation (Brigandt, 2004a: 4).

Although the entire conceptual role individuates the concept so that no two individuals have the same concept, we can appeal to partial sharing for certain explanatory purposes. What must be shared is different depending on what those explanatory purposes are, so there is variation in what it is important to share.

[30] Brigandt's is not the only theory positing two levels like this. Galison, in his work on pidgin languages in science, talks about 'representing meaning as locally convergent and globally divergent' (Galison, 1997: 47).

At the level of the community, Brigandt says:

> I view a concept as a cluster of similar individual meanings or conceptual roles. Taking a concept as a group-level entity abstracts from the inter-personal variation and focuses on the more substantial difference between different concepts. Thus I follow Harman (1973), Block (1986), and Jackman (1999) in assuming that merely similarity, not necessarily identity in conceptual role is sufficient to share the same concept. (Brigandt, 2004a: 4).

This does not seem too dissimilar to the picture at the individual level in that we can choose to focus on certain similarities in conceptual role for certain explanatory purposes. However, at the community level, this difference in explanatory purposes corresponds to a difference in conditions for concept individuation:

> [F]or the study of conceptual change in science pragmatic and case by case criteria for the individuation of concepts can be used. I assume that two terms can be viewed as corresponding to two distinct concepts in case they make inferences or explanations possible that are relevantly dissimilar. What counts as relevant is dependent on the scientific standards of the given situation… The point of my approach is that while I stress that the variation of individual mental representations is real, the existence of clusters is real as well. And we can pragmatically pick out some clusters for a certain philosophical purpose, and compare these concepts and explain and assess the origin of these conceptual differences (Brigandt, 2004a: 5–6).

It is not clear why similarities should constitute concept individuation conditions at the community level, but not at the individual level. I will aim to give a more unified account in this respect.

Brigandt is also a little unclear in the above quotation as to whose interests the individuation conditions are dependent on. He says that '[w]hat counts as relevant *is dependent on the scientific standards* of the given situation', but also that 'we can pragmatically pick out some clusters *for a certain philosophical purpose*' (my emphases). Is it our purposes as analysts that is important, or the scientists' purposes as actors? Clarifying this issue is another way in which I will try to improve on Brigandt's account.

I also need to show that the changes I have made to Brigandt's account in the first section of this chapter – broadening conceptual role to include the whole framework, going beyond just public language to include practice, and including both goals and criteria and a link to the world as parts of the conceptual role, rather than having reference and epistemic goal as separate components of meaning – do not prevent me from using the main insight of Brigandt's approach. We now have several desiderata for the solution to the concept individuation problem. The solution will involve referring back to my notion of levels of grain, and discussing Khalidi's approach to the individuation problem.

According to my account, a relatively coarse-grained framework needs to be shared between two scientists in different disciplines for them to count as sharing concepts. Two scientists in the same discipline need to share a finer-grained framework, and two in the same research team need to share a finer-grained framework still. Rather than just having two levels – individual and community – we have a continuum of progressively finer-grained sharing.

Recall the example from chapter two of a psychologist who says "we have discovered that a rat's memory for a maze is better after consuming caffeine" to both a colleague in her research team, and to a physicist friend. I said that the two psychologists share a framework at quite a fine-grained level. Working in the same research group, they have shared goals, knowledge of methods and practices, etc. The psychologist and the physicist however share a much coarser-grained framework. I said that my interest is at the level of subdisciplines, and at this level communication has succeeded in the first case because the framework is shared, but not in the second case. The psychologists can be said to have the *same concept* of memory, while the psychologist and physicist do not. However, working at a more coarse-grained level of analysis (for example science as a whole), the psychologist and physicist could be said to share a concept. Concept individuation conditions therefore vary with level of grain, and the relevant level is set by our aims as analysts. I want to find out about variation between subdisciplines in the sciences of memory, so my goals set the level of grain for my project to that of the subdiscipline.

It may look like this response hasn't advanced things very much. There is no level at which an *entire* framework is shared. The interconnectedness of the nodes in the framework still seems to push us towards radical holism. However, this objection misunderstands my solution.

Recall that a framework contains aims and goals for the science, (construed at the appropriate level of grain). What it is to share a framework at the relevant level of grain is to be able to employ the concepts, theories, methods etc. at that level to meet the goals specified at that level. There is no need to agree on all commitments construed at that level. Whether communication succeeds is therefore taken on the terms of the scientists communicating, but without allowing that anything goes. This completes the clarification of whose aims are relevant for concept individuation; it is the analyst's aims that set the level of grain, but within that, it is the scientists' own goals as actors according to which their concepts should be judged. This fits well with Shapere's insight that we should not impose a metascientific standard on science.

There is still a practical problem for the philosopher: How do we go about carrying out this judging, when we stand outside the framework? And how can we break out of the circularity using scientists' goals, given that these are composed of concepts, are theory-laden, etc.? This is where Khalidi's work comes in.

As I said above, Khalidi suggests that the Prototype Theory involves viewing the human cognizer from the *design stance*, while the Theory Theory involves taking the *intentional stance* and treating the human cognizer as an agent with rational beliefs. When it comes to the Theory Theory's problem with individuating concepts in a non-circular way, Khalidi makes use of this idea that we are treating the subject as a rational agent. Inspired by Quine and Davidson, Khalidi 'locates the notions of meaning and belief in the process of translation or interpretation' (Khalidi, 1995: 411).[31] He draws on the Quinean and Davidsonian idea of radical interpretation,

[31] Brandom also makes an appeal to Davidsonian interpretation (Brandom, 1994). Fodor and Lepore, in their review of Brandom, say that '[t]here is…considerable irony in the spectacle of Brandom, the arch-inferentialist, appealing to Davidson, the arch-Tarskian, in hopes of saving his bacon. Brandom's appeal to Davidson here sounds to us a lot like panic.' (Fodor and Lepore, 2007: 192, footnote 13). There is something importantly different about inferentialist versus Tarskian semantics, but I see no

according to which translation between two frameworks can come about by adopting a principle of charity or similar (see Khalidi, 1995: 411, footnote 7) and assuming that your informant is at least minimally rational. This involves assuming a certain fit between the informant's meanings and beliefs into a coherent whole. Then, as Khalidi says:

> We aim at constructing a mapping between our respective vocabularies that exhibits a certain overall fit. This overall fit concerns a term's position in the linguistic practice of the agent being interpreted, in short, its place in that person's entire corpus of beliefs and intentional actions (Khalidi, 1995: 412).

This approach provides a way of interpreting a scientific framework from the outside, even if only by finding our way inside. By taking the intentional stance to scientists (as, if Khalidi is right, we do when adopting the Theory Theory of their concepts), we try to make sense of them as rational agents, and can therefore find a way in to understanding their goals, concepts, theories etc. despite holism. In Khalidi's words:

> The interpretation is not being driven by the presence of a requisite set of beliefs or features but by the need to make overall sense of the informant in an intentionalistic idiom. In other words, the ascription of concepts is subordinated to the need to make sense of the rational agent; the agent is not viewed merely as a complex feature detector, as on the design stance (Khalidi, 1995: 413).

We must become acquainted with many aspects of the framework and begin to see how they fit together before we can come up with an interpretation.[32]

In fact, our job is considerably easier than that of the radical interpreter. There will be a (relatively coarse-grained) level on which we share a framework with the

---

reason why this should stop an inferentialist making use of the idea of interpretation. Doing so need not import any other commitments of Davidson's.

[32] This interpretive solution is not dissimilar to Galison's notion of a "Trading Zone" (Galison, 1997), and in fact Khalidi uses the analogy of an economy, where interpretation is compared to working out both the exchange rate and the value of particular goods, with nothing to go on other than an assumption that there is a coherent economic system at work (Khalidi, 1995: 415–416). A key difference is that, in a Trading Zone, two groups of scientists are interpreting each other in order to work together, whereas in mine and Khalidi's version, we are interpreting the community (in my case the scientists) from the outside as analysts.

scientists we are interpreting, and we can make that sharing somewhat finer-grained by further study. It is appeal to this coarser-grained sharing that gets a principle of charity off the ground for us. The reason we can treat scientists as minimally rational is that we share some idea of what it is to be rational. We thus already share enough to quite adeptly take the intentional stance to scientists and come to understand their frameworks well enough for our purposes.

There are several advantages to the interpretationist way of thinking. One is that we have no need to share Brigandt's assumption that we are dealing purely with public language. Practices are as much open to interpretation as concepts and theories. The advantages of expanding the conceptual role beyond just inference and explanation can therefore be had without the cost of exacerbating the holism problem, as I suggested in the first section of this chapter.

Another advantage is that the interpretationist strategy makes sense of how goals, criteria and fit with the world can also be brought in as part of the conceptual role, rather than being added as extra components (Brigandt's epistemic goal and reference). Consider for example Quine's radical interpretee pointing at a rabbit. While what he is pointing at is radically underdetermined, the coherence of the overall picture that emerges on further interactions with the interpretee must fit with the world. To put it simply, the pointing brings the world into the picture. Other things than pointing, such as apparatus and experiment, can perform the same function when we are interpreting scientists.[33]

A major objection to this way of thinking might be that it assumes what philosophers of science are often trying to assess, namely the rationality of the scientists under study. If this is right, it would make the method inapplicable for my project (and many others).

---

[33] Some empirical work (Goldstone and Rogosky, 2002) suggests that translation between two conceptual systems can be achieved based only on within-system relations, not extrinsic information about the world. However, this work also found that translation is much easier when both intrinsic, within-system information and extrinsic information are used. Brigandt cites this work in a footnote, saying it shows 'the fact that basing concepts on similarities of syntactic entities is possible' (Brigandt, 2004a: 5, footnote 8). It may also offer support to the interpretationist strategy. Much more would need to be done in order to cite this as evidence, but it is interesting and suggestive.

I do not think this is an insurmountable problem. What we assume is a fairly minimal kind of rationality; just enough to allow us to make sense of the subject from the intentional stance. This does not mean that every individual decision or research direction taken must be the most rational it could be, and therefore immune to criticism.

In fact, there is an important sense in which we must assume scientists have some level of rationality at the outset. Shapere points out that terms such as "rational" (and "knowledge") have a use. A major part of that use is in describing science. Shapere offers this point as part of a general criticism of relativism. He says:

> Those relativistic views [held by the critics of logical empiricism] have, I believe, been effectively criticized; and, standing above all the criticisms, is the point that science is, after all, a paradigm case of the knowledge-acquiring process. To deny that science and its development *can* be rational – a denial that seems to be the conclusion of the relativist position – fails to recognize that the terms "rational" and "knowledge" have a use. It is a condition of the adequacy of any philosophy of science that it show *how* rational change in science is possible, and a philosophy of science which, after asking *whether* scientific change can be rational, denies that it can be, must be rejected (Shapere, 1977b: 200).

For Shapere, what we mean by "rational" is something that has grown out of scientific development (one of our paradigmatic rational processes) and has changed over time as science has changed.

In my framework language, there is a standard for rationality which has emerged from the coarse-grained framework of epistemic enquiry that we all share (and of which science is a major part). The world and our means of epistemic engagement with it have allowed a conceptual role for rationality, and therefore a concept of rationality, to emerge. In a sense, this is the most fundamental step in Shapere's learning how to learn (see chapter 3). It may be that the concept of rationality could have been otherwise, it may not be completely precise, and it may change, but given the position we are in, it is available for us to appeal to, and so we can get a principle of charity off the ground. In practice, in most (perhaps all) cases, much more will be available to us than this, because we will share a more fine-grained framework with those we are trying to interpret than that found at the level of epistemic enquiry in general.

We cannot lay down a standard of rationality from a metascientific perspective, and then check science against that. However we can make use of the standard that has emerged out of science itself, along with any more fine-grained aspects of the framework that we share, to examine on their own terms the factors shaping a particular concept. It is the examination of these factors according to the scientists' own terms that I intend to carry out in this thesis, but with the focus of the analysis (MEMORY) and the level of grain (the subdiscipline) set by my own aims as an analyst.

Conclusion

I have presented a version of conceptual role semantics in keeping with the view of science laid out in chapter 2. According to my view, a concept is defined by its place in the dynamic framework, or its role within that framework, at the relevant level of grain.

I have compared this account to Ingo Brigandt's version of conceptual role semantics in philosophy of science, arguing that my account is better because it allows all of the important aspects of the scientific enterprise to shape the concept (to be part of its epistemic niche in my terms), rather than being limited to inferential and explanatory role. This means we have no need to separate reference and epistemic goal as additional components of conceptual content as Brigandt does.

The middle section of this chapter briefly introduced the psychology of concepts, in particular the Theory Theory, and its similarity to conceptual role semantics. The importance of the Theory Theory for my purposes is that it seems to suffer from the same problem of radical holism and impossible communication as conceptual role semantics. This problem is the major challenge facing accounts such as mine.

I have argued that there is no single solution to the problem, but that how concepts are individuated depends on the level of analysis. In particular, the level of grain at which the framework is viewed depends on the analyst's aims, but whether communication at that level succeeds depends on the scientists' aims. Concept application in science is therefore to be judged on the scientists' own terms, in line with Shapere's rejection of imposing metascientific standards.

Khalidi's interpretationist strategy to the concept individuation problem for the Theory Theory explains how we can analyse concepts despite standing outside the framework. This strategy involves treating the concept users as intentional agents, and assuming they have a minimal level of rationality. I have argued that the concept of rationality involved in doing this has arisen from epistemic enquiry, particularly scientific enquiry, and is therefore a concept we share with scientists at this coarse level of grain.

Now that I have given a way of talking about science (the dynamic framework account) and the place of concepts within that account (my version of conceptual role semantics), I have the vocabulary and background needed to explain my intended methods in detail, and this will be the job of the next chapter. This will explain how, given the tools I now have, I can assess case studies to see whether any of them are employing a situated cognition concept of memory, and if so, whether it is functioning as an investigative kind (and is therefore a legitimate concept). My methods for identifying such legitimate concepts will refer back to this chapter, making use of the interpretationist strategy, and fulfilling the task I deferred above of accounting for conceptual change as rational. This will allow me to make normative claims about scientific concept application.

5. Methods

Introduction

I said in chapter 1 that I would be using a historically situated case-study-based conceptual ecology to analyse the concept(s) of memory that scientists employ in their research. I set out my intention to look for situated cognition concepts of memory that are in use in current practice, and assess the extent to which they are legitimate, i.e. whether they are functioning as investigative kinds (an epistemological question about how the concepts function in scientific theory and practice). In this chapter, I will describe my methods for doing this in detail, before going on to apply those methods in the next three chapters.

I will be making use of the dynamic framework account introduced in chapter 2, but introducing a way of making normative assessments of frameworks. Recall that the framework is a network of theories and practices (e.g. goals, methods, concepts, criteria for what counts as an observation, criteria for what counts as a good explanation, etc.) that mutually influence and constrain one another. Scientists can be grouped at different levels, sharing frameworks at different levels of grain. I will predominantly be interested in the level of the subdiscipline.

Frameworks change over time. I am particularly interested in the change in MEMORY (its conceptual development) but this is constrained by surrounding aspects of the framework, so cannot be looked at in isolation. I am calling those aspects of the framework that constrain or apply pressure to a particular concept the *epistemic niche* of that concept. These aspects of the framework define the concept according to a variety of conceptual role semantics (see chapter 4).

As I said in chapter 2, different subdisciplines studying memory have different but overlapping frameworks, which may well result in different concepts of memory. More than one of these concepts may function equally well as investigative kinds, so my methods must be open to the possibility of finding this. In chapter 3, I offered some support for the likelihood of this pluralist outcome, arguing that the sciences of memory should not be using unification for its own sake as a criterion to shape their theories.

This was part of a more general discussion in chapter 3, concerning the fact that many of the sciences of memory are immature. I will be returning in this chapter to the picture of immaturity I outlined there, according to which an immature science is one which does not yet have a clearly delineated internal/external distinction, in Dudley Shapere's sense of that distinction. The concepts of memory in many of these sciences are weakly constrained by internal factors. In other words, they are at an early stage in the process of conceptual development, so the epistemic niches for MEMORY found in the various subdisciplines can be expected to contain few internal considerations and many untried hypotheses. Despite this, I said that it is important to try to begin to answer the situated cognition question because of the importance of memory research, and that doing this must involve looking at the criteria that *have* been internalised, and whether they constrain MEMORY sufficiently to determine whether it should be a situated cognition concept, and if so, of which type. I will flesh out that idea further here.

My methods for identifying legitimate concepts will also refer back to chapter 4, making use of the interpretationist strategy it introduced, and fulfilling the task I deferred there of showing how we can account for conceptual change as rational. In particular I will be interested in any situated cognition concepts I find, and whether the conceptual change involved in adopting them was rational. This will link the interpretationist strategy to Shapere's work on rationality and the internal/external distinction, leading me to advocate using historical analysis of the epistemic niches for MEMORY in particular subdisciplines.

Investigating the sciences of memory: Descriptive and normative projects

My project can be divided into a descriptive project and a normative one, where the descriptive project is to find out what concepts are in use in particular pieces of research in the sciences of memory, and the normative project is to find out whether these concepts are legitimate, and make recommendations for scientific practice accordingly. The descriptive project is therefore the search for situated cognition concepts of memory, and the normative project is to assess whether any such concepts I find are functioning as investigative kinds.

Recall that the question of whether a concept is functioning as an investigative kind is the question of whether it can support fruitful science, a question about how the concept functions in scientific theory and practice (e.g. in giving explanations). As we have seen, the criteria for what counts as fruitful theory and practice (e.g. criteria for what counts as a good explanation) are also part of the framework, and emerge from the scientific research as the framework changes over time. Answering both the descriptive and normative questions will therefore involve some consideration of the framework more broadly.

In the next section, I will discuss two possible ways of tackling the descriptive question: experimental philosophy, and a case study approach. I will advocate the second of these. In the following section, I will introduce the normative question in more detail, and propose an historical approach to answering it. This will involve looking at how the framework (in particular the parts of it forming the epistemic niche of MEMORY) has developed over time. In my ways of tackling both the descriptive and normative questions, I will be treating the scientists as rational intentional agents, and trying to understand the frameworks on their own terms, in the manner of the intepretationist strategy introduced in chapter 4.

The descriptive project

Experimental philosophy

Experimental philosophy is a burgeoning field which now encompasses a variety of techniques including citation analysis of papers, observation of participants' behaviour in a laboratory, and analysis of questionnaires. It is this last method I will focus on here because it has been used already by Stotz, Griffiths and Knight (hereafter SGK) (2004) to study the variation in GENE, a similar problem to our issue with MEMORY. They refer to their method as a variety of "conceptual ecology".

The essence of this method is testing participants' responses to vignettes designed to probe a particular concept, and comparing those responses to various demographic data about the participants. Correlations can then be sought to see whether, for example, there are typically differences in concepts possessed by people of different races, genders, etc.

This kind of experimental philosophy can be used to survey a range of subjects to find out whether there are significant disagreements between their concepts, so it seems to be ideal for the kind of variation found in the cases of GENE and MEMORY. For the version of this approach practiced by SGK, scientists from a range of backgrounds were presented with a range of examples in a questionnaire, and had to decide which of them conformed to their concept of gene and which did not, or which of several examples best conformed. A large number of scientists' concepts were probed and correlated with data about the subdisciplines in which they were trained and in which they worked. Other information such as age and gender were also collected to build up a full picture of the variation.

Importantly, SGK say that a scientist's implicit concept (the concept he actually uses) may be different from his explicit one (the concept he thinks he uses). Their questionnaire is very carefully designed to account for this. As well as "direct" questions about the definition and function of the gene and the methodological value of the concept, the survey contains "indirect" questions. Here, the scientists are required to actually apply their concept of the gene, rather than just answer questions about it.

As well as describing the variation in the concept, SGK's work explains this variation in terms of the epistemic niche inhabited by the concept. It is from this paper and other work by Griffiths and Stotz that I took the term "epistemic niche"; for them the epistemic niche is the needs the group of scientists have in their investigation (Griffiths and Stotz, 2008: 508), although they do not share my framework account of how to cash out what the concept is needed to do.

Here, I will argue that there are some problems with the experimental method, the biggest of which being the risk of artefacts of the method contaminating the results. The problem, in short, is that scientists answering questionnaires know that they are being tested.[34] The situation of answering a questionnaire is very different from the situation of designing and carrying out an experiment. In fact the situations are different enough to constitute partially different epistemic niches; we should

[34] This is a general problem for questionnaire-based experimental philosophy. I do not think it is a fatal one, but something that should always be borne in mind in this kind of work.

therefore suspect that different concepts may be most appropriate in each case. This is the very phenomenon we set out to investigate in the first place.

Using the language of frameworks, we can expect fine-grained variation in the framework when context changes. Although the coarse-grained framework remains the same, there will be fine-grained differences in the epistemic niche between the case where the scientist confronts a vignette in the questionnaire, and a case where she encounters exactly the same situation in real life.[35] These differences will prevent us from clearly seeing the differences between concepts for members of different subdisciplines (or genders, or ages, etc.) Therefore another way of putting the problem with the experimental method is that it does not allow us to isolate variations at a particular level of grain.

This objection, and the idea of levels of grain that it draws on, take the notion of the epistemic niche beyond that discussed by SGK. In their work, conceptual variation is based on an evolutionary model of divergence over time from a common conceptual ancestor. On the view I am advocating here, concepts may also vary with their context of application, and we can see why when we look at the epistemic niches for those concepts at a fine enough level of grain.

If this is right, then it is a serious concern that the questionnaire context is not the same as the normal context of research. The epistemic niche may be different in two ways. For one thing, all the factors that are part of the niche in scientific research will not be fully replicated in the questionnaire context. The needs of scientists employing the concept (the niche) are partly constituted by a wider research context, not just the narrow issue at hand on an occasion of concept application. A question in a questionnaire cannot entirely replicate this because it is not part of a wider research project. Secondly, there may be extra factors in the niche in the questionnaire context, such as giving a particular impression to the experimenter assessing the answers. Those who chose to fill in the questionnaire must have had some motivation for doing so, and this may well have been a factor in shaping how they

[35] In fact, the epistemic niche will also vary in different real-life cases where the scientist confronts the case from the perspective of involvement in different research projects. See the case discussed in chapter 2 for the kinds of factors that might vary from one research project to another.

answered; on this occasion they needed the concept to help them communicate a certain idea about their field of research.

I am arguing that we shouldn't necessarily expect a single scientist's concept to be stable across different contexts. Others do raise worries about ecological validity with respect to experimental philosophy (e.g. Knobe and Samuels, 2013: 84). However, it is not clear whether they have in mind something like this contextual variation resulting in multiple concepts, or a somewhat different worry that the experimental setting does not allow us to adequately capture a single concept.

SGK's acknowledgement that a scientist's implicit concept may be different from his explicit one seems to show that they accept that when a scientist is asked a direct question, he may at least *apply his concept differently* from when he is applying the concept in his work. I am arguing that he may in fact be *applying a different concept* when the individuation conditions are considered at a fine enough level of grain (compare chapter 4 on the topic of concept individuation). In other words, the very fact of answering a questionnaire may be enough to cause the concept to vary from that usually applied in scientific work.

It is taking seriously the experimental philosophers' own idea of the epistemic niche that brings this issue into focus. In particular, it is considering what this niche is made up of. I have argued that it is composed of aspects of the framework. These vary with context, so the pressures that the scientists' concepts respond to will also vary.

There is another disadvantage of experimental philosophy that affects its application to my project specifically, namely that it does not allow one to do the detailed historical work that I will advocate. I will argue later in this chapter that such work is necessary for the normative project.

All this is not to say that the experimental approach is useless. One important thing to come out of SGK's work is the distinction between implicit and explicit concepts. It is true that a scientist's concept need not be what he thinks it is, and in the case study method I advocate in the next section, I will discuss how to get at the implicit concepts.

Experimental conceptual ecology provides results that are very suggestive of the extent of conceptual variation and the factors that cause it. It has been very valuable work in the case of GENE, where such variation needed to be demonstrated. However, for MEMORY, it is widely known and hard to deny that there is such variation. What we want is a more detailed analysis of that variation, and there are reasons to think that the results of the experimental approach may not be an accurate reflection of the variation in actual scientific practice at the finer level of grain required for this. The case study method aims to get around this problem.

Case studies

In this section, I will discuss a case study approach that allows us to look at current concepts in use in science. In the next section I will recommend improving upon it by taking an historical approach, which will allow me to carry out the normative project.

The case study method involves looking closely at particular pieces of current or recent research. One way to do this is by interdisciplinary collaboration with scientists; another is by detailed study of published papers. Interdisciplinary collaboration between philosophers and cognitive scientists is beginning to be carried out in the study of memory, so this is perhaps an appropriate method here (see Sutton 2004, 2007; Craver, 2002; Barnier and Sutton, 2008). But for collaborative work, in order to understand and analyse concept use, the philosopher needs to become proficient in the vocabulary and practices of the sciences in question. In Harry Collins' terms, she needs to develop contributory expertise in the science (Collins and Evans, 2002).

The high level of expertise required of the philosopher is a huge challenge when studying MEMORY because of the sheer diversity of disciplines. The question at issue is best addressed by comparing and contrasting a variety of subdisciplines in search of situated cognition concepts. The philosopher would therefore need a high level of expertise in a range of contrasting sciences. She would need to be fluent at shifting between the languages and concept use in different subdisciplines in order to understand the similarities and differences between them. This is a very tall order although, if there are any individuals with such expertise, this kind of active interdisciplinary collaboration is a very good way to find out what concepts different groups of scientists are using. The research being done in such an interdisciplinary vein is promising, suggesting such expertise is possible through collaboration, although probably no single individual could develop enough expertise in all of the sciences of memory to carry out a full survey.

Here I will be taking the second option of looking closely at published papers. While this still requires a good understanding of the work being studied, it does not require the same level of expertise because there is no need for the philosopher to actively participate in the research. The reduced need for depth of understanding allows for more breadth of understanding which is an advantage given that the nature of the project means getting familiar with various different research frameworks.

Case study work can be done within the broader methodology of conceptual ecology, and SGK do mention the possibility of carrying out conceptual ecology by looking at published work from different scientific fields (SGK, 2004: 648). The emphasis on the epistemic niche can be retained, giving us a way to talk about the reasons for scientists having the concepts that they do. This is the approach I advocate because it retains the advantages of the experimental work discussed in the previous section (the idea of the epistemic niche, and separating implicit and explicit concepts), while improving on it in an important respect, namely that it studies the use of concepts in the normal course of scientific practice rather than in a questionnaire setting. It studies scientists "in the wild" to borrow Edwin Hutchins (1995) phrase. This means that the concepts identified are closer to those actually used by scientists. The epistemic niche is the one we intend to study, not one constructed by philosophers.

The distinction between implicit and explicit concepts drawn by the experimentalists works slightly differently for the case study method. The distinction is between any explicit definitions of a term given in a paper in a case study (the explicit concept), and how the concept is being employed in the rest of the research of which the written paper is a part (the implicit concept). It is primarily implicit concepts I will be interested in drawing out for my project, since situated cognition perspectives are rarely explicitly adopted. Getting at these implicit concepts using the case study method involves careful reading to identify the role the concept plays in the theory and practice underlying the research being described. The implicit concept may be revealed in any experiments or apparatus chosen, or the background theories invoked, as well as specific uses of the concept in the paper itself. The paper is therefore functioning as an integral part of the scientific research – it is not simply a factual record of research produced after the actual work is done – but also as a source of clues to other parts of the research, including those that may not be explicitly documented.

There are some apparent downsides of the case-study method. Three of these are mentioned by Machery and Cohen (2012), and I will discuss ways to alleviate these problems here. While the case study method is not perfect, I argue that it is not as bad as critiques from experimentalists make out.

The first problem is that the sample size is much smaller for a case study approach than for an experimental method and '[a]s a result, this method is not optimally tailored to examine whether different subgroups… endorse different norms, methods, or assumptions.' (Machery and Cohen, 2012: 186).

This is not a particular problem for my project because my question is just whether there are *any* instances of situated cognition concepts, and then whether those instances are legitimate. Just one instance would answer my question of whether there are any, so a sample of statistically significant size is not required.

In addition, the historical approach I advocate in the next section involves looking at how the framework in use in a particular research project evolved from prior research, particularly focussing on the subdiscipline level of grain. I will therefore be looking at a wide range of research over time, often looking at how a subdiscipline arose and developed. There is still something to be said from this work about differences between the subgroups considered, even without surveying a wide range of contemporary work in each subgroup.

The second problem identified by Machery and Cohen is that ‘it might also be problematic to extrapolate from these few alleged paradigmatic articles to a whole field since the research commonly done in a scientific field can substantially differ from the research done in the articles singled out by philosophers’ (Machery and Cohen, 2012: 186). Machery and Cohen are quite right that we should avoid extrapolating to a whole field of research; however, more local results can still have value, even if we cannot generalise them with much certainty beyond the case studies we are looking at. Again, this is not a criticism that particularly affects the kind of project I am engaged in. If one did want to find out more about whole fields and the variation between them, this is a good place for experimental and case study methods to work together. The experimental approach can uncover the possible extent of variation and suggest what kind of case studies might be appropriate, and the case study approach can give detailed and in-depth analysis of contrasting concepts in play.

The third problem mentioned by Machery and Cohen is that philosophers tend to focus on paradigmatic articles and books that defined the relevant field, so more recent changes in the field can be ignored (Machery and Cohen, 2012: 186). This is an excellent point about philosophers in general, but again, for my project, any instance of a situated cognition concept would be relevant, paradigmatic or not. Even for other projects than mine, once we are aware of this problem, we can relatively easily avoid it by selecting a broad range of case studies, both in terms of their age, and how central they are to the field. For MEMORY, the problem of focussing on paradigmatic articles is not too severe in any case given that there is no paradigm as such – that is part of the reason why the conceptual diversity is so great.

The case study method is therefore a better approach than its critics suggest, in particular for a project such as mine that does not require statistically significant sample sizes. In some cases, it can fruitfully be used alongside an experimental approach. However, for my project specifically, the experimental approach is problematic, and the case study method avoids these problems, while retaining the advantages of being able to talk in terms of epistemic niches and implicit and explicit concepts.

To close this section, I will give some brief guidance on how to choose case studies for research on MEMORY. Because I am looking for situated cognition concepts, I want to select cases where it seems *prima facie* most plausible that such concepts would be found. Because I am interested in variation between subdisciplines, my case studies will focus on debates over memory, where it seems that contrasting concepts may be being employed on different sides of the debate. Each case study will feature two contrasting subdisciplines or groups of subdisciplines. I will introduce the specific cases I have chosen in a short section at the end of this chapter.

The case studies should be analysed carefully to determine the type of concept of memory in play in each case. These concepts should then be looked at alongside the epistemic niche in which the researchers were employing the concept in each case. This will involve looking at the wider dynamic framework in which the concept is embedded. In this way, we will be able to get some idea of how the niche applies pressure to scientists' concepts, or in other words why the scientists have the concepts that they do.

The normative project

It may appear that conceptual ecology is a purely descriptive tool, but this is not the case; there is also a normative dimension. I will be advocating a second and broader kind of normative work, but first it is worth looking at that already inherent in the conceptual ecologists' work. What they have to say on this topic refers explicitly to the experimental approach, but applies to conceptual ecology generally, i.e. to both the experimental and case study methods.

Normative force on a concept is provided by the surrounding framework of theories and practices – the concept should be the best tool for the job it does. There is conceptual change and diversification over time as the role the concept is needed to play (the epistemic niche) changes. The concept should change in response to this pressure, but it may not do so. This leaves room for the conceptual ecologist to recommend a way of improving the fit between concept and niche.

As Griffiths and Stotz say '[experimental conceptual ecology] allows philosophers to embrace and study conceptual diversity, and hence to gain new insights into the process of science... It can provide insights for *normative* work in philosophy of science – scientists may be using conceptual tools that are not well suited to the job in hand.' (Griffiths and Stotz, 2008: 518, emphasis in original). In his review of Beurton et al.'s (2000) book *The Concept of the Gene in Development and Evolution: Historical and Epistemological Perspectives*, Griffiths argues that his approach '…can suggest better ways to conceptualize the subject matter and even diagnose a persistent conceptual problem in a scientific tradition…' because scientists' concepts may not always be perfectly adapted to the epistemic niche (Griffiths, 2002: 276).[36]

Conceptual ecology would allow us to analyse the extent of the variation in MEMORY and to explain it by using the notion of the epistemic niche, and I have argued that the case study approach is the best way of doing this. It could also suggest improvements to the concept or resolve confusions by analysing fit between concept and niche. However, I think there is still room for improvement on this method. This introduces the need for the broader kind of normative work I mentioned above.

While the method I have discussed so far allows us to analyse fit with the epistemic niche, it says nothing about whether the niche is appropriate in the first place. It may be that a concept is perfectly adapted to its epistemic niche, but the factors constituting the niche are not those that would produce the best science. Arguably Lysenkoist science furnishes us with an example of this. The name refers to the work of Trofim Lysenko, a Soviet biologist and agronomist whose work is widely decried as pseudoscience, but which was endorsed by the Stalinist regime for political purposes. At one stage, Lysenko's theories were the only theories allowed to be taught.

The epistemic niche for a particular concept used in Lysenko's work (for example INHERITANCE) includes factors such as supporting the political regime, and suppressing alternative work. The concept may be perfectly adapted to this niche (let us say so for the sake of argument) but we still want to say that something has gone

[36] This is a discussion of what Griffiths calls "conceptual archaeology". This seems to be a general term encompassing experimental conceptual ecology and more historical approaches.

wrong. We want to be able to criticize factors in the epistemic niche, not just the fit between concept and niche. This is where the truly normative part of my project comes in.

We need a means of deciding what constitutes a good reason for having a particular concept, where a "good" reason is one that produces fruitful science, as this notion has emerged from the rest of the relevant framework. Because what it means to be successful or fruitful science is something that emerges over time, we need to look at the history of the framework, in particular the development of the epistemic niche for the concept in question. In other words, we should study the evolution of the environment in which current concepts are employed.

I have already discussed an approach in the philosophy of science that allows us to assess a framework in terms of its historical development – Shapere's internal/external distinction – and I propose making use of it again here. We need to take an historical approach to see whether the considerations shaping the concept are of a kind that the science has internalised. Recall from chapter 3 that this will involve seeing whether they have a track record of success that entitles the scientists to rely on them as background information. This approach leaves open the possibility that the science could have internalised different considerations, and so captures the contingency required by both my openness to the possibility of pluralism, and Shapere's "Rejection of Anticipations of Nature" (see chapter 3), while still giving us a basis for making normative claims.

We can now begin to see that the problem with something like the Lysenko case is that it arose in a climate where suppressing opposition and having the results of science be dictated by political regime were external considerations. They should therefore not have been treated as legitimate parts of the framework, being allowed to shape concepts, etc. Showing this in detail here would require too long a diversion; it would need the kind of in-depth analysis I undertake in my case study work in chapters 6, 7, and 8. However, I can outline the shape such an analysis would take. This will serve as a brief and preliminary illustration of my methods.

Taking a concept such as INHERITANCE in the Lysenkoist framework,[37] one would need to analyse the epistemic niche to find out the dominant factors shaping the concept. These would include suppressing opposition, and having the results of science be dictated by the political regime, amongst other things. One would then need to look at the history of science to show the process by which these factors had already been found not to result in successful science. This might include showing the internalisation of factors in direct conflict with these, such as the exposure of scientific theories to opposing views, and attempts to isolate scientific methods and results (although not necessarily aims or choice of projects) from political concerns. The employment of all these factors would need to be measured against the evolving standards for successful science for the relevant frameworks. These could be the frameworks for the disciplines concerned (biology and agronomy), or the coarser-grained framework of science in general. In this case, assessment at either level of grain would show that the relevant factors were external.

Assessing concepts in this way requires the interpretationist strategy I discussed in chapter 4; to assess the factors in the epistemic niche involves getting into the framework in order to judge it on its own terms. One could thereby show that, by the time Lysenkoism came along, the dominant factors in the epistemic niche for INHERITANCE in that framework had already been demarcated as external at a coarser level of grain than that of the research program. We could then conclude that the concept was not legitimate because it was chiefly shaped by these factors.[38]

In subsequent chapters I will carry out this kind of analysis, examining the epistemic niches for particular concepts of memory described by my case studies. As well as needing much more detail than I can give here of the Lysenkoist case, there is a further methodological complication: There are *three* possibilities for any factor in

[37] I phrase my criticism here in terms of identifying a problematic *concept*, but the method could be applied to any other aspect of the framework, in terms of pressure from external factors in the surrounding parts of the framework.

[38] This is necessarily (given constraints of space) a schematic analysis. A deeper analysis may reveal that some of my speculations here are wrong, but I think the form of the argument is right. For some details on Lysenko's work in relation to the scientific and socio-political climate of its time, which gives an idea of the complexity of the epistemic niche one would need to discuss in order to tell the full story, see Roll-Hansen, 1985.

the epistemic niche: it may be ***internal***, an ***untried hypothesis***, or ***problematically external***:

- ***Internal*** factors are those that have been internalised due to a track-record of success.

- ***Untried hypotheses*** are things that have not yet had chance to acquire a track-record of success, but are just being tentatively tried out by the science in question. Because many of the sciences of memory are immature sciences and do not yet have a firm internal/external distinction, we might expect to find many of these factors (see chapter 3). Note that this is not the usual sense of "hypothesis", such as might be used to talk about a "scientific hypothesis". It refers to *any factor* in the epistemic niche which has not yet been sufficiently tested to be either internal or external (see chapter 3).

  Recall the case of the Cognitive Atlas project discussed in chapter 2. The big data approach in use there has not been internalised into cognitive neuroscience, and in fact is being employed partly because of a similar method in biology (in the Gene Ontology project) so may look like it is being imposed from biology as an external factor. However, it is not being systematically and unquestioningly used, it is merely being tested to see whether it leads to fruitful scientific research. If it does so, it will be internalised; if not, scientists should stop using it. It is therefore an hypothesis in the sense in which I am using the term here.

  In fact, it is sensible to employ techniques that have been successful in neighbouring disciplines, provided that they are treated as hypotheses, not relied upon as though they had been internalised.

- It is only ***problematically external*** factors that it is irrational to rely on. These are considerations that are being treated by the science as though they are internal – they are being relied upon and taken for granted as background information – despite the fact that they have been tried out without acquiring a track-record of success. In chapter 3, I argued that unification is one such factor in the cognitive and social sciences. Suppression of conflicting research in the Lysenkoist case briefly discussed above would be another example.

A third example is journal publications. Just because a particular approach is more likely to get published than another, this does not make it a better approach for producing fruitful science. Brigandt and Love talk about this factor and its leading to a focus on the concept NOVELTY in evolutionary and developmental biology: '…workshop participants drew attention to the possibility that the invocation of evolutionary novelty serves more as a rhetorical device in the process of grant writing than as an important biological concept.' This is because such invocation tends to result in publication in high impact journals (Brigandt and Love, 2010: 7). This factor is external, not because it is a "sociological" factor, but because its employment has not resulted in a track-record of producing successful science.

I can now spell out my search for legitimate concepts – those functioning as investigative kinds – in these terms: To the extent that a concept is shaped by internal factors, I will consider it to be a legitimate concept of memory, and to the extent that it is shaped by hypotheses and external factors, I will consider it to be non-legitimate. I have chosen this term rather than "illegitimate" to reflect the fact that untried hypotheses and external factors have not *been legitimated*, but that doesn't mean they never could be.[39]

To the extent that a concept is shaped by internal factors and hypotheses, it is rational, and to the extent that it is shaped by external factors, it is irrational. Hopefully a schematic representation will make this clearer:

[39] This is true even of external factors. Although such factors lack a history of success, we cannot demonstrate that they will always lack such a history in all circumstances. To do so would be to prove a negative claim about an infinite domain. The problem is in scientists' reliance on these factors as though they were internal, despite their lacking the right history of success.

Internal	Hypothesis	External
Legitimate	Non-legitimate	Non-legitimate
Rational	Rational	Irrational

Any epistemic niche is likely to contain a mixture of the three types of factor, so a concept is not simply legitimate or non-legitimate, but is legitimate to some degree – the degree to which its niche is made up of internal factors. It is also rational to some degree – the degree to which its niche is made up of internal factors and/or untried hypotheses.

Note that this is a rather special sense of "legitimate". A concept that is held primarily as a result of untried hypotheses is not irrational to hold (what else could the scientists do in the absence of more internal factors?) but it does not count as legitimate in the sense I am interested in here. This special notion is important for distinguishing concepts that are held as a result of accepted background information that has a track-record of success. These legitimate concepts pick out investigative kinds. To see this recall that investigative kinds are those that support fruitful scientific theory and practice, i.e. which have the kind of successful track-record that is necessary for internalisation. Concepts without such a track-record have not been legitimated in this way.

My work will hopefully allow a conclusion to be drawn as to whether any situated cognition concepts of memory that I find are legitimate. On my view, finding situated cognition concepts in this category would be the nearest thing we could have to an answer saying that the situated cognition perspective is correct. Finding situated cognition concepts that are rational but non-legitimate would mean that we cannot yet say whether the perspective is correct, but that there is nothing wrong with the scientific practice employing it. This seems to be the right thing to say, given that such concepts would largely be the result of untried hypotheses. Finding situated cognition concepts that are shaped primarily by external factors would indicate that something has gone wrong with the scientific practice, and that these are not the concepts that those sciences should be operating with.

I can now complete the task I deferred in the last chapter, of giving a way to account for conceptual change as rational.[40] A change is rational to the extent that it comes about as a result of pressures from internal factors and/or untried hypotheses, and irrational to the extent that it comes about as a result of pressures from external factors. Although I am looking at case studies to determine whether their concepts are rational (or legitimate) at a particular time, it is important to remember that the framework is dynamic and, according to the historical approach, it is a chain of good reasons for conceptual change over time that leads to a particular concept being rational at a time. I have therefore been able to give a way to account for the rationality of conceptual change without reference to an extra component of the concept outside of its conceptual role (such as Brigandt's epistemic goal), but instead have used the conceptual role's own history.

Before closing this section, I want to deal with two kinds of case that might be used to object to the method I have outlined here: coherent Lysenkoist cases and Swampman cases. I will treat each in turn, arguing that raising these kinds of cases misunderstands the approach I am taking here.

Coherent Lysenkoist cases

What I am referring to as "coherent Lysenkoist cases" are frameworks that are completely internally coherent, i.e. successful according to their own standards, but which have aspects that we would not want to endorse for social or political reasons. These are like the Lysenko case introduced above, except that we assume that suppressing opposition and being subservient to political aims have allowed the science to be successful on its own terms, and have therefore been internalised. This perhaps does not seem implausible, since the same political factors that resulted in this kind of science coming into being also shape what counts as success for that science. The practice of suppressing alternatives then increases the success of a theory on its own terms, so it might seem that it should be internalised according to what I have said here.

[40] The challenge was to do this without separating conceptual role and epistemic goal as different components of conceptual content as Brigandt does (see chapter 4).

This case gets at worries that relativism is inherent in the idea of judging a framework on its own terms; but such worries misunderstand what it means to judge a framework in this way. Frameworks are not isolated, incommensurable systems. The historical approach I am advocating involves looking at how a framework developed and how its criteria came to be internalised. There is a pre-existing context for this process; the new framework does not emerge in a vacuum. We should not let the possibility of treating the framework for a research project as a relatively discrete entity for some methodological purposes blind us to its connections to the frameworks of other research projects. It is part of the same framework as them when viewed at a coarser level of grain, and some criteria have emerged as external considerations at this coarser-grained level. When we consider the Lysenkoist case, suppressing conflicting research is one such criterion. Even at the very coarse level of grain that is epistemic enquiry in general, this is an external consideration. It would therefore never even be a candidate for internalisation into the Lysenkoist research programme.

This is the picture I drew on in the last chapter, when I noted that interpreting a scientific framework is never truly *radical* interpretation in the sense that Davidson and Quine discuss. If frameworks were entirely isolated and did not draw on other frameworks at their inception, each new research project would start with a "blank canvas" where anything might or might not turn out to be internal. Research under these circumstances would be impossibly slow.

The problem with something like the Lysenkoist approach then, is that it arose in a climate where factors like suppressing opposition were already external considerations. They may never have been used as criteria by this particular framework when viewed at the research project level of grain, and so may look like untried hypotheses (and thus at least candidates for internalisation) from this point of view. However, they have been demarcated as external considerations across the subdisciplines (at a coarser level of grain), and therefore in similar research projects. According to what I am saying here, it would be irrational to rely on such criteria. This is why the distinction between untried hypotheses and problematically external considerations is important. In the coherent Lysenko case, we are dealing with the latter.

It may be that even the coarse-grained framework of human epistemic enquiry could have developed otherwise. In other words, it may be that nothing about the world would have prohibited the development of a kind of enquiry according to which coherent Lysenko cases could be rational and successful sciences. We have no way of knowing whether this is the case without re-running most of human history, to see whether the world would ever allow such frameworks to develop. If they did, they would look so radically different from anything we can envisage now, including having very different meanings for the terms “rational” and “successful”, that it is hard to know what purpose understanding them could serve for us here and now.

Swampman cases

Swampman cases are scientific frameworks that spring into life fully formed, but are nonetheless completely correct and successful. Like Davidson’s Swampman in the philosophy of mind (Davidson, 1987) after which they are named, they lack the relevant history, but have the appropriate present-day characteristics enabling them to go on appropriately in the future.

As might already be clear from my response to the coherent Lysenkoist case, from the perspective I have taken so far in this thesis, there is something wrong with trying to set up such cases. Terms like “correct” and “successful” get their meaning in part from what emerges from the framework over time, so it does not make sense to say that a framework that has no history is correct or successful. It does not yet have internal criteria for what it is to be successful, so there is no way to judge it on its own terms. We cannot say exactly how it would have to go on in order to count as successful, except in very coarse-grained terms, for example that it should allow good explanations, predictions and practical applications. The only reason we can say even these things is by assuming that the Swampman research project’s framework is part of coarser-grained frameworks which do have a history, for example at the levels of science, or epistemic enquiry in general. It is only through investigation at these levels that we have learned that anything we could call “successful enquiry” was possible at all; it might have been the case that the world was too chaotic to allow this.

To see the difficulty, try setting up a Swampman case in more detail. Perhaps it seems that a research project to find a drug to cure cancer which works with the first randomly chosen substance it tries could serve as an example. But this cannot work. What cancer is, what would count as curing it, and what a drug treatment is, are all things that have long histories, and the research project could only be considered successful in the light of these things.

Perhaps we can try a more outlandish case, where not only the project, but all the scientists, and the society of which they are a part are Swamp-creatures. Imagine a planet that comes into existence with a race of beings, some of whom (the Swamp-scientists) give pills to others (the Swamp-patients). The Swamp-patients survive, but without the pills, they would die. The Swamp-scientists can give what look to us exactly like explanations of how this happens, and make predictions about it, as soon as they emerge from the swamp.

Here we genuinely have a case where no part of the framework has a history, but this is not a successful science by its own standards, because it does not have any standards for success; all we can say is that it is successful according to our standards. Because our standards are not those of the research project in question, they do not count as standards for success in the sense I am interested in. (If they did, our history would be the Swamp-science's history, and once again we would not really have a framework without a history).

The Swamp-scientists may stipulate standards that they call "standards for successful science", but according to the perspective I am taking here, they cannot really be such standards. A notion of success is something that can only emerge over time, through prolonged contact with the world.

Although a science with no history could not be successful, it would not necessarily therefore be irrational, because the considerations it employed would fall in the untried hypotheses category, rather than the problematically external one. As I said in chapter 3, it is possible for a science to mature very quickly, so a long history is not necessarily required. However, nothing can demonstrate a history of success when it has no history at all, so the framework of our Swampman science has no internal factors.

To both the Swampman and coherent Lysenko problems, the answer is that the cases cannot be set up. According to my account, there is no such thing as a successful Swampman science, and no such thing as a coherent Lysenkoist science.

To summarize what I have said about the normative project: in order to find out whether any particular factor in the epistemic niche for a concept is internal, hypothesis, or external, its history must be traced to see whether it has established a track-record of success. Therefore, when good case studies have been identified and the epistemic niches of their contrasting concepts investigated, the history of those niches should be traced. The factors that make up the niches can be assessed in terms of how they came to be seen as important by the scientists in question, and therefore whether they are internal factors, untried hypotheses, or problematically external.

If there are internal considerations shaping multiple different concepts of memory, I take this as an indication that memory science should retain its current plurality, because multiple concepts are legitimate. If there are multiple different rational concepts of memory, even if these are not all legitimate (because they are held as a result of untried hypotheses), memory science should still retain its plurality for the time being. We cannot predict that it should continue to do so when the hypotheses are given fair trial, but nor can we predict that it should not. If there is only one legitimate concept of memory, the science should be monist and adopt that concept. In this way, it is possible to make a local investigation into whether plurality in the sciences of memory is a good thing (local to whether MEMORY should admit of situated cognition concepts), although we cannot draw conclusions about whether any of the particular scientific subdisciplines within memory science should internalise unity or pluralism. This is a way around the immaturity of the sciences of memory discussed in chapter 3; although a particular subdiscipline may not have sufficient internal criteria to answer the question of whether it should be pluralist or unified overall, we can make some headway with local questions about particular concepts.

One might ask just how revisionary the upshot of such work could be. To what extent can the work I undertake here recommend changes to scientific concepts? Some philosophers taking a similarly pragmatic approach to concepts recommend conceptual engineering. For example, in her work on KNOWLEDGE Sally Haslanger, somewhat like I have done here, suggests that it is our purposes in using the concept that are important, and that we need to ask whether those purposes are legitimate (Haslanger, 1999: 468). Haslanger says that the best way to go about her project 'is first to consider what the point is in having a concept of knowledge: what work does it, or (better) could it, do for us? And second, to consider what concept would best accomplish this work' (Haslanger, 1999: 467). The role the concept plays is primary here. For Haslanger, it is this role that we should look for first, then think about the concept that could fill that role. To put her approach in SGK's terms, we should design the epistemic niche first, then find a concept that adequately fits that niche.

My project is similar to Haslanger's in spirit, although I am looking at current concepts and the roles they play, before seeing whether they are the roles they should play. Haslanger, on the other hand, recommends beginning with working out what role a concept should play, then seeing what could fill that role. In Fisher's (2006) terms, I am looking at "tried-and-true" ways that past employment of a concept has regularly delivered benefits, while Haslanger advocates looking for "novel" ways of delivering benefits by engineering our concepts for future use (Fisher, 2006: 96–97). But could I go further and recommend conceptual engineering?

In a sense I am recommending some degree of engineering – we should let go of problematically external criteria, and let concepts be shaped only by internal criteria and, where these fall short, untried hypotheses. This allows some scope to alter science, for example if its goals are judged morally suspect. However, we cannot change concepts in any way we like. Consider the Lysenko case again: our criticisms of the epistemic niche must arise from practice over time. Even the very idea that science should be responsive to moral concerns is something that we have learned and internalised. To go beyond this in our engineering would be to problematically anticipate nature in the way that Shapere schools us against. We cannot engineer concepts according to external standards.

With this in mind, my aim is not to radically change the practices of the memory sciences according to philosophical goals, but only to look for ways that the unfolding of the disciplines might run smoother. The point is not to impose a situated cognition perspective, but to see whether or not the time is ripe for one.

I will now go on to say a brief word about how to choose case studies for the approach I have outlined here.

Choosing case studies

I have said that I will be looking at cases that seem *prima facie* most likely to involve situated cognition concepts. I have also said that it is useful to choose cases where we would expect to find contrasting concepts of memory in order to get the most out of this method. What choices of case studies can meet these criteria?

The first case study I will look at concerns memory in locked-in syndrome patients. I will be contrasting neuroscience and neurology with philosophy. *Prima facie*, we would expect neuroscientists and neurologists to have a brainbound concept of memory. The philosophers I consider in this case study are directly discussing, and in some cases endorsing, various situated cognition perspectives. If their implicit concepts match their explicit ones, we might therefore expect them to have situated cognition concepts.

My second case study concerns analyses of memories of two political scandals. Here I will be contrasting cognitive psychology with discursive psychology. Unlike in the first case study, both sides here are good places to look for situated cognition concepts of memory. The cognitive psychologist whose work I will be looking at is Ulric Neisser. Neisser argued that cognition in the laboratory is often very different to real world cognition and that laboratory-based experiments therefore lacked *ecological validity*. This is an important idea in embedded, extended and enacted cognition approaches.

The link between discursive psychology and situated cognition is more obvious. Discursive psychologists see something like a memory as being constructed through discourse. For them, discourse is therefore more than merely an input or output to cognitive processes, it is part of such processes. This parallels Clark's arguments for extended cognition, where external processes are seen as part of cognitive processes, not merely as inputs and outputs to them.

Despite both Neisser and the discursive psychologists having possible situated cognition readings in common, there is still a contrast in concepts here. *Prima facie*, we would expect cognitive psychologists to have a very individualist concept of memory, while discursive psychology sees memory as constructed between participants in a discourse, and thus involving multiple people, perhaps in the manner of distributed cognition.

My third case study concerns Transactive Memory Systems – a kind of system for which memory is a property of a group. I will be contrasting cognitive and social psychologists with researchers in communication studies. This case is a good place to look for situated cognition concepts of kinds where memory is extended or distributed over multiple people, rather than in physical tools. There is an expected contrast between the two disciplines because psychologists are interested in individual cognition and how it gives rise to group phenomena, while those in communication studies are interested in how the communication between group members meets the aims of the group. This difference is expected to lead to a different way of construing group memory.

In subsequent chapters, I will apply the methods discussed here to these case studies.

Conclusion

The method for my project can be divided into descriptive and normative components. The descriptive part will involve a case study-based conceptual ecology, to determine what concepts are in play in various examples of memory science, and what the epistemic niches for those concepts are. The normative part will involve finding out whether those concepts are legitimate – whether they are functioning as investigative kinds – by tracing the history of the factors in their epistemic niches.

To the extent that a niche is made up of internal factors, the concept it shapes is legitimate, and to the extent that it is made up of internal factors and untried hypotheses, the concept it shapes is rational.

6. Locked-In Syndrome and Brain-Computer Interfaces: A Case Study

Introduction

This case study compares neuropsychologists and neurologists to philosophers. It focuses directly on the various situated cognition perspectives, because the philosophical papers on one side of the comparison discuss these perspectives explicitly. Specifically, they discuss whether brain-computer interfaces (BCIs) being used by Locked-In Syndrome (LIS) patients should count as instances of situated cognition, in particular extended or enacted cognition. The neuropsychologists and neurologists do not discuss this issue explicitly, but they provide a different perspective to the philosophers on the use of BCIs in LIS. We might *prima facie* expect them to have a brainbound concept of memory, which should provide a contrast to the philosophers.

I will first explain the details of the case study, then try to tease out the implicit concepts of memory in play. Note that the contrast I have highlighted above is only a *prima facie* one which makes this case study a promising place to start; whether the implicit concepts match this initial suspicion still remains to be seen. This is the case even with the philosophical papers, where we might expect the authors to be clearer about the concepts they are using. It is important to remember that the concepts being used are not necessarily the concepts that the authors think they are using and explicitly endorse.

When the implicit concepts of memory have been identified, I will use conceptual ecology, in particular the idea of the epistemic niche, to explain why the different subgroups have the concepts that they do, and attempt to determine which features of the relevant epistemic niches are internal considerations, which are untried hypotheses, and which are external, by looking at the niches in their historical context. This will allow me to make a claim about which of the concepts in play are legitimate, i.e. which are functioning as investigative kinds. Obviously I will not be able to consider all the aspects of theory and practice that make up the epistemic niche here, but I will try to identify those that are most important in determining whether the authors use a situated cognition concept, and if so, which one.

The case study

Locked-in syndrome (LIS) is a rare condition caused by brain damage to the ventral pons, usually caused by a stroke or degenerative disease like motor neuron disease, but also occasionally by trauma. Patients are paralysed and unable to speak, but consciousness is preserved. LIS is usually classified into three categories:

1. In *classic LIS*, patients retain vertical eye movement and/or eyelid movement.
2. In *incomplete LIS*, some other voluntary movement is preserved.
3. In *total LIS*, there is no ability to move or communicate at all.

In all cases consciousness is preserved and there appears to be a high level of cognitive functioning, although testing cognitive functioning will be the topic of the papers discussed below. (For a good brief overview of the condition, see Smith and Delargy, 2005).

A brain-computer interface (BCI) is a device that allows brain activity normally resulting in movement of the body to be read and translated into other outcomes such as the movement of a robot arm or a cursor on a screen. This exploits brain plasticity, as the BCI

> attempts to assign to cortical neurons the role normally performed by spinal motoneurons. Thus, a BCI requires that the many CNS areas involved in producing normal motor actions change their roles so as to optimize the control of cortical neurons rather than spinal motoneurons. The disconcerting variability of BCI performance may stem in large part from the challenge presented by the need for this unnatural adaptation (Wolpaw, 2007, quoted in Fenton and Alpert, 2008: 122).

BCIs can be used for various applications including movement of a robot arm, and control of an avatar. However, the most developed application and that which has received the most attention is communication. Research has focussed here because patients report that the inability to communicate is the most traumatic and frustrating aspect of LIS (Fenton and Alpert, 2008: 127) and because of the obvious clinical benefits of being able to communicate with a patient.

As the above quote suggests, outcomes are variable, but one of the most successful methods for communication is to use a virtual keyboard. Electrodes, usually placed on the scalp, are used to detect electrophysiological and chemical activity of the brain. The user imagines bodily movements in order to move the cursor, for example imagining moving the right hand to move to the right and the left hand to move to the left. Letters are selected by imagining another movement, for example squeezing the hand. In this way, words can be spelled out and communication is possible (Heersmink, 2013: 209).

Older communication methods exploit the preserved eye movement in classic LIS. In the simplest versions, an interlocutor holds an alphabet board with the letters of the alphabet, often in order from most commonly used to least. The interlocutor reads out the letters one by one and the patient blinks when the desired letter is reached. A version using a similar set-up but not requiring an interlocutor has also been developed. In this version, the patient focusses on the desired letter and a sensor detects the gaze (León-Carrión et al., 2002, review: 565). Unlike the BCI, neither of these older methods offers hope for the patient with total LIS.

These methods of communication look at first glance like good examples of extended cognition since the vehicles of cognition seem to be partly in the world outside the head. Their relevance to the concept of memory will be made clear below.

For this chapter, I focus on a series of papers that appeared in the journal *Neuroethics* from 2008–2013, and three papers by a combination of neurologists and neuropsychologists that focus on similar issues. Two of these three papers are cited by the debate in *Neuroethics*, and the third (Allain et al.) is cited by both of the other two. The papers from *Neuroethics* are all written by philosophers, or authors with a philosophical background, and they constitute a debate over the status of BCIs for LIS patients.

The first paper in this debate (Fenton and Alpert, hereafter F&A, 2008) claims that BCIs can be seen as examples of extended cognition in action, and that they extended the selves of LIS patients. The next paper (Walter, 2010) disputes this, claiming that F&A’s claim about extended selves is unfounded because the extended cognition framework is not typically applied to selves, but to classically cognitive processes (of which memory is one example). When it comes to the impact of BCIs on such cognitive processes, the capacities enhanced by BCIs are all bodily, not cognitive, so BCIs are not an example of extended cognition. He suggests that enacted cognition is a better framework for understanding the value of BCIs. Kyselo (2013) then responds, agreeing that the enactive approach is the best way of thinking about BCIs and their effect on the self, but arguing that BCIs can be seen as vehicles of certain extended cognitive processes. Finally, Heersmink (2013) claims that BCIs are not yet able to extend cognitive processes, but with technological improvements, they could do so.

In contrast to these papers, I consider the following:

Allain et al. (1998) from the journal *Cortex*: This is a paper about the testing of cognitive functioning in LIS patients, with a specific focus on testing memory.

Schnakers et al. (2008) from the *Journal of Neurology*: This paper is concerned with neuropsychological testing of cognitive functions (including memory) in LIS patients.

León-Carrión et al. (2002, survey) from the journal *Brain Injury*: This paper is a questionnaire survey of patients with LIS, collecting data (including some on cognitive functions such as memory) with the hope of improving diagnosis and rehabilitation.

There is a contrast here between papers which test for memory to answer the question of whether cognition is intact in LIS patients, and papers for which the nature of cognition and therefore of memory is under question. Despite focussing primarily at the subdiscipline level of grain, I am not assuming that the philosophers share one concept of memory and the neuropsychologists and neurologists share another. There could be variation within groups, but for my purposes here, the most important contrast is that posited between the two groups. In the next section, I will attempt to describe this variation for MEMORY in more detail.

Concepts of memory

Neuropsychologists and neurologists

Allain et al.

Allain et al. is the oldest of the papers I consider here, and the technology for communicating with LIS patients has developed since it was written, but it is helpful for its explicit focus on memory. The authors tested two patients with incomplete LIS for both verbal and visual memory. Conventional psychometric tests were used, adapted for the unusual means of communication required. The tests administered were as follows:

> Verbal immediate memory was assessed using the digit forward and backward tasks of the Wechsler Memory Scale Revised (WMS-R; Wechsler, 1991). Verbal learning was investigated using the Rey's 15 words test (Rey, 1958) which requires free recall of a list of 15 common words read out by the examiner. The task was repeated five times and the sum of the recalled words was the score for immediate memory. Verbal and visual paired associate learning were assessed with the subtests of the WMS-R (Wechsler, 1991). The score was the sum of the recalled pairs after the 3 first trials. Delayed verbal and visual recognition were tested using a subtest of the BEM 144 (Signoret, 1991), in which the patient is asked to remember 24 sentences and 24 figures presented one by one. After a delay, each stimulus is presented with 3 alternatives and the patient has to recognize the target. (Allain et al., 1998: 632).

Verbal abilities and general intelligence were also tested.

The results showed both patients to be in the normal range for all memory tests, although they were at the bottom end of this range for the backward digit task, which the authors attribute to 'the cumbersome procedure required by the communication aid system' (Allain et al., 1998: 633). They conclude that there is no cognitive impairment of language, memory or intellectual functioning.

The concept of memory here is revealed by the methods chosen to test it. The recall tasks are adapted from those used on able-bodied subjects, but when an able bodied subject tackles these tasks they are in a very specific setting. They are likely to be in a laboratory or perhaps a doctor's surgery, probably alone apart from the experimenter, and if someone else is present they are likely to be acting only as an observer. The subject does not use any external tools to aid memory, because for memory as it is being tested here, that would be cheating. The prospects for any kind of situated cognition are therefore none to very slight. The memory being tested here is a capacity of the brain. This should come as no surprise, since Allain et al.'s paper appears in the journal *Cortex*, and most of the authors are neuropsychologists (the remainder working in medicine and rehabilitation of brain injury patients).

The use of the BCI here is for communication alone. Any part that communication may play in remembering for an able-bodied subject in an everyday context is "screened-off" by the testing situation for both LIS patient and able-bodied subject. There is no prospect of talking to oneself or another person, or writing or sketching to aid recall. Potential external vehicles for cognition are placed out of reach because the type of memory that has external vehicles is just *not memory* in this context. Imagine if Clark and Chalmers' notebook-wielding amnesiac Otto (Clark and Chalmers, 1998) were to be tested with this battery. Otto writes things down in his notebook, and later consults his notes to "remember". In the context of these tests, that is cheating and Otto has not remembered. It is according to tests like these that he is described as an amnesiac in the first place. Any use of a BCI as a vehicle of memory rather than purely as a communication tool would similarly be deemed cheating. Allain et al. are employing a brainbound concept of memory.

Schnakers et al.

Schnakers et al. similarly set out to test cognitive abilities in LIS patients. They improved on the methods of Allain et al. because they used a simpler method of communication, relying on preserved eye movement. As the authors say:

> Our results also corroborate two case reports [one of which is Allain et al.] having investigated short- and long-term memory, language and general intelligence in LIS. However, the latter responded to the test using a sophisticated, cognitively demanding computerized communication aid. It is therefore not surprising that these patients showed no cognitive impairment as they were already selected based on their ability to comply with the cognitive requirements of the implemented communication device. (Schnakers et al., 2008: 328).

Schnakers et al. tested long- and short-term memory, attention, executive functioning, phonological and semantic processing and verbal intelligence. Short-term memory was tested using forward and backward digit spans, similarly to Allain et al. Long-term memory was tested using the Doors test from the "Doors and People" battery. This test measures non-verbal episodic memory.

> The Doors test consists in learning two lists of twelve photographs depicting different doors (lists A and B). The learning phase for each list is followed by a recognition phase where 12 sheets are presented, containing one previously presented picture and three distractor pictures. The distractor pictures for the second list are more difficult to reject as they are more closely matched to the target picture (Schnakers et al., 2008: 324–325).

The results showed some memory impairment in some patients, but in each case the experimenters were able to find something to attribute it to other than the lesion of the ventral pons. Such factors included: other lesions in areas of the brain known to be involved in memory, fatigability of subjects, other brain damage resulting in some uncontrolled motor activity which affected the communication method (Schnakers et al., 2008: 327). The authors conclude that it appears that patients with pure brainstem lesions can recover intact cognitive levels, including long- and short-term memory.

As for Allain et al., the concept of memory in play is revealed by the methods chosen to test it. Again, a standard testing battery is used (with some necessary adaptations), such as would be used to assess the memory of able-bodied subjects in isolation from cognitive props that may support cognition. The kind of memory that is being tested here is an ability of the brain.

This is brought out even more strongly in Schnakers et al. by their use of healthy controls who were asked to pretend they were in LIS and communicate using only eye movements (Schnakers et al., 2008: 324).[41] The controls are not people in a normal everyday context with access to cognitive scaffolding in the form of physical tools and other people. The extended/embedded etc. aspects of memory that a BCI might be expected to replace are not present in the healthy controls either; they are not part of memory as it is being tested here. Again this should perhaps come as no

[41] For some tests, published data were used instead of these controls.

surprise as the paper was published in the *Journal of Neurology* and authored by a team of neuropsychologists and neurologists.

León-Carrión et al.

León-Carrión et al.'s paper appears in the journal *Brain Injury*, and the authors are neuropsychologists and those with a specific interest in the rehabilitation of brain injury patients.

This paper is different to the previous two studies in that the testing procedure involved asking LIS patients to assess their own condition:

> A 36-item survey questionnaire…was specifically designed to be filled-out by the member of the family most directly living with the person affected by LIS. In most cases this person was the individual's spouse, and the procedure was carried out in collaboration with the direct participation of the patient with LIS. (León-Carrión et al , 2002, survey: 572).

The questionnaire asked about personal details (marital status, age, gender etc.), the cause of the LIS, treatment received, and about various features of the psychological and emotional state of the patient. Question 13 asked 'Does the patient have memory problems?' with simple yes/no tick boxes for the answer (León-Carrión et al., 2002, survey: appendix, 581). Only 18.6% of the 44 people surveyed answered "yes" (León-Carrión et al., 2002, survey: 571, 577).[42]

The interesting thing about this paper in contrast to the other two is that it was left up to the patients to decide whether they felt that their memories were impaired, and therefore it was the patients' concepts of memory, not the experimenters', that were used in collecting results. Because of the method used to answer the questions, the main caregivers' concepts of memory and opinions of the patients' memory skills were also important. It may be that in some cases, a joint decision was reached between the patient and the caregiver over how to answer the question. Answers may

[42] Interestingly, '[w]hen one analysed for memory problems, it was found that, significantly ($p < 0.002$), those subjects with LIS consequential to TBI [traumatic brain injury] were more likely to have memory deficits than those subjects with LIS caused by stroke.' (León-Carrión et al., 2002, survey: 577).

therefore be a combination of whether the patients felt themselves to be memory-impaired, and whether they appeared to be so to the main caregivers.

The authors mention Allain et al. as a study in agreement with their own results on memory (León-Carrión et al., 2002, survey: 577–578). This suggests agreement with the brainbound concept of memory found in that paper. Further evidence that León-Carrión et al. have a brainbound concept of memory can be found in a literature review article in the same journal issue as the survey I am looking at here. In this review article, they briefly discuss a study by Onofri et al. (1997), concerning measurement of cognitive event-related potentials. This is a measurement of brain activity, of which they say ‘it is possible to record ERPs [event-related potentials] in patients shortly following an acute ischemic lesion, and thereby to *objectively assess cognitive activities*.’ (León-Carrión et al., 2002, review: 558, my emphasis). Endorsement of this method as a measurement of cognitive activity strongly suggests brainbound concepts of cognitive processes like memory.

However, the survey method used in the main article tells a more interesting story. The focus here is much more on memory in an everyday context; 63.6% of the patients tested lived at home. Because no special test of memory is performed, the answers to the questionnaire presumably reflect whether memory-impairment poses any problems in everyday life. If it does not, memory is likely to be rated as unimpaired. Any tools the patient has access to that are used to scaffold memory would probably therefore be included. Memory here is not necessarily an isolated brain state or process, and external vehicles could be involved to the extent that the patient has access to them. There seems to be at least openness to an embedded cognition perspective here, where the cognition still happens in the brain, but is partially dependent on the external world.

As further evidence that it is memory in an everyday context that is being tested, question 23.10 in the survey asks the caregiver whether the patient ‘[r]eminds you of important things that are pending’ (again with yes/no tick boxes for answers) (León-Carrión et al., 2002, survey: 582). The results for this question are not given, but the nature of the question reflects the role of memory in everyday family or social life, not the isolated experimental context seen in the previous two papers discussed. The choice of question suggests that, not only were patients and caregivers answering the question with this everyday practical concept of memory in mind, but the experimenters were asking the questions using a similar concept. The concept of memory in play here therefore seems to be at least open to an embedded cognition perspective, although memory is still seen as a state or process in the brain.

All three of the papers considered so far are interested in the question of whether cognitive states or processes like memory are impaired in LIS patients. The nature of these states or processes is not in question. I will now go on to consider some more philosophical papers, in which the nature of cognition, and therefore of memory, is the subject of the debate.

Philosophers

Fenton and Alpert (F&A)

F&A (2008) discuss the idea that BCIs can be seen as examples of extended cognition in action, and that they extend the selves of LIS patients.

The following claims can be extracted from their paper:

- BCIs promise to enable individuals with conditions like LIS to re-engage with their physical and social worlds (F&A, 2008: 119).
- Functionally integrated BCIs extend the minds of individuals with LIS beyond their bodies (F&A, 2008: 119). They extend their cognitive as well as physical capacities (F&A, 2008: 127).
- From a perspective rooted in embodied theories of cognition, BCIs have the potential to change the individual users themselves (F&A, 2008: 124).

These three claims are linked for the authors because they are writing about a view of cognition according to which ‘various physiological processes are implicated in human cognition. Additionally, the social contexts in which an individual matures can also profoundly affect how she cognitively develops and responds to her environment’ (F&A, 2008: 125). Re-engagement with their physical and social worlds therefore extends the cognitive abilities of the patients; if BCIs can facilitate the re-engagement, they can also extend cognition.

F&A spend much of the paper on the topic of the embodiment or extension of the self, but my focus here is primarily on cognition, not self-hood.[43] F&A say that they use the extended cognition framework of Clark and Chalmers (1998) merely as a *heuristic*, or 'a lens through which we learn to re-see particular aspects of human cognitive engagement with the relevant physical or social environment' (F&A, 2008: 126). In defence of this, they say '[i]nstead of becoming embroiled in metaphysical debates about the nature, or extention [sic], of mind that might threaten to undo any possible philosophical advance arising from a re-seeing of mind as extended, we use extended mind theory heuristically.' (F&A, 2008: 126, footnote 10).

Note that this way of approaching the debate is very similar to that seen in much of the extended cognition literature. The benefits that can be drawn from using an extended cognition approach are all that some proponents argue for (a form of inference to the best explanation, see Clark and Chalmers, 1998: 14; Clark, 2008: 80). Clark sometimes talks about using the extended cognition framework as just one perspective among many from which cognition can usefully be viewed for different purposes (Clark, 2008: 139).[44] It is therefore not clear that F&A would need to do any more than they do in order to count as endorsing the extended cognition approach.[45] However, the attitude they display in their claim to only use the perspective heuristically tells a more interesting story than simple implicit assumption of an extended cognition concept. To understand this story, it is useful to look at what they say about memory.

There are two mentions of memory in F&A's paper. In the first, the authors quote León-Carrión et al.'s study described above, stating that 18.6% of LIS patients report memory problems. Here they also quote a study (New and Thomas, 2005) suggesting cognitive deficits in LIS patients. The other discussion of memory in the

[43] Extension or embodiment of the self may entail extension or embodiment of cognition for any cognitive processes involved in selfhood, see Walter (2010: 64, footnote 6).
[44] For more on this, see my discussion of the epistemic strand of the extended cognition debate in chapter 1, subsection: "From natural kinds to investigative kinds".
[45] Walter (2010) doubts it is possible to use the extended mind thesis heuristically, because it is inherently a metaphysical claim about cognition. According to my reading of the extended cognition literature, only one strand of the debate is metaphysical. Another important strand, the one I focus on here, and which includes some of Clark's own contributions, is not metaphysical but epistemic. See chapter 1, subsection: "From natural kinds to investigative kinds".

paper is part of an explanation of extended cognition. They discuss examples due to Clark and Levy (Clark and Chalmers, 1998; Levy, 2007) which claim that a notebook or personal digital assistant can be part of an extended cognitive mechanism (F&A, 2008: 125). They discuss the example of using a paper and pen to aid in a lengthy calculation, where 'as a repository of our thought, the used paper seems to be relevantly similar to the neural repositories of other thoughts (i.e., memories) constitutive of the relevant calculation. The paper, that is, is an analogue to the neural sites implicated in explicit memory' (F&A, 2008: 126).

The authors discussed in these two passages (Clark and Levy, and León-Carrión et al.) do not have the same concepts of memory. Clark and Levy endorse an extended cognition approach, whereas León-Carrión et al. do not discuss any of the situated cognition perspectives directly. I have argued that they implicitly use an embedded cognition approach, where cognition is still something that happens in brains, albeit with the possibility of external scaffolding. It is therefore open to question whether the extended cognition framework can be used to account for the memory deficit experienced by 18.6% of patients in León-Carrión et al.'s survey. If cognition is extended, a cognitive process such as memory could be expected to be severely impaired by LIS. The percentage of patients reporting memory impairment could be expected to be much higher than 18.6%.[46] What does the mention of these different perspectives tell us about F&A's concept of memory, bearing in mind that they claimed not to endorse extended cognition, but only use it heuristically?

It may suggest that they are just confused, in that they haven't noticed the potential problem with taking data from a study that construes memory differently. Alternatively, their use of this data may just reinforce the idea that they see extended cognition as only one possible option, rather than assuming its truth across the board. Although these two options say different things about their explicit thinking, they both suggest that their implicit concept of a cognitive state or process like memory is flexible. A flexible concept of memory is one that is open to interpretation from different perspectives. This diagnosis fits with the overall aim of their paper, since it

[46] Note that 65.8% in this survey could communicate without technical aid, so it is possible that the degree of impairment was not severe enough in the patients surveyed to cause more memory problems.

is the nature of cognition that is up for debate. They do not implicitly assume any one of the positions on offer, because what the concept should be is what is in question. The best description of their implicit concept is therefore that it is flexible.

Walter

Sven Walter (2010) replies to F&A, arguing that F&A's claim about extended selves is unfounded because the extended cognition framework is not typically applied to selves, but to classically cognitive processes (of which memory is one example). He writes that extended cognition 'has so far been restricted to paradigmatically cognitive abilities like perception, memory, thought, and language' (Walter, 2010: 126). F&A's use of the extended cognition framework to support their claim about the extension of selves is therefore problematic.

Turning from selves to cognition, Walter explicitly endorses the extended mind hypothesis (Walter, 2010: 61), but when it comes to the particular case of the impact of BCIs on cognitive processes, he says that the capacities enhanced by BCIs are all bodily, not cognitive, so BCIs are not an example of extended cognition (Walter, 2010: 67–68). He suggests that enacted cognition is a better framework for understanding the value of BCIs. This explicit endorsement of both extended and enacted cognition for different purposes suggests a flexible concept, but the implicit concepts do not quite line up with his explicit avowals.

Despite his explicit support for extended cognition, the claim that BCIs cannot extend cognition because they only enhance bodily capacities shows that Walter's implicit concept of cognition is not extended. For an extended cognition theorist, the capacities of the bodily parts contribute to an *overall process* which is cognitive. Walter seems to recognise this (Walter, 2010: 69), but still denies that it allows BCIs to extend cognition because the part of the process they replace is bodily, not cognitive. This misunderstands the extended cognition perspective. It is processes, not component parts of the underlying substrate of those processes, that can be said to be cognitive. We would not describe a neuron in the brain as cognitive, but enhancement of its activity could still contribute to a cognitive process. The same is true for bodily processes for the extended cognition theorist. As Clark says: 'The question that needs to be addressed, then is: When is some physical object or process acting as part of a larger cognitive routine? It is not the much murkier (probably unintelligible) question: When should we say of some such candidate part, such as a neuron or a notebook, that it is *itself* cognitive?' (Clark, 2008: 87–88).[47]

Walter also notes in passing some doubts 'that the intimate couplings Clark appeals to so frequently are, all by themselves, sufficient to establish cognitive extension' (Walter, 2010: 68). Walter's implicit concept of memory therefore cannot be extended in Clark's sense, which he finds too liberal (Walter, 2010: 68). He seems to have a view of cognition according to which bodily (and so presumably extrabodily) contributions to the process are mere inputs and outputs. This could be either a brainbound concept, or an embedded cognition concept. His explicit support for extended cognition makes the latter much more likely – even if he does not properly recognise the constitutive role of extra-neural scaffolding, he is sensitive to its importance. In a reversal of the stance adopted by F&A, who claim to use extended cognition only as a heuristic but implicitly endorse it, Walter claims to endorse it but does not really do so.

Does the embedded cognition concept he uses here remain consistent throughout the paper? When he comes to discuss his favoured approach – enacted cognition[48] – Walter begins to more explicitly question his concepts of cognitive states and processes. From the enacted perspective, as quoted in chapter 1, '[c]ognition is the relational process of sense-making that takes place between an autonomous system and its environment.' There is a strong emphasis on the active nature of cognition, to the extent that proponents 'seem committed to the view that in the absence of agency there can be no cognition' (Walter, 2010: 69). This is a radical departure from the embedded cognition concept that Walter seemed to be implicitly employing when he was discussing extended cognition.

[47] A similar objection to Walter is also raised by Kyselo, and I will mention this in the next subsection.
[48] It should be noted that Walter expresses doubt over whether his characterisation of the enacted cognition perspective is adequate (Walter, 2010: 66, footnote 12). This explicit doubt would not prevent Walter from implicitly making use of an enacted cognition concept, but it should make us wary of attributing one to him based only on his explicit support for the perspective with respect to BCIs.

Walter says something about concepts which may help us to figure out what is going on here, when he goes on to ask whether the enacted approach would require mere *possession* of sensorimotor skills, or the *actual exercise* of such skills for cognition to take place. In the latter case, LIS patients' cognition would be seriously impeded or prevented altogether because they cannot exercise sensorimotor skills; testing LIS patients' cognitive abilities could therefore answer his question. However, Walter goes on to note that such testing is only effective

> provided, of course, that the studies that test the cognitive capacities take into account what ENC [enacted cognition] says about cognition. That no cognitive impairments are found in LIS patients counts against ENC only if the notion of the "cognitive" underlying the experimental tests is the same as the notion appealed to by the enactivist. Moreover, the choice of which cognitive capacities to test may make a difference—a fact that may go unnoticed unless situated approaches to cognition, and in particular ENC, are taken into account. (Walter, 2010: 70).

He refers to the experiments of Schnakers et al. discussed above, noting that the cognitive capacities tested 'are more or less the sort of "offline" cognitive capacities at the focus of classical cognitivism' (Walter, 2010: 70). Here, we see Walter conceding that how cognitive states are conceptualised varies in this debate, and how they should be conceptualised is what is in question.

In light of this, the best way to see Walter's implicit concept of memory is as flexible, or open to being viewed from different perspectives. As we found when discussing F&A's concept, this flexibility of concept is important, because how to conceive of cognition is what is being debated.

Kyselo

Kyselo (2013), agrees with Walter that the enacted (enactive in her terminology) approach is the best way of thinking about BCIs and their effect on the *self*, but argues that BCIs can be seen as vehicles or realizers of certain extended *cognitive* processes. She argues effectively that, in his discussion of BCIs and extended cognition, Walter fails to realize that bodily processes are part of cognitive processes for an extended cognition advocate. This is similar to the point I made above, although Kyselo goes about making it in a slightly different way (Kyselo, 2013: 581).

Here, she seems to be adopting a similar concept of cognition to the extended cognition theorist. This can be seen particularly in her example of a patient using the cursor of a BCI to draw a virtual mind map:

> The proponent of EXT [extended cognition] could hold that this is a case of extended cognition, where in order to perform a complex mental task movement of an external tool allows outsourcing parts of the memory process as well as of our abstraction capacities. Bringing items into a logical order, or conceptualizing a dynamic relationship clearly counts as cognitive. The only difference in the case of BCI is that it is partly externally realized. (Kyselo, 2013: 581).

However, when she discusses the self, the approach advocated is an enacted one rather than extended, because extended cognition lacks a notion of what the self is (Kyselo, 2013: 582). Although I am interested in cognition rather than the self here, the enacted approach has no clear-cut separation of consciousness and cognition, and puts conscious experience at the heart of cognitive science (Kyselo, 2013: 588). This suggests that Kyselo would defend enacted *cognition* as well, and in fact there are a couple of hints that this is the case.

Introducing the criticisms of Walter that I have already discussed, she says: ‘He has sought to clarify these distinctions [between the different situated cognition perspectives]; subsequently assigning the role of the appropriate situated approach to the enactive approach to *cognition*. […] *Walter is right in this*, but for the wrong reason.’ (Kyselo, 2013: 580, my emphasis). She also says in a footnote that it might be better to use a single unified framework ‘if explanatory efficiency matters’ (Kyselo, 2013: 584, footnote 6). This unity would make sense for an enactivist, particularly given the lack of separation between cognition and consciousness on the enacted perspective. Taken together, these things suggest that her defence of an enacted approach to the self is also implicitly a defence of enacted cognition.

Again we see more than one kind of concept within a single paper – extended cognition and enacted cognition – because how to conceptualise cognitive states or processes is part of what is in question. The best way to see Kyselo's implicit concept is as flexible, or open to being viewed from different perspectives.

Heersmink

The final paper in this debate is Heersmink (2013). He focusses on the idea of tools used for extended cognition being trusted and transparent in use. Otto's notebook (Clark and Chalmers, 1998) is like this. It can play the same functional role as biomemory in a healthy subject because it has become a transparent cognitive tool that is consulted without conscious thought. BCIs, due to the difficulty of using them, are not like this.

> They [BCIs] are difficult to use and only work in highly controlled laboratory settings; they require a long, complex training period and a high degree of concentration. They have furthermore very limited control options, a highly variable use-efficiency and effector motions are often slow, clumsy and sometimes unsuccessful. Because of these limitations, an agent cannot fully trust these systems to perform their function adequately and they are not experienced as transparent extensions of the human body. (Heersmink, 2013: 214).

These limitations of BCIs could be overcome as technology improves. Heersmink endorses the extended cognition approach, so believes that this is a possibility. He discusses the use of pen and paper as an external memory store, arguing that another way in which it differs from a BCI is in the possibility for external manipulation of ideas (Heersmink, 2013: 217). Based on this observation, he suggests ways in which the interface of BCIs could be improved, for example a larger workspace where emerging text can be manipulated more easily (Heersmink, 2013: 218–219).

Heersmink adopts an extended cognition concept of memory according to which vehicles of memory can be found spread across the brain, body and world. The external vehicles must be connected in a particular way to the controlling human, viz. they must be trusted and transparent in use. Heersmink's implicit and explicit concepts are extended. Unlike the other philosophers considered here, the topic of the paper is not which perspective to use, but how BCIs can be made better vehicles for cognition on the extended cognition approach. Heersmink's concept therefore, unlike the others, does not need to be flexible within this paper.

The four philosophical papers discussed above are all part of the same debate over the role BCIs play in cognition, but the debate is approached by questioning the conceptualisation of cognition itself, and therefore of specific cognitive states or processes like memory. It is therefore not surprising that in three cases out of four it has been impossible to pin a single kind of implicit concept on the authors. Reasons for this will become clearer in the next section.

Summary

Before moving on, it will be useful to summarize what I have said about the implicit concepts of memory in each paper:

Neuropsychologists and neurologists	Allain et al.	Brainbound
	Schnakers et al.	Brainbound
	León-Carrión et al.	Embedded
Philosophers	Fenton & Alpert	Flexible (extended)
	Walter	Flexible (embedded / enacted)
	Kyselo	Flexible (enacted / extended)
	Heersmink	Extended

(Types of concept listed in brackets represent positions that seem to be held at some point in the text, see above for details.)

The concept of memory has been shown to vary in terms of whether a situated cognition perspective is appropriate, and if so, which one. There are a number of possible positions on offer, as shown in this section, with the main contrast being between neuropsychologists and neurologists taking memory to be something that happens in brains, and philosophers having a concept that is more open to situated cognition perspectives, but is flexible between which of them should be used. I will now go on to explore how the epistemic niches for MEMORY shape the concepts, and whether those epistemic niches are composed of internal considerations, untried hypotheses, or external considerations.

Identifying and assessing the epistemic niches

Neuropsychologists and neurologists

As outlined in previous chapters, the epistemic niche of a concept is constituted by the surrounding parts of the framework which apply pressure to that concept, defining the role it is needed to play. Here I will examine some of the most important pressures on MEMORY for this case study, beginning with the neuropsychologists and neurologists.

Focus on the brain

The first three papers considered here (Allain et al., Schnakers et al., and León-Carrión et al.) share the common idea of testing for cognitive impairment in LIS patients, and all include testing memory as part of this. They are all co-authored by a mixture of neuropsychologists and medical practitioners or researchers specialising in rehabilitation of brain injury patients. All three come from journals with a focus on the brain, or neurology more generally. There is therefore a pressure towards a neuro-centric concept of cognitive capacities like memory from both the training of the authors, and from the journals they are submitting to. Allain et al. and Schnakers et al. use standard psychological tests for memory. As neuropsychologists, they see these tests as tests of brain function, although the testing method is psychological.

Is a focus on the brain an internal consideration for neuropsychology and neurology? It looks as though the answer to this is obviously "yes" because such a focus defines the fields. But this alone only makes it a presupposition, not necessarily an internalised criterion. Domain formation is connected to conceptual development and should also become increasingly dependent on internal factors. In light of this, the best approach here is to look at the development of these neuro-centric fields to find out and examine the reasons for their foundation. The important question is whether their domains were formed for internal reasons.

The term "neuropsychology" '…is relatively new (Bruce, 1985), having gained currency only in the 1950s when it displaced older terms, e.g., psychoneurology (Bekhterev), brain pathology (Kleist).' (Benton, 2000: 3). The 1950s was a time when behaviourism was still significant, but beginning to be challenged by cognitivism. The neurosciences grew out of this focus on the physical basis of the mental. I will consider two particularly important factors in the formation of these domains here: localisation in the brain, and the development of technology.

Localisation in the brain was of particular relevance for the development of the neuro-sciences, and this goes back at least to phrenology in the late 19th century, and to discoveries that damage to particular brain areas resulted in particular pathologies (e.g. Broca's discovery in 1861 of an area of the brain related to speech production). This idea is linked to a faculty view of psychology, and the theory that a particular brain area corresponds to a particular mental faculty. Traces of this view can be seen in the papers described here, with memory being used as an example cognitive capacity (a particular faculty) that is explicitly linked to the brain. Discoveries such as Broca's represent successes because they have led to fruitful subsequent research and applications (e.g. for brain damaged patients). Therefore localisation to particular brain areas was internalised by the sciences of the mind (although as with everything else in science, this was a defeasible assumption).[49] The new disciplines such as neuropsychology coalesced around this success.

A second major factor in the formation of the domains of neuropsychology and neurology was the development of various technologies. I will look at whether this factor is internal to neuropsychology first, then move on to consider neurology.

Neuropsychology is an interdisciplinary enterprise, therefore 'the history of neuropsychology has been one of irregular progress as advances in one or another of its contributory disciplines were achieved and made an impact on thinking and practice in the field.' (Benton, 2000: 4). Many of these advances have been technological, for example recent improvements in brain imaging techniques such as MRI and CAT scans. Shaping research in tandem with such technological improvements has been successful in terms of linking research into the mind with medical applications (see also Rand and Ilardi, 2005: 15–16). Following technological advancements in this way was therefore internalised into neuropsychology.

[49] The idea that particular areas of the brain are responsible for particular mental states or processes has not continued to be successful, and in fact the data suggests very strongly that this is not the case. Instead it now appears that each region of the brain is involved in several mental states and processes, and each mental state or process requires several brain regions. The response to this may be to change our categories to ones which fit the neural data more neatly, or to accept that mental states and processes are multiply realized at the neural level. Finding out which of these is the case is part of the motivation for the Cognitive Atlas project (see chapter 2), which includes fMRI data for each mental concept listed. The focus on the brain in some form remains an internal factor however.

Neurology, unlike neuropsychology, is a branch of medicine.

> [T]he origins of the specialty in hospitals and universities can be located only in the mid-nineteenth century and afterwards, a fact that makes it difficult to identify the way neurological knowledge was systematically acquired before this period (McHenry, 1969; DeJong, 1982). Moreover, the inclusion of neurology as a discipline recognized within the medical profession in terms of autonomous hospital departments, university teaching and professional certification occurred only around the period of World War I (Magoun, 1975). (Casper, 2010: 638).

There were some social reasons, particularly linked to the war, that were important in neurology's formation:

> Historians routinely associate the intellectual origins of American neurology with studies of wounded soldiers first conducted by Silas Weir Mitchell and William A. Hammond during the American Civil War…As these physicians began elaborating a set of practices that would trigger the local emergence of modern neurology throughout North America, a simultaneous convergence of social, economic and intellectual forces in the commercial and industrial milieu of the post-bellum United States made the formation of clinical specialties increasingly acceptable to the medical profession and the public. Thus, within a few years of the conflict of 1861–65, specialist organizations devoted to the diagnosis and treatment of nervous conditions began to decorate the American medical landscape. (Casper, 2010: 638).

The history of neurology is closely tied to changing dissection techniques and results. Vesalius dissected out nerve tissue in the 16th century, but he did not typically remove the brain from the skull. Later, this began to change:

> Perhaps Thomas Willis (1621–1675) was the founder of neurology… and the arterial circle at the base of the brain is one of Willis' eponymous claims to fame…Willis removed the whole brain from the body instead of dissecting from above and the cerebral body thus removed was seen to contain important solid portions. Earlier workers had concentrated on the ventricles, perhaps echoing William Harvey's emphasis on the solid portions of the heart rather than its cavities—the empty areas. Solid organs influenced the movement of fluids, the opposite view to that of early classical physiologists and Cartesians. We might think of a bottle turned upside down where the fluid drains, the fluid moving and the solid walls remaining unmoved. Here we have the solid portions of the body being the more active. Removal of the brain from the body liberated the theory of the humours of coldness,

> moisture, dryness and heat. The cerebral body was part of the whole of man's body, a body which was beginning to be anatomized. (Gardner-Thorpe, 2000: 2573–2574).

In addition to changing dissection techniques, the invention of and subsequent improvements in microscopy were important. E. Purkinje (1787–1869) in 1837 gave the first description of neurones, a very early description of cells of any kind. Electricity was another relevant technology, from Galvani's work in the 18th century on the effect of electricity on the muscles, to the use of electricity as a therapy (e.g. ECT).

The development of neurology was therefore led by dissection techniques, microscopy and the use of electricity in the same way that the development of neuropsychology was led by technological advances. Following technology in this way was successful, resulting in the medical applications mentioned (brain scanning techniques, treating wounded soldiers) and to some degree of predictive success, e.g. predicting what a brain scan of a person will look like while they are doing a certain kind of task (e.g. a motor task), although more specific prediction (mind reading as the press would have it) has proved elusive.

Shapere says of technology:

> [I]n particular, the instruments that particle physicists or astronomers or molecular biologists employ, while their detailed design and construction require an engineering expertise, are now almost invariably proposed and constructed in the light of some scientific theory in the area in which they are to be used. To that very great extent, technology has itself been "internalized" into the disciplines in which it is employed, affecting the problems and strategies of that discipline as the latter affects it. (Shapere, 1986b: 21).

The same seems to have been the case for technology in neuropsychology and neurology, so this consideration is internal to those subdisciplines.

The development of neuropsychology and neurology out of existing disciplines such as psychology and medicine followed a path at least largely dependent on internalised considerations – here I have looked at localization and technology. The focus on the brain is therefore internal to these neuro-centric disciplines.

Need for easy testability

Another pressure on the neuropsychologists' and neurologists' concept of memory is the need for diagnostic criteria. The papers aim to investigate the assumption that LIS patients have unimpaired cognition and consciousness. This has never been an unquestioned assumption, but it is notoriously difficult to investigate, given the challenges in communicating with patients, the difficulties of differential diagnosis between LIS and minimally conscious state and persistent vegetative state, and the occurrence of additional brain damage that may result in cognitive impairment in some cases. Diagnosis is often slow (over two months on average, León-Carrión et al., 2002, survey: 573) and this is obviously traumatic for the patient and their friends and family. This creates a pressure to find clearer diagnostic criteria that can be easily and quickly applied.

Given the aim of finding out exactly what state LIS patients are in, and how they should best be treated, there is pressure towards a concept of memory according to which it can easily be tested. Because LIS results from brain damage, it is natural to test for brain impairments. The need for easy testability therefore creates pressure towards a brainbound concept of memory.

For neurology, because it is a branch of medicine, the biggest part of what it is to be successful is improved patient outcomes, and the differential diagnosis referred to here is an important part of that. For severely brain-damaged patients, other factors include establishing communication of some kind, and getting the patient into a stable enough condition that they can live at home. Focussing on easily testable measures has been internalised because it has led to these kinds of success. Neuropsychology on the other hand is a research discipline, not an applied medical one. Results that lead to improved patient outcomes are *part of* what it means to be successful, but prediction and explanation of the world is a larger component of success for a discipline of this sort. Pressure to design testing methods that are as quick as possible is an external pressure that comes from the applied medical disciplines, not one internal to neuropsychology itself. It is relied on as a criterion, but lacks the relevant history of success in neuropsychology, so for them, selecting a testing method dependent on a purely brainbound concept just because it is quicker and easier to administer is an external consideration.

This conflict between the two disciplines comes about because they have different histories, different aims and measures of success, and therefore have internalised different criteria. It highlights the problem of interdisciplinary collaboration; what considerations should be used when members of two disciplines work together?

Brigandt and Love's problem-centred approach to multidisciplinarity is useful here (Brigandt, 2010; Love, 2008; Brigandt and Love, 2010). "Multidisciplinarity" is the term used to describe disciplines transiently working together, while the term "interdisciplinarity" is reserved for new disciplinary units formed by the merging of existing disciplines (Love, 2008: 876, note 1). What we have here seems to be a multidisciplinary collaboration between neuropsychology and neurology (although this doesn't rule out some stable new discipline forming from their collaboration in the future).

According to the problem-centred approach, the contribution each discipline should make and which discipline(s) should be explanatorily fundamental varies with the problem at issue (Brigandt, 2010: 295). Each problem will have its own set of questions, or problem agenda (Love, 2008: 877). For Allain et al. and Schnakers et al., the project is to find out whether LIS patients have unimpaired cognition. This problem agenda is mostly research-based, rather than having directly medical aims (although of course there will be ultimate benefit for patients and this is one important motivation). Therefore it seems that the considerations internalised by neuropsychology should have more weight in this case.[50] The pressure to use a purely brainbound concept of memory just because it is easier to test is an external consideration for the authors of these two papers; although it may serve the aim of a medical discipline, it does not serve the research aim.

[50] This is my own application of Brigandt and Love's central idea, and is somewhat different from their own examples, which involve integrating approaches from subdisciplines working at different levels in biology. They do not use the notion of internalisation.

For León-Carrión et al., the main aim is slightly different: rehabilitation of the patient. This is a more directly medical aim, so neurology is more fundamental than neuropsychology in this case. Easy testability is therefore an internal factor for León-Carrión et al. This factor therefore should be allowed to shape León-Carrión et al.'s concept, but not Allain et al.'s or Schnakers et al.'s. However, easy testability applies pressure towards a brainbound concept, and according to what I argued above, León-Carrión et al.'s is the only paper to *not* employ such a concept. As I said above, León-Carrión et al.'s paper has at least some openness to an embedded cognition concept. So far I have only discussed factors that apply pressure towards a brainbound concept, so nothing I have said accounts for this. I will now attempt to remedy that by discussing a central consideration that is particular to León-Carrión et al.'s work.

Patient-centred care

Care and attempted rehabilitation of the patient after diagnosis is the major concern for León-Carrión et al. This includes more effective communication to improve quality of life, and a realistic understanding of the prognosis of patients. Doctors often offer little or no treatment in the belief that LIS patients have no prospect for recovery (47.1% of those surveyed in León-Carrión et al.'s study were not currently being treated). Challenging this is one aim of León-Carrión et al.'s work (León-Carrión et al., 2002, review: 561–567).

I have argued that León-Carrión et al. have an embedded cognition concept of memory that is at least open to the importance of external scaffolding. The aim of their paper is to survey patients to collect data which suggest diagnostics and rehabilitation procedures (León-Carrión et al., 2002, survey: 571). While all the papers I am looking at are interested in diagnostics and rehabilitation, León-Carrión et al. differ in their focus on the patients' point of view. The tests applied by Allain et al. and Schnakers et al. are to be administered by a specialist experimenter who will decide, based on the results, whether memory is impaired. León-Carrión et al.'s method of asking the patients means that the judgement of impairment is left to the patients themselves. The patients do not need a concept of memory that can be tested using a neuropsychological battery, or one that measures the extent and type of their brain injury; they need one that represents how they remember in their everyday lives. Therefore, there is no pressure to exclude external scaffolding for memory, and some pressure to include whatever they may have access to via technology such as a BCI, or communication with other people.

While the authors are from the same subdisciplines as those of the previous two papers discussed, what counts as success for them is different (or rather, while the same factors are important for success, the emphasis placed on them is different). Brigandt and Love's problem-centred approach can be used again to clarify this. As I said in the last subsection, the problem agenda is different for León-Carrión et al., so the weights given to the contributions of the different subdisciplines will be different too, with neurology having priority over neuropsychology. Of the factors internalised by the relevant subdisciplines, the important ones are those that were internalised due to past successes relative to the aims that León-Carrión et al.'s research emphasizes. In this case, that means factors that led to improved diagnostics and rehabilitation procedures. Patient-centred care is one such factor.

Taking patients' own views seriously (and the views of their friends and families) is an important component of patient-centred care. This is an approach that has proved important in the treatment of LIS patients, for example in management of anaesthesia. Yoo et al. (2012) report how the use of a bispectral monitor can help to determine whether a patient is awake or not, and therefore allow an appropriate level of anaesthesia during an operation and better post-operative care. This is particularly important because, as they demonstrate in their paper, an LIS patient can be awake while his eyes are closed, even if he has preserved eyelid movement and at other times opens his eyes when awake. Because the LIS patient can neither move nor speak, he has no way of communicating that he is awake, so this monitoring is of crucial importance during and after an operation. In the case Yoo et al. describe, it was initially used because the wife of the patient raised concerns that nurses were treating him as though he was asleep when his eyes were closed. She claimed to be able to tell that sometimes he was awake in these situations, and expressed concerns about pain management for an upcoming operation for a pressure ulcer.

Yoo et al. provide a list of '[p]atient-centred concerns for locked-in syndrome (LIS)', the first of which being '[p]atient-centred care focuses on the patient and family concerns and needs throughout the hospitalization.' (Yoo et al., 2012: 7, table 2). Taking the patient's point of view and his wife's concerns seriously in this case contributed to the aim of improved patient care. Before, during and after the operation, the surgical team and nurses were able to tell when the patient was awake and when he was asleep, and treat him accordingly.

More widely, patient-centred care has proved effective for chronic conditions. In the introduction to a theme issue of the British Medical Journal on managing chronic diseases, Edward Wagner says:

> Despite the clinical differences across these chronic conditions, each illness confronts patients and their families with the same spectrum of needs: to alter their behaviour; to deal with the social and emotional impacts of symptoms, disabilities, and approaching death; to take medicines; and to interact with medical care over time. In return, healthcare must ensure that patients receive the best treatment regimens to control disease and mitigate symptoms, as well as the information and support needed effectively to self manage their health and, in many instances, their death. (Wagner, 2002).

The introduction is subtitled "The efficacy of coordinated and patient centred care is established, but now is the time to test its effectiveness". The points made in the passage just quoted make this efficacy unsurprising, and they are very much applicable to LIS. LIS patients and their families, like sufferers of other chronic conditions, must adjust to a new way of life.

Recognising this has led to improved patient care, and in some cases evidence suggests it can improve physical health. For example, Wagner cites Norris et al.'s study (2002) on the use of self management education (a form of patient-centred care) for type 2 diabetes. The conclusion of Norris et al.'s study is that '[s]elf-management education improves GHb levels at immediate follow-up, and increased contact time increases the effect.' (Norris et al., 2002).[51] Taking patients' and their families' concerns and points of view seriously is a consideration which has been internalised due to past success, both in terms of physical outcomes and patient experience.

[51] Although the conclusion continues '[t]he benefit declines 1–3 months after the intervention ceases, however, suggesting that learned behaviors change over time. Further research is needed to develop interventions effective in maintaining long-term glycemic control' (Norris et al., 2002).

This is the context of the consideration of the patient's point of view that I have argued pushes León-Carrión et al.'s concept of memory in the direction of being open to external scaffolding. Their implicit embedded cognition concept of memory is therefore legitimate to the extent that it is shaped by this factor.

Summary

In summary of what I have said about the epistemic niche for the neuropsychologists and neurologists:

Factor applying pressure to concept	**Concept encouraged by this pressure**	**Internal / Hypothesis / External**
Brain focus	Brainbound	Internal
Testability	Brainbound	External for Allain et al. and Schnakers et al.; Internal for León-Carrión et al.
Patient-centred care (mainly affects León-Carrión et al.)	Embedded	Internal

For Allain et al. and Schnakers et al., focus on the brain is an internal factor, and easy testability is an external factor. Both of these factors apply pressure towards a brainbound concept.

Recalling that a concept is both legitimate and rational to the extent it is shaped by internal factors, the brainbound concepts of memory found here are legitimate and rational to the extent that they arise as a next step in the history of successful brain-focussed subdisciplines, but non-legitimate and irrational to the extent that they are shaped by the need for a concept which makes memory easily testable. Therefore, although holding a brainbound concept may well be legitimate for the time being, in the event of new factors applying pressure towards a situated cognition concept becoming relevant, these authors and others like them are at risk of holding too tightly to a brainbound concept in response to the external pressure towards easy testability. One such new factor may be patient-centred care, for example if they were to undertake a project like León-Carrión et al.'s.

For León-Carrión et al., all the factors considered here are internal. Focussing on the brain, and easy testability apply pressure towards a brainbound concept, while patient-centred care (taking patients' own views seriously) applies pressure towards an embedded cognition concept. Here we see a tension between two aims focussed on patient welfare: the aim to diagnose patients as quickly as possible (best facilitated by a brainbound concept) and the aim to aid their recovery (best facilitated by a situated cognition concept). This is perhaps why León-Carrión et al. are using an embedded cognition concept, rather than a "stronger" (i.e. less brainbound) situated cognition perspective such as extended cognition. To the extent that it is shaped by the pressures considered here, this concept is legitimate and rational. It is impossible to say more than this for sure here, but the tension discovered suggests that different kinds of concept will be applicable in projects that have different aims, so there is no way to stipulate a single type of concept that neuropsychologists and neurologists should be using. I will return to this point in the conclusion to this chapter.

There is some evidence here that both brainbound and embedded cognition concepts are currently functioning as investigative kinds in neuropsychology and neurology, although there are some tensions in the epistemic niches such that we cannot conclude that both kinds of concept should continue to be used in future projects in these subdisciplines. I will now go on to give the same treatment for the philosophical papers.

Philosophers

The philosophers I am looking at here are all working within a part of philosophy that falls under the interdisciplinary banner "cognitive science". They also all fall under the banner of ethics – the papers were all published in the journal *Neuroethics*. Much of what I say here is perhaps only applicable to this particular subset of philosophers. Talking about philosophy more broadly would require an argument that my Shaperean methods can be applied outside the sciences broadly construed, as well as much more discussion of the diverse aims and standards for success in philosophy as such.

When compared to the neurologists and neuropsychologists, the epistemic niche for the philosophers' concept is very different. While the philosophers are interested in rehabilitation and the quality of life of patients, their interest is from the point of view of philosophical ethics, rather than immediate bedside care. None of the authors are medical practitioners, so the demands on them are very different. The most important pressure on their concepts is perhaps that towards being able to question their concepts.

Questioning concepts

The philosophers are open to questioning their concepts of cognitive states or processes like memory, and this openness to questioning creates a pressure towards having flexible concepts. Questioning their concepts of important phenomena is a major part of what philosophers are trained to do. In the context of the situated cognition debates, what cognition *is*, and therefore what particular cognitive states are, is what is in question. This can be seen, for example, in the debate over the "mark of the cognitive" in the extended cognition literature (e.g. Adams and Aizawa, 2001; Clark, 2008). This is a debate over what features something must have in order to be cognitive, and over whether listing such features is the right way to go about analysing cognition in the first place (see Hurley, 2010 for an example of someone who believes it is not).

Is this questioning of concepts internal to philosophy? I suggest that it is, because it has allowed philosophers to successfully meet the philosophical aim of engaging in debate. Questioning concepts is a very important part of debate in traditional philosophy; but more than this, philosophers debate over their aims, what would count as a good theory or explanation, what their methods should be etc. Philosophy seems to display far more radical plurality and disunity than even psychology.

According to the account I developed in chapter 3, this is not a sign of immaturity in philosophy. I would argue that, unlike psychology which has not yet internalised either unity or plurality, philosophy has internalised plurality. Over time, plurality has proved more successful than unity within philosophy. The nearest philosophy seems to come to unity is widespread general movements like logical positivism, although there is of course still much disagreement within them. But such movements have always collapsed, with logical positivism now being famously "dead, or as dead as a philosophical movement ever becomes" (Passmore, 1967).

Debate about the fundamentals of the discipline is a proper part of what philosophy is, with metaphilosophy and the philosophy of philosophy being hotly contested. Timothy Williamson says that '[t]he philosophy of philosophy is automatically part of philosophy, just as the philosophy of anything else is' (Williamson, 2007: ix).

This radical plurality requires concepts that are open to question in order to debate effectively. Given that there is disagreement about almost all aspects of theory and practice in philosophy, it is not surprising that there should also be disagreements about concepts, given what I said in chapter 2 about conceptual development. Concepts are shaped by the theories and practices of which they are a part. In order to debate fruitfully with this level of disagreement, it must be possible to develop a reasonably high level of understanding of other disputants' points of view, and this is why philosophers must have some flexibility in their concepts. They must be able to envisage how a different concept would fit in with a different framework of theory and practice, and where the main tensions with their own perspective lie.[52]

In more modern philosophy of cognitive science, openness to questioning concepts has also proven successful by allowing philosophers to engage with a variety of scientific literatures and to collaborate with scientists from a variety of subdisciplines whose concepts are not the same as their own. This is another source of success that has led to the internalisation of having concepts that are open to question.

In the sections that follow, I want to address several factors that arise from the situated cognition literature. All the philosophers I have looked at in this case study explicitly advocate one of the situated cognition perspectives (although, as discussed

[52] This flexibility of concepts is perhaps part of the reason why attempts to elucidate our concepts by means of lists of necessary and sufficient conditions have met with no success.

above, their implicit concepts may not match their avowals). All four papers refer to Clark and Chalmers' paper proposing the extended mind thesis (Clark and Chalmers, 1998), in which extended memory in the form of a notebook is one of the main examples. Addressing this example, and how BCIs are relevantly similar or dissimilar, is important for all of them. Because of this, certain features of the literature that has grown up around Clark and Chalmers' paper are major factors in the epistemic niche of MEMORY for these philosophers.

Science fiction

The first of the factors derived from the extended cognition literature is the focus on new technology, particularly that with a science fiction flavour. The extended cognition literature is very much concerned with the so-called cyborg fantasy examples (the term is due to Wilson, 2010: 173) and claims to offer us a way of seeing cognition that is appropriate for the modern world and our increasing integration with technology (see e.g. Clark, 2003). BCIs bring out this element in the debate. The idea of the completely paralysed patient using a technological avatar is very similar to the science fiction scenarios that inspire a lot of writers on extended cognition. The prevalence of this literature creates a pressure towards extended cognition concepts of cognitive states and processes.

This pressure is clearly external to philosophy. While science fiction scenarios can provide a useful heuristic, they do not provide criteria that philosophical accounts should meet. There is no history of success attached to endorsing theories that sound particularly like science fiction scenarios.

Social explanations

Another factor deriving from the literature is the social nature of cognition. The dependence of the LIS patient on communication with others highlights this factor. For someone who wants to argue that social interaction is essential for or constitutive of cognition, the LIS patients' situation is likely to be interpreted as one in which their cognitive abilities are dependent on or partly constituted by communication with their care-givers. This creates a pressure towards seeing cognitive states or processes from a situated cognition perspective, particularly from the distributed cognition perspective (although also to some extent from the extended and enacted cognition perspectives).

To see whether this consideration is internal or external or a hypothesis for the authors in question, we need to assess whether giving socially based explanations has a history of success in philosophy. In many branches of philosophy (e.g. philosophy of science, epistemology) social theories do have a long history. They have provided a valuable foil for more individualist accounts of knowledge, enquiry etc. and in this sense have led to some success. However, providing socially based explanations is not something that has come to be used as a criterion in philosophy generally, or in any of these sub-branches of philosophy. It is merely one type of theory on the table here, albeit one the authors in some situated cognition traditions tend to favour. To favour such explanations just because they are social is to prejudge the question of whether these views of cognition are adequate, and it is therefore an external factor in the epistemic niche for MEMORY in philosophy.

Ethics

A third important factor coming from the situated cognition literature is a concern with the ethical implications of a situated cognition perspective. The debate in the papers I am considering here takes place in an ethics journal, so it considers the potential ethical implications of BCIs playing a role in cognition. These focus particularly on autonomy, selfhood and patient choice, all things which may be enhanced by a BCI.

Considerations about autonomy and selfhood are particularly philosophical ways of looking at the issue, compared to the more obvious medical concerns with pain and being able to express it, and the ability to carry out everyday tasks. The philosophical literature on topics like the self and its relation to cognition is highly complex, and this complexity is part of the epistemic niche for MEMORY here. There is much debate in the literature over the role of cognition in selfhood, the role of a BCI in cognition, and whether considering such enhancement from a situated cognition perspective would have good or bad consequences. The number of competing perspectives here means that consideration of ethical issues applies pressure towards flexible concepts of cognitive states; having such flexible concepts is the only way to adequately assess the competing perspectives.

Ethical considerations are internal to philosophy. Analysing ethical implications has allowed philosophy to have success in its aim to answer questions about the sorts of things people care about. However, perhaps it is not always the case that such implications should be allowed to lead and shape philosophical theories. It might be thought that it is appropriate that they do so in value philosophy, but not in metaphysics for example. Although the ethical implications are important, we might think that the world is as it is independently of them, and we should theorize about it as such. If we grant that this is right, would the question about MEMORY be similar?

Recall from chapter 1 that the question I am interested in is an epistemic one about investigative kinds. The concept used should be that which best facilitates fruitful science. Ethically sound science is a part of this, so the question is not entirely separate from ethics, even if we grant that a metaphysical question might be. Ethical considerations are internal to this kind of philosophy.

One possibility that might concern us is that these authors are only considering the moral dimension because they are trying to get published in *Neuroethics*. Publication in a particular journal is not the sort of thing that can establish a track record of success, so cannot be an internal consideration. Is this the motivation in this case?

I think actual motivation is not the important question here. Recall that Shapere's internal/external distinction is not about motivation, but rather the theories and practices of the discipline. Ethical considerations are internal to philosophy, and so they should be allowed to shape philosophers' concepts, even if they are being employed for other reasons like journal publications in a particular case. If a particular experimental method has a proven track record in science, experiments done using that method should not be considered less reliable just because the scientist has been bribed to perform them on this particular occasion (provided we can be sure that the method has been followed properly). Similarly, because ethical considerations have a proven track record in general, the fact that they might be being employed for other reasons in this specific case is not important, provided those other reasons do not interfere with the application of the considerations. And in this case there seems to be no reason to think that the philosophers concerned will be any worse at thinking through the ethical dimension of the problem just because their motivation may be publication in an ethics journal rather than thinking about this aspect for its own sake. Ethical considerations are therefore an internal factor in the epistemic niche for MEMORY in this case.

Having an impact on another field

Another pressure on philosophers' concepts that is relevant here, although it is not specific to the situated cognition literature, is the pressure to make a difference in a field outside of philosophy. Because the kind of philosophy I am considering falls under the banner of cognitive science, the ability to do interdisciplinary work is an important aim. Philosophy that can make a difference to another field can most successfully meet this aim, so this is an internal factor.

In this case, this factor applies pressure towards the situated cognition concepts, in particular extended cognition. This is because encouraging scientists and medical practitioners to view BCIs as a part of the substrate of the cognitive process is a concrete recommendation that philosophy can make that would affect practice in these other disciplines. Because the extended cognition approach is probably the one to have received the most philosophical attention, there is most to say from that perspective. We can see the response to this pressure particularly in Heersmink's paper, with his analysis of how BCIs might be improved to better meet the criteria for realising part of an extended cognitive system.

Summary

In summary of what I have said about the epistemic niche for the philosophers:

Factor applying pressure to concept	Concept encouraged by this pressure	Internal / Hypothesis / External
Questioning concepts	Flexible	Internal
Science fiction	Extended	External
Social explanations	Distributed (/extended / enacted)	External
Ethics	Flexible	Internal
Impact in another field	Extended	Internal

The internal factors for the philosophers' concept are: openness to questioning concepts, considering the ethical implications of theories, and having an impact on another field. Openness to questioning concepts creates pressure towards a flexible concept of memory, which may is to being viewed from different perspectives on cognition. Considering ethical implications also creates some pressure towards flexible concepts. Having an impact on another field creates some pressure towards the situated cognition concepts, particularly extended cognition.

The external factors are: science fiction scenarios, and giving explanations from a social perspective. The first of these applies pressure towards an extended cognition concept, and the latter towards a distributed cognition (or possibly extended or enacted) concept.

Most of the authors considered here have flexible concepts, and this is legitimate and rational to the extent that it is shaped by openness to questioning concepts, and ethical implications. Heersmink's extended cognition concept is perhaps too rigid according to my analysis, maybe as a result of giving too much weight to external considerations such as science fiction scenarios. However, it is legitimate and rational to the extent that it is shaped by the pressure to have an impact on another field. In addition, having flexible concepts may mean defending a certain concept at a certain time for the sake of argument (as all the authors here do at some point in their papers). It would therefore be premature to condemn Heersmink's use of an extended cognition concept as non-legitimate. What is suggested here is that he (and others in the field) should not hold too rigidly to this concept in all circumstances.

There is evidence here of a flexible concept of memory functioning as an investigative kind in philosophy, and some evidence, albeit more tentative, of an extended cognition concept also doing so.

Conclusion

The above work has shown that cognitive scientists working mainly in neuropsychology and neurology tend to have a different concept of memory from those working mainly in philosophy, although there is also some variation within these groups. This difference is due to differences in the epistemic niches of the concepts – what the authors need their concepts to achieve in the work they are doing. These factors create pressure on the concept of memory, which the concept is shaped by.

Whilst it is not possible to consider every factor in any scientific framework, I have tried to identify the most significant in the epistemic niche for MEMORY. Some of the factors identified have been dismissed as external to the relevant subdisciplines, and the concept of memory therefore both non-legitimate and irrational to the extent that it is shaped by pressure from these factors. No untried hypotheses were found among the factors considered in this case study, perhaps because all the subdisciplines in question are relatively mature.[53]

When just the internal considerations are taken into account, neuropsychologists and neurologists have some pressure towards a brainbound concept, and some pressure towards an embedded cognition concept, particularly when they are carrying out work that aims to take patients' own views into account. These are the two kinds of concept that I argued the scientists considered here are in fact employing. The philosophers have pressure towards having flexible concepts which can be considered from various perspectives, including the situated cognition perspectives. They also have some pressure towards accepting a situated cognition concept (particularly extended cognition) to the extent that their work tries to have an impact in fields outside philosophy. I argued that most of them do employ a flexible

[53] The next case study will provide a sharp contrast in this respect.

concept, and the one who does not (Heersmink) employs a situated cognition concept (in this case extended).

Various different concepts are currently functioning as investigative kinds in the subdisciplines considered here. There is quite strong evidence for brainbound and flexible concepts functioning as investigative kinds, and some more tentative evidence for embedded and extended cognition concepts also so functioning. These subdisciplines should therefore remain pluralist with respect to their concept MEMORY, at least for the time being. A possible advantage of this can be seen, particularly with respect to the tension identified in León-Carrión et al.'s work between easy testability applying pressure towards a brainbound concept, and patient-centred care applying pressure towards an embedded cognition concept. This tension suggests that different kinds of concepts (and their accompanying frameworks) are best for different kinds of projects.

7. Constructing Memory in Political Scandals: A Case Study

Introduction

This case study focuses on a study of memory by cognitive psychologist Ulric Neisser, and a methodologically different study of a similar case by discursive psychologists Derek Edwards and Jonathan Potter. Edwards and Potter first criticise Neisser's work, before going on to present their own study as an improvement in the respects on which they base their criticism. In this sense, the two works constitute a debate over how memory should be studied, fulfilling my desideratum that contrasting work should be the basis of case studies to get the most out of my methods. There are also various similarities between the cases that provide useful points of comparison, particularly regarding veridicality and the notion of memory as constructed, as will become clear below.

Both case studies concern political scandals. Neisser's work (Neisser, 1981) concerns a case from the U.S. in the 1970s, namely John Dean's testimony in the Watergate scandal. Edwards and Potter's case (Edwards and Potter, 1992) is a political gaffe from the UK in 1988, involving the then chancellor, Nigel Lawson. Both Neisser and Edwards and Potter are interested in the role of context, and both are looking at discourse of a kind, although they disagree on how these things should be treated.

Neisser advocates the so called "ecological" approach (Gibson, 1979), in which subjects are observed in real world contexts rather than in artificial laboratory tests. His work on John Dean's testimony is an example of this approach, and is reprinted in a collection of papers demonstrating this approach for memory (*Memory Observed*, Neisser, 1982).

Edwards and Potter on the other hand are among the founders of discursive psychology, a form of discourse analysis applied to human psychology that was developed in the late 1980s and 1990s. This is an approach that takes ordinary contexts even more seriously by focussing on the discourse that takes place in these contexts. They accuse studies like Neisser's of only looking at the world through the laboratory window (Edwards and Potter, 1992: 32).

This focus on the importance of context, but disagreement over its precise role, makes these cases good places to look for situated cognition perspectives, since it is the role of context (that which is external to the organism) that is of interest in situated cognition, and is disputed between its different varieties. It is this dispute which makes this case study a good one for my project.

I will first outline the two cases, then go on to identify the implicit concepts present in each, arguing that both are using situated cognition concepts of memory, although of different kinds. I will then go on to identify the most important features in the epistemic niche for MEMORY in each case, and discuss which of them are internal, which are untried hypotheses, and which external to the branches of science in question.

The case study

The Watergate scandal occurred in the USA as a result of a break-in at the Democratic National Committee headquarters at the Watergate office complex in 1972. It was alleged that the break-in was organised by the president Richard Nixon and his Republican government, and that this was subsequently covered up. The scandal resulted in Nixon's resignation in 1974. The investigation involved the trial of many of Nixon's advisors, including White House Counsel John Dean. Dean volunteered himself as a key witness for the prosecution, and in return pleaded guilty to only one offence, resulting in a reduced prison sentence. It is on his testimony that Neisser focusses his investigation.

Shortly after Dean had given his testimony, it was discovered that conversations taking place in Nixon's office had been covertly tape-recorded. There was therefore a record of important conversations that Dean had recounted in court, to which his account could be compared. Neisser notes that much of Dean's testimony was wrong both in terms of verbatim recall, and the gist of the conversations (two categories of memory used in cognitive psychology). Yet Dean was not taken to be lying. Neisser therefore proposes a third category of memory, "repisodic memory", to describe what *was* correct about Dean's testimony. Repisodic memory is memory which appears to be episodic memory – recall of particular autobiographical episodes – but is not. Instead, 'what seems to be an episode actually *re*presents a *re*petition…He is not remembering the "gist" of a single episode by itself, but the common characteristics of a whole series of events' (Neisser, 1981: 114). Dean's testimony was therefore taken to be an honest account of what was going on at the White House, and of Nixon's involvement, although it did not have many of the features we would typically take an accurate memory to have.

Neisser takes many of the more specific errors in recall to be the result of Dean recounting the conversations as he would have liked them to have gone, not because he was trying to mislead, but because this was how he remembered them.

Edwards and Potter criticise Neisser's study for the way it treats discourse, for Neisser's attribution of errors to features of Dean's personality (his remembering conversations as he would like them to have gone), and for his notions of truth and error.

For Edwards and Potter, discourse should be the object of study. They accuse Neisser of seeing it instead only as a way to discover the inner cognitive states. They say:

> Instead of following Neisser's ecological cognitivism and attempting to use Dean's testimony as a pathway to the nature of the cognitive processes that allow him to remember correctly, we took Dean's testimony, and the various reports of events in the Oval Office it contained, as discursive acts which were part of broader activity sequences involving blame, responsibility and mitigation. (Edwards and Potter, 1992: 156–157).

This is an expression of the attitude to discourse that characterises discursive psychology. Discourse is not just an expression of internal memory states on this view; rather, memory is constructed through the discourse in a way that is appropriate to the context and what the rememberer is trying to achieve in that context (for example apportioning blame to others, or mitigating one's own responsibility). This attitude constitutes a general move advocated by the discursive psychologists away from the psychology of inner cognitive states and toward the public acts of discourse and what they achieve. Discursive psychology is not necessarily anti-representationalist, but discourse is not chiefly interesting for the inner states it represents, but in its own right. In Edwards and Potter's words: 'The study of situated discourse redefines and relocates the relations between language and understandings, and it does this by placing language as representation (whether of cognition or of reality) in a position subordinate to language as action.' (Edwards and Potter, 1992: 158).

Their criticism of Neisser's explanation of Dean's errors can be seen in a similar light. Neisser attributes misremembering to Dean's ego – his desire for the meetings to have gone a certain way. However, there is no corresponding account of his accurate memories. For Edwards and Potter, both accurate and inaccurate memories should be explained symmetrically, by reference to the pragmatics of speaking (Edwards and Potter, 1992: 47). Again, this is a refocussing on speech acts as acts which are designed to achieve something in the context in which they are uttered. Not only is Neisser's explanation asymmetrical with respect to accurate and inaccurate memory, his account of Dean's errors is too rooted in Dean's individual dispositions, rather than his pragmatic aims.

Even Neisser's notions of accurate and inaccurate memory are criticised by Edwards and Potter. They argue for a better account of what counts as getting it right in remembering, based on the rememberer's context, rather than the analyst's categories of truth and error. Describing their recommended reorientation, they say '[i]n effect, we are moving from a view of people struggling to remember with the aid of their mental faculties to a view of people struggling with one another in their talk and texts over the real nature of events' (Edwards and Potter, 1992: 57). In other words, the focus is shifted from an individual using their inner cognitive processes to remember, to groups of people employing discourse to construct memories. As with their other criticisms, this advocates a stronger role for context than ecological cognitive psychology seems to have room for, given the latter's focus on the inner cognitive states.

Explaining why they chose to write about Neisser's work on Dean, Edwards and Potter say it is 'as close a study as we can find to the kinds of materials, concerns and methods of discourse analysis, while still retaining an explanatory base in the cognitive workings of mind.' (Edwards and Potter, 1992: 32–33). The way they approach their own case study – Nigel Lawson's gaffe – therefore really highlights the disagreements they have with Neisser, because the cases are so similar in other respects.

Like Neisser's, their study is focussed on a political scandal, but one on a much smaller scale. In 1988, the British Chancellor of the Exchequer, Nigel Lawson, attended a press briefing. The type of briefing is described by Edwards and Potter as 'universally characterized as a regular event, one of a series of "off-the-record" briefings in which senior politicians are able to "float ideas in the press": forthcoming policies, plans and so forth, to which they do not yet want publicly to commit themselves.' (Edwards and Potter, 1992: 58).

At this briefing, Lawson was alleged to have proposed a policy of means-testing pensioners for benefits. Although some would be better off on this system, others would be worse off. Subsequent headlines were not favourable. For example, the Mirror's article from November 7th (three days after the briefing) is headlined "Fury at Tory Blow to Old Folk". It talks about a scheme that 'would bring financial misery to millions and a further serious erosion of the welfare state' and describes the way the plans were revealed at the briefing as "underhand" (via ukpressonline).

In the face of this media furore, Lawson initially denied that he had said any such thing,

> claiming at one point that the journalists had got together and their "fevered imagination" had produced a "farrago of invention", "inaccurate, half-baked" accounts which "bear no relation whatever to what I said" (quotations from *The Times* and the *Guardian*, Tuesday 8 November). (Edwards and Potter, 1992: 58).

As the media storm continued, with the journalists vigorously denying his accusations, he backed down and admitted that he had talked about the issue, but had been misunderstood; the actual policy would involve some being better off, but none being worse off. This policy was eventually implemented at great expense, resulting in accusations that the government was covering up for Lawson's gaffe (e.g. 'Extra

200 million pounds for poorest pensioners; Surprise announcement defuses Lawson row' *The Times*, 25 Nov 1988; '200 million pounds windfall for old: Extra benefits for 2.6 million pensioners rushed through to cover Lawson's means test gaffe' *Guardian*, 25 Nov 1988, quoted in Edwards and Potter, 1992: 59).

Edwards and Potter's study focusses on the act of remembering what happened at the lobby briefing, via the newspaper reports and the official parliamentary record of debates, *Hansard*. Through this written discourse, the participants construct a memory of the briefing. In contrast to the Watergate testimony, there was no tape recording of the meeting. There should have been, but something went wrong with the tape, a fact which led to its own chain of speculation about what the problem could have been and whether it was accidental. This absence of a record for comparison is just the situation we most often find ourselves in in ordinary life, and this is one of the discursive psychologists' important points. Although ecological cognitive psychology attempted to move away from the laboratory and its method of recalling lists set by the experimenter, it is still tied to the comparison of memory with "what really happened".

It is their disagreement over the role of context that makes these cases a good place to look for situated cognition concepts. In terms of my project, the debate constitutes a disagreement over MEMORY. In the next section, I will look in more detail at the concepts in use in each case.

Concepts of memory

Neisser

Ecological cognitive psychology is premised on the idea that cognition in everyday contexts may be different to cognition in the laboratory. Laboratory experiments are said to lack "ecological validity". This means that the results found by traditional experimental psychology may not translate back into real contexts, which are the very things they are supposed to illuminate. To put this another way, the context in which the person is embedded makes a crucial difference. This is exactly what is argued for by embedded cognition theorists.

The role given to the environment is the main thing that changes in the move from traditional to ecological approaches in cognitive psychology. It is not just that being in their typical environment allows subjects to remember better than they would in the laboratory; MEMORY and the other cognitive states or processes are conceptualised as embedded in a context, and this has implications for how they should be studied.

Robert Rupert, a proponent of embedded cognition, says '[a]ccording to HEMC [the Hypothesis of EMbedded Cognition], the human cognitive system does not extend beyond the boundary of the organism, although during cognitive processing, the human exploits environmental objects and structures in surprising and extensive ways.' (Rupert, 2009). This seems to be the assumption behind ecological cognitive psychology. If the cognitive system depends on the environment in important ways, normal cognition should not be studied outside the environment in which it normally takes place. On a more traditional brainbound approach, the situation shouldn't make a relevant difference to the cognition that takes place, so laboratory studies should provide results that are typical of cognition in any other context.

Eugene Winograd sees both the ecological and cognitive approaches as descendants of functionalism, marking a move away from the tradition of experiments on recall of nonsense syllables that began with Ebbinghaus. He discusses this in his introduction to a collection of papers he co-edited with Neisser, *Remembering Reconsidered*. Here, "functionalism" refers to the function *in a particular environment*, as he says '… early functionalism, that is, the importance of adaptation to the environment' (Neisser and Winograd, 1988: 16). Note that this is not the same thing as functionalism in modern philosophy of mind. Functionalism of the environmentally-embedded variety was increasingly popular in the 1930s with several figures who rejected behaviourism, including Bartlett, Dewey, Mead, Vygotsky, Baldwin, Bergson, von Uexküll, (see Wagoner, 2013: 555–556). In this kind of functionalism, the environment is given the important causal role it has in embedded cognition theory, although not the constitutive role it has in fully extended cognition perspectives.

Neisser, with his ecological approach, can therefore be seen as at the forefront of a move from brainbound cognitive concepts to embedded ones. (Recall that it is implicit concepts I am interested in here, so it does not matter that Neisser would not put his own work in these terms.)

As further support for this claim, I will briefly consider how Neisser sees external memory aids like notebooks and knots in handkerchiefs. In *Memory Observed* (Neisser, 1982), the collection of papers edited by Neisser in which the Dean case study is reprinted, external memory aids are of particular relevance in papers by Harris, Kreutzer et al., and Istomina. In his introduction to Harris's paper (Neisser, 1982: 337), Neisser talks about external memory aids as parallel to internal memory techniques like the method of loci. This is a memory technique which involves picturing the items to be remembered at particular spatial locations, such as rooms in a building, or landmarks along a well-known route. According to Neisser's comparison here, external memory aids like notebooks therefore remain as mere memory *aids* ("mnemonic devices"), rather than playing a constitutive role as they would on an extended cognition view. They are however of crucial importance. In his introduction to Istomina's paper, which concerns experiments on children's memories, Neisser acknowledges that '…at all ages, even the most knowledgeable age, performance on laboratory memory tests is markedly poorer than in the more natural kindergarten setting.' (Neisser, 1982: 350). This shows the environment playing the kind of scaffolding role it plays in modern embedded cognition theories.[54]

Neisser has since been seen by some in the literature as a forerunner of the situated cognition movement, which perhaps provides more evidence for my claim that he was beginning a move in this direction. For example, Dahlbäck et al. say in a recent paper on "distributed remembering": 'The discussion on memory, starting with Neisser's anthology *Memory Observed* (1982), was a forerunner that bolstered the

[54] Neisser even notes that whether we use internal or external resources as memory aids will depend on context (Neisser, 1982: 337). How the choice is made between internal and external resources is something that modern situated cognition theorists have devoted some time to, e.g. Clark (2007) and Rupert's (2004: 31–35) criticisms of him on the subject of Gray et al.'s experimental work (Gray and Fu, 2004; Gray and Veksler, 2005). Here the debate is concerned with the processing cost of accessing the resource, while Neisser suggests instead reliability as a basis for choice.

emergence of the notion of distributed cognition (Salomon 1993).' (Dahlbäck et al., 2013). Their use of the word "distributed" here seems to be a general umbrella term like my use of the term "situated". They cite both Hutchins (proponent of distributed cognition more narrowly construed, see chapter 1), and Clark and Chalmers (proponents of extended cognition). If Neisser was one of the first to begin this move, it is perhaps not surprising that his implicit concept would be embedded, as I have argued, rather than full-blown extended. That cognition is embedded is the weaker claim, so adopting it requires a smaller change from the purely brainbound perspective that is implicit in traditional cognitive psychology.

For Neisser then, memory is still something that takes place in brains, but it depends in interesting and important ways on the environment or context of the rememberer. In other words, his implicit concept of memory is an embedded cognition theorist's concept.

Edwards and Potter

I have said that the discursive psychologists' approach takes context and environment even more seriously than Neisser's ecological approach does. Does this indicate a concept of memory based in one of the "stronger" situated cognition perspectives than embedded cognition? I think it does.

For discursive psychologists, the discourse itself is primary, and discourse is not brainbound. More than this, it is *shared between* participants. Even if only one person writes or speaks the piece of discourse in question, it is part of a process taking place between that person and the intended audience. This is part of the aim of the speaker or writer. This observation suggests a concept of memory that is less centred on a particular individual. In fact, Edwards and Potter specifically discuss the fact that their approach is not the basis for an individual psychology:

> Discursive psychology is concerned with the way psychological entities and processes are constituted in discursive acts, and there are two senses in which these acts are not reducible to individual psychology. First, as we have emphasized in various places above, even where an utterance has an individual speaker, this speaker is not necessarily considered to be in sovereign control of her talk. Simple examples of this occur in the sorts of discourse we have

> concentrated on where news reports are sifted and edited, and television interviews may be scripted and preplanned and in various ways collaborative and collective products. At a more profound level, post-structuralists have produced searching critiques of the very notion of sovereign control in this way (cf. Sampson, 1988).
>
> Secondly, when the psychology of agents or entities is constructed in talk and texts, these things are not necessarily correlated with the unitary subjects who form the basis of much of psychology. We have been concerned with the construction of a whole variety of actors, sub-agents and collectives. Motivation, for example, may be attributed to some sort of sub-system of self ('a part of me wants to get really angry with you'), to a more or less standard individual ('Shelley wants an ice cream') or to a wide range of collectives ('the Kurds have been wanting autonomy for decades'; 'children from broken homes are looking for security'). For these reasons, we have seen it as very important to eschew the image of the single individual accounting for the actions of another individual; in some ways this most common of psychological paradigms is the least interesting to study. (Edwards and Potter, 1992: 171–172).

Something like Hutchins' distributed cognition framework is a good fit with this non-individualist approach. Recall that Hutchins characterises distributed cognition as follows:

> [C]ognitive processes may be distributed across the members of a social group, cognitive processes may be distributed in the sense that the operation of the cognitive system involves coordination between internal and external (material or environmental) structure, and processes may be distributed through time in such a way that the products of earlier events can transform the nature of later events. (Hutchins (2000: 1–2).

Constructing a memory through discourse involves distribution across the members of the group involved in the discourse, and also coordination with material structures in the case of written discourse (the journalists' notebooks, the newspaper reports, and the record in Hansard in this case). In a case like the Lawson scandal, there is also distribution through time; later pieces of discourse are responses to earlier ones, and this is part of negotiating the memory of the event.

The only other kind of situated cognition perspective that can involve the spread of cognition across multiple people is extended cognition, so this possibility is also worth considering. Although Edwards and Potter's work could be seen from an extended cognition perspective, it would diverge somewhat from Clark's proposal. Hutchins, in a review of Clark's book *Supersizing the Mind*, proposes

> the hypothesis of enculturated cognition: The ecological assemblies of human cognition make pervasive use of cultural products. They are always initially, and often subsequently, assembled on the spot in ongoing cultural practices. With *Supersizing the Mind*, Clark has delivered the sciences of mind to a prospect from which the field can turn from the tunnel vision of brainbound thinking to the panorama of the enculturated Supersized Mind. (Hutchins, 2011: 445).

This proposal is briefly made in a short review, but it seems to offer a way of construing the extended approach that is much closer to the distributed view in its being non-individualist. It is proposed as an alternative to what Clark describes as the "Hypothesis of Organism-Centred Cognition" (Clark, 2008: 139). Hutchins is directly challenging the individualism of the Organism-Centred approach. If his enculturated version of extended cognition could be fleshed out, that too could be a good framework for describing the discursive psychological concept of memory.

More generally, the discursive psychology literature at times sounds as though it is arguing for parity or impartiality in considering inner and outer elements of putative extended cognitive systems, much as the extended cognition literature does. For example, Smith, Harré and Langenhove suggest that we 'keep in mind the important observation that remembering is a task for people and that the memory "machines" in their heads are of no more and no less significance than the tape-recorders and diaries they also use' (Smith, Harré and Langenhove, 1995: 156). Similarly, Harré and Gillett discuss the idea of the brain as a custom-made instrument (Harré and Gillett, 1994: 96). These examples refer to tools or artefacts rather than other people, suggesting that discursive psychology may be compatible with these versions of extended cognition as well as social versions.

I leave here these tentative suggestions of a type of extended cognition approach that fits with the discursive psychological concept of memory. The view would not be the extended cognition view that is currently prominent in the literature, and would need to be fleshed out considerably – a task beyond the scope of my project. We already have one framework that fits Edwards and Potter's implicit concept of memory well, namely the distributed cognition approach.

For Edwards and Potter, memory is something constructed between participants in a discourse, as part of an occasioned discursive achievement. It is *distributed* between the participants in this act of discourse.

Summary

In summary, Neisser has an implicit embedded cognition concept of memory, and Edwards and Potter have an implicit distributed cognition concept of memory.

Identifying and assessing the epistemic niches

Neisser

As outlined in previous chapters, the epistemic niche of a concept is constituted by the surrounding parts of the framework which apply pressure to that concept, defining the role it is needed to play. Here I will examine some of the most important pressures on MEMORY for this case study, beginning with Neisser.

Rejection of the dominant paradigm

One important pressure on Neisser's concept of memory was for it to play its part in the reaction against traditional cognitive psychology and in setting up ecological cognitive psychology. Being part of cognitive psychology creates a pressure towards a concept of memory as an internal state of an organism, but setting up the ecological version creates pressure towards this internal state having a strong dependence on the environment, i.e. towards an embedded cognition concept. This pressure comes in part from the experimental methods employed to investigate memory. These include a move away from Ebbinghausian recall of nonsense syllables, and eventually a move away from the laboratory altogether.

This pressure created by reacting against traditional cognitivism and setting up the ecological approach is particularly important at the time of the John Dean case study. Neisser talks about the importance of the "low road" in memory research, and the papers collected in his *Memory Observed* (1982) are intended to be examples of this. In the preface, he says that 'those on the low road want to understand the specific manifestations of memory in ordinary human experience' as opposed to those on the high road who 'hope to find basic mental mechanisms that can be demonstrated in well-controlled experiments' (Neisser, 1982: xi). His paper on John Dean's memory is printed in this collection, and thus is an example of how to travel the low road, and part of his support for the claim that people should take this route. Neisser explicitly acknowledged later that this had been an aim of his: 'The principal goals of *Memory Observed* were to illustrate the possibilities of naturalistic memory research and encourage people to do more of it.' (Neisser and Winograd, 1988: 2).

To decide whether this factor is internal to the science, it is helpful to look from Neisser's point of view at how the shift to the ecological approach came about. First it is worth saying something briefly about the rise of cognitive psychology and the decline of behaviourism, because Neisser was a key figure in this change in the years before writing the John Dean case study (see especially Neisser, 1967), and early prefigurings of his ecological approach can already be found here.

According to Winograd, as I said above, both the cognitive approach and its ecological version are descendants of the anti-behaviourist approach of functionalism, in the sense of being interested in function *in a particular environment*. Whilst the functionalists had some influence on Neisser, he himself attributes the collapse of behaviourism to ethology:

> The fundamental blow was struck by a small group of scientists who called themselves "ethologists", and who were not concerned with learning theory at all…They were not so much interested in hypotheses as in the animals themselves…The work of the ethologists showed that the concepts and methods of learning theory were simply irrelevant to the understanding of natural behaviour.' (Neisser, 1982: 10).

The idea of natural contexts that became so important for the ecological approach was therefore there right from the start of the cognitivist movement for Neisser. That influence only grew stronger, and by the time of his keynote address at the first "Practical Aspects of Memory" Conference in 1978, Neisser was advocating its application specifically to the study of memory from an ecological perspective. The address is printed as the opening chapter of *Memory Observed*, entitled "Memory:

What are the Important Questions?" In his introduction to this chapter as editor, Neisser says that

> The orthodox psychology of memory has very little to show for a hundred years of effort, perhaps because it has always avoided the interesting issues. Just as the naturalistic, ethological study of animal behaviour has proved to be more rewarding than traditional research on "learning", so a naturalistic study of memory may be more productive than its laboratory counterpart' (Neisser, 1982: 3).

In the address itself, Neisser memorably said 'If X is an interesting or socially important aspect of memory, then psychologists have hardly ever studied X.' (Neisser, 1982: 4). *Memory Observed*, including the paper on John Dean's memory, was supposed to be a collection of papers dedicated to the underexplored but interesting and socially important work of ecological memory research.

Did Neisser's transition from traditional cognitive psychology to ecological cognitive psychology result in successful science?[55] Laboratory-based work was and is fruitful research; it has not been replaced by ecological work. Nonetheless, I would argue that ecological cognitive psychology grew out of legitimate concerns with the laboratory-based approach, and important questions that it was ignoring, or was not equipped to investigate. Laboratory work was not shown to be useless, or the assumptions behind it completely incoherent, but something else was needed, and the ecological approach seemed to fit the bill.

Judging from the Neisser quotation above, it may look like it was imposed externally from ethology, but for this to be the case, it would have to have been relied upon without question, and continued to be relied upon without its demonstrating any record of success. When Neisser was writing about John Dean's memory, ecological psychology was simply too new for this to have been true. The ecological approach was more of an untried hypothesis at the beginning of its career than an external imposition from outside the science. It had not yet demonstrated the history of

[55] Another interesting question is whether the transition from behaviourism to cognitivism resulted in successful science. I will not address that question here because it would require a wider-ranging history than I have space for. I will take it that the cognitive approach was established, so rejecting behaviourism was not a factor in the epistemic niche for Neisser's concept by the time of the John Dean case study.

success required for internalisation, but nonetheless was rational to employ as a hypothesis.

Is there any evidence to suggest that Neisser thought of it in these terms, or was he more evangelical? In an obituary article about him, Ira Hyman recalls:

> In the late 1980s, ecological memory research in general, and Neisser's argument in particular, came under fire. I asked him if he had ever regretted his strongly worded assault on traditional laboratory memory research. He stated that he was right when he said it, and that the field had needed the push. Neisser was always proud that by championing the cause of ecological memory research, he helped open the field to a greater variety of research methods and questions. (Hyman, 2012).

There might be a worry that Neisser was partly motivated by pushing the field in a new direction just for the sake of its novelty. Later in the same piece, Hyman says that

> [d]uring his career, Neisser was awarded a long list of honors, and he occasionally found himself in the center of broad movements. Neisser, however, always thought of himself as an outsider challenging psychology to move forward. He worked to create an alternative to behaviorism. He then tried to make sure that cognitive psychology was concerned with meaningful problems. (Hyman, 2012).

This notion of Neisser as an outsider challenging one dominant view after another (behaviourism, then traditional cognitive psychology) perhaps indicates a desire to challenge orthodoxy for the sake of challenging it. This wouldn't be an unusual attitude. Many want to make their mark on a field by being at the forefront of change. But in this case there were good reasons to seek change, and other evidence suggests that Neisser's attitude to ecological psychology was more tentative.

For example, in his introduction to Cole et al.'s chapter in *Memory Observed*, he says

> …they [the authors] believe that "ecological validity" may be an impossible goal. For my part, I think it is still too early to decide such questions. *The study of memory in natural settings may or may not yield further insights; time will tell.* Meanwhile, we must follow the

> example of this study by looking at those settings as carefully as possible. (Neisser, 1982: 367).

This is just what I have described as the proper attitude to an untried hypothesis that is being newly tested in a particular field. The reaction against traditional cognitivism doesn't seem to have been an external imposition.

In fact, in his later writings, Neisser seemed to become less antagonistic to the laboratory-based approach. His being open-minded enough for this change in view to take place offers further support for the claim that the pressure to set up a new discipline was not an unquestioned external factor shaping his concept of memory. As an example of his changing view, in *Remembering Reconsidered*, he says:

> It is no longer enough to denounce the old laboratory methods and call for more ecologically oriented studies; we have now to examine the findings of those studies and try to understand them. And in doing so we must not make the mistake of supposing that the "traditional" psychology of memory has simply been standing still, waiting for the ecological approach to come along. Since the mid-1970s, the laboratory-based study of remembering has undergone what amounts to a revolution of its own. Tulving's distinction between semantic and episodic memory, the postulation of "schemata" for everything from stories to selves, the research on scripts and even representations, the rush of new findings on memory development in children – all these are signs of renewed vigor and creativity in the field that I criticized so sharply a decade ago. In chapter 14 of this book, I try to interpret these new developments in ecological terms: In effect, they expand our definition of what kinds of real things exist to be remembered. Partly for that reason, I believe that future relations between ecological and traditional studies are more likely to be complementary than antagonistic. In any case, it seems that the time may be ripe for another reconsideration of remembering. (Neisser and Winograd, 1988: 2–3).

His co-editor Winograd's introductory chapter of *Remembering Reconsidered* is also about the convergence of ecological and traditional approaches.

It seems that the pressure on the concept of memory created by trying to set up ecological cognitive psychology as a new approach is not problematically external. Neisser treats it in the tentative and exploratory manner that one should treat new approaches, and there were good reasons that a new approach was needed. This factor in the epistemic niche was therefore an untried hypothesis at the time of this case study. As such, to the extent that MEMORY was shaped by it, the concept was rational, but non-legitimate because ecological cognitive psychology was too new to have yet demonstrated a history of success.

Having considered the pressure to be novel, I now want to consider another part of the epistemic niche here: the pressure for continuity. In his early work, Neisser worked within the dominant paradigms, first behaviourism, and later cognitive psychology. The pressure for continuity with his own earlier work within these paradigms strengthened a pressure for continuity with his contemporaries who still worked within them. This is a pressure to which even revolutionaries are not immune. In addition, the popular research frameworks of the day dictate the problems, concepts etc. available, which even someone reacting against them must work in terms of to a large extent. Here I will discuss two aspects of this pressure towards continuity: the influence of Bartlett, and the computer metaphor. What will ultimately be of interest is Neisser's attempt to combine these two influences in his earlier work (Neisser, 1967).

Bartlett

The most important aspect of Bartlett's influence is his notion of a schema (although he has some importance in other ways, and I will mention him again in the "constructed memory" subsection below). Bartlett does not give a completely clear account of schemata in his work, and in fact disliked the term (see Wagoner, 2013).[56] In his (1932), Bartlett says:

> 'Schema' refers to an active organization of past reactions, or of past experiences, which must always be supposed to be operating in any well-adapted organic response. That is, whenever there is any order or regularity of behaviour, a particular response is possible only because it is related to other similar responses which have been serially organised, yet which operate, not simply as individual members coming one after another, but as a unitary mass. … All incoming impulses of a certain kind, or mode, go together to build up an active, organised setting. (Bartlett, 1932: 201).

[56] See Wagoner for the history leading up to Bartlett's use of the term "schema", the factors shaping his concept, and how it has changed in modern cognitive psychology. Wagoner's analysis is very much in keeping with the kind of project I am carrying out here.

The notion was part of his attempt to set up a rival to the trace theory of memory. Wagoner says that '[f]or him, schema was to provide the basis for a theory of remembering that was embodied, dynamic, temporal, holistic, and social.' (Wagoner, 2013: 553).

It can be seen here that Bartlett was an early proponent of the real world rather than laboratory-based approach, even as early as the 1930s. He worked in social psychology, with an interest in anthropology amongst other things, contributing to an approach not dissimilar from that of the ethologists in their research on animals. It is therefore perhaps not surprising that Bartlett is often hailed as one of the fathers of cognitive psychology, and part of its overthrow of behaviourism.

He also criticised Ebbinghaus (see e.g. Bartlett, 1932: 2–7), and worked on memory outside the laboratory, in a way that resembles that advocated by modern situated cognition proponents. For example, he studied memory for playing cards during a game of bridge, versus when they are presented randomly in an Ebbinghaus-style laboratory experiment. As Wagoner says:

> In bridge there will be an active interest in remembering the cards to meet the needs of the game... Bartlett is emphatic that it is the former ability to remember, as part of a whole living social activity, that is needed for general functioning in everyday life. (Wagoner, 2013: 560).

This embeddedness in environmental and social context is accompanied by embodiment of the organism for Bartlett. Here is Wagoner again:

> Despite its sometimes vague and sketchy formulation, it is clear that Bartlett emphatically rejects the trace theory of remembering and wants to replace it with one in which the whole active organism takes central place. "Attitude," "schemata," and "image" are all organismic concepts, for Bartlett, which are implicated in a person's dynamic relating to the world. They are functions coordinated within a total system, which must make a unitary response in its environment. (Wagoner, 2013: 562).

It seems from this that the influence of Bartlett may have been a pressure on Neisser's concept of memory, pushing towards an embedded cognition approach.

(Recall that in chapter 1, I defined embedded cognition such that it includes dependence on the extra-cranial body, as well as the extra-bodily environment).

However, the treatment of the bodily and external components in Bartlett is not that of a modern embedded cognition theorist. His work is closer to the radical enactivist perspective, in his focus on “dynamic relating to the world”, and his holistic rather than representationalist approach. This shows up particularly in a passage in Wagoner’s paper, discussing Bartlett’s dissatisfaction with the term “schema”, and some alternatives he uses elsewhere in his work. Wagoner says that

> …by 1932 Bartlett was growing dissatisfied with the word “pattern.” He ultimately prefers the term “organised setting,” which better highlights that schemata operate at the developing transaction between organism and environment, rather than being a purely cognitive phenomenon (i.e., a mental representation). (Wagoner, 2013: 559).

Wagoner also says:

> In contrast to the trace theory, which treats memory as an isolated mental faculty, Bartlett starts with a whole organism actively involved with its environment. The mind is taken “out of the head” and situated in the ongoing transactions between a person and his or her environment. [4] (Wagoner, 2013: 555).

This is very much like the enacted cognition approach. Footnote 4 reads: ‘The contemporary equivalent of this position is ecological psychology, which follows the work of J. J. Gibson.’ As we will see in the section on the Gibsons below, J. J. Gibson is a major figure in the history of modern enactivism. The influence of Bartlett’s work was therefore applying pressure towards an enacted cognition concept.

Neisser’s use of Bartlett involved both his ideas and his experimental results (see especially Neisser, 1967). By the time Neisser was working, Bartlett’s ideas had already been taken up widely in psychology. He was hugely influential, becoming a Fellow of the Royal Society in 1932, and being director of the Unit for Research in Applied Psychology in Cambridge from 1944–1951. His work had practical applications, leading to him being knighted in 1948 for work in applied psychology with the Royal Air Force during the Second World War. His experimental results, such as those on constructed remembering which I will discuss below, and his notion of schema, were fruitful in pointing the discipline in new directions. His influence had therefore been internalised into the discipline by the time Neisser was writing.

However, the enactivist Bartlett I have described here, with his dynamic, contextually-embedded and socially-oriented schemata, was not Neisser's reconstruction. Therefore, although the influence of Bartlett was an internal consideration, it seems that it was moderated, perhaps even distorted, by some other influence. The major candidate for this other influence is the computer metaphor.

The computer metaphor

Wagoner says: 'In an early foundational book of cognitive psychology, Neisser (1967) drew heavily on Bartlett's (1932, 1958) work, but did so very selectively in order to fit Bartlett's ideas into the computer metaphor of mind.' (Wagoner, 2013: 565). In effect, Neisser makes Bartlett's schemas equivalent to programs, i.e. software. Later in the same paper, Wagoner says that '[s]chema was transformed into a static knowledge structure composed of different slots or nodes that either accept incoming information or fill in default values where input is lacking. Schema is here severed from an organism's functioning in the world.' (Wagoner, 2013: 569). The pressure on MEMORY from the computer metaphor by the time of the John Dean case study is therefore in part a pressure towards continuity with Neisser's own earlier work. We can see from the Wagoner quotation that the computer metaphor creates pressure towards a brainbound concept of memory.[57]

There was already precedent for this reading of schemata in some work done in Bartlett's own laboratory by one of his students. Oldfield (1954) argued that using modern computers as a model of memory could make more sense of Bartlett's experimental results than older static storage models based on metaphors such as the

[57] Brainbound concepts are not a necessary feature of the computational approach. Clark and Chalmers' (1998) extended cognition can be given a functionalist (in the modern sense) reading, and this version of the position is arguably enabled by the computer metaphor. Roughly the idea is that there is nothing sacred about neural "hardware" because anything (including parts of the outside world) could implement the relevant "software". However, the version of the computer model in play when Neisser was writing was a brainbound one.

photograph or gramophone record. This was the beginning of a more representational, atomist reading of Bartlett's schemata (see Wagoner, 2013: 565).

This view of schemata had been widely taken up in cognitive science by the time Neisser was writing the John Dean paper: 'In the 1970s, a number of new concepts in cognitive psychology were explicitly derived from Bartlett's schema, including Minsky's (1975) frames, Shank and Abelson's (1977) scripts, and Mandler and Johnson's (1977) story grammar' (Wagoner, 2013: 566). The departure from Bartlett's conception in these versions is stark according to Wagoner: 'In short, these early cognitive schema theories are spatial (not temporal), static (not developing), focused on elements or nodes (not holistic), passive (not active), individual (not social), and structural (not functional).' (Wagoner, 2013: 567). This information processing view of schemata has been with us ever since. According to online encyclopaedia "PsychCentral", modern cognitive psychology defines schema as 'A mental structure that represents an aspect of the world, and streamlines information processing by categorizing objects.' (Fournier, 2009).

We can see here the tension between the pressure towards an enacted cognition concept from Bartlett's influence, and the pressure towards a brainbound concept from the influence of the computer metaphor. What should be done in this situation depends on the status of the two considerations. I have argued that the influence of Bartlett was an internal factor, but what about the computer metaphor?

In the last case study, I argued that using science fiction scenarios was external to science. Such scenarios seem to be metaphors, and I argued that they should not be allowed to apply pressure to scientific concepts. However, many metaphors for memory have resulted in some success e.g. the written word, the calculator, the photograph, the telephone exchange (see Draaisma, 2000). The computer metaphor is just the latest in this line of technological metaphors for mind.[58] This type of metaphor seems different from the science fiction type; I suggest that science fiction

[58] Perhaps our new metaphor for the future is the human + computer integrated system (cyborg! see Clark, 2003) and this will apply more pressure towards situated cognition concepts. Such a metaphor would help the science to be relevant to everyday life (smart homes, wearable tech etc). I leave this suggestion as speculation here; we cannot anticipate future science.

cases are not true scientific metaphors because, unlike the computer and the telephone exchange, they are not *models* for the mind. They are therefore not a formal part of the science. Modelling is a scientific technique that has been internalised because it has been very fruitful in many scientific disciplines.

At one stage, the use of metaphorical or figurative language was considered to be very bad form in science. In 1666, one member of the Royal Society, Samuel Parker, advocated banning it altogether. Draaisma quotes him as saying that metaphors' 'wanton and luxuriant fancies climbing up into the Bed of Reason, do not only defile it by unchast and illegitimate Embraces, but instead of real conceptions and notices of Things impregnate the mind with nothing but Ayerie and Subventaneous Phantasmes' (S. Parker, 1666, quoted in Draaisma, 2000: 55). This (highly metaphorical) criticism of metaphor was far from unique at the time, with others such as Thomas Sprat and Francis Bacon also advocating simple language, free from metaphor and imagery, as best for science (Draaisma, 2000: 54).

In 1682, Robert Hooke delivered a lecture to the Royal Society, which was later published as *An Hypothetical Explication of Memory: how the Organs made use of by the Mind in its Operation may be Mechanically understood*. The lecture was full of metaphorical language. Hooke argued that this language was indispensable. He was presenting a materialist theory, and the metaphors allowed him to talk about the *location* of memory, among other physical properties (Draaisma, 2000: 62). He is quoted as saying: 'It is not, I conceive, possible to be truly understood or described, but only by Similitude' (Hooke, 1682, quoted in Draaisma, 2000: 63).

To this day, metaphors have retained this indispensable quality in the sciences of memory. We often don't have a literal way of saying what we say metaphorically. Draaisma says:

> Whereas in the case of physical processes like the interaction between immune cells and pathogens one can form some kind of idea about a literal description, the literal description of *mental* processes seems to be fundamentally excluded. What is the literal equivalent of "search processes" in the memory? How do you literally describe a process such as "storing"? If "filtering of information" is a metaphor, what literal description does it replace?

> The problem with figurative usage in psychology is that no literal alternative is available (Draaisma, 2000: 11).

Not only can we not do without metaphors, they seem to have positive virtues. Hooke suggested that metaphors and figurative language were essential for understanding memory processes, and also provided a familiar and intelligible basis for formulating hypotheses (Draaisma, 2000: 63–64). These features are still important today. Draaisma talks about two types of heuristic values of metaphors, theoretical and empirical:

> *Theoretical* heuristics means that a metaphor introduces new theoretical notions, brings coherence to hypothetical processes or is able to resolve apparent contradictions between experimental results, while *empirical* heuristics describes the degree to which a metaphor produced new topics for research. (Draaisma, 2000: 18).

The use of metaphors therefore has positive benefits, which have contributed to a history of success and its internalisation into science. In psychology, it may well even be indispensable because we lack literal means of saying the same things.

The computer metaphor in particular was also starting to lead to success when Neisser was writing. It was an integral part of the birth of cognitivism, and as I have said already, Neisser's own earlier work was an important part of this.

Both the influence of Bartlett and the computer metaphor were internal, but there was some tension between them. The pressure from the computer metaphor is towards a representationalist concept, so the enactivist tendencies coming from Bartlett's influence are lost. The upshot of this is an embedded cognition concept, where the environment is important, hence Neisser's ecological approach, but cognition is still construed in terms of inner representations.

I have argued (see chapters 2 and 5) that, when faced with two fruitful frameworks, a science should remain pluralist. The same is true here – the appropriate strategy when faced with conflicting internal pressures would have been a pluralism of concepts. However, faced with Bartlett's legacy, and the popularity of the computer metaphor, Neisser tried to integrate them. Attempting to integrate rival strands can be a part of the pluralist strategy, as discussed in chapter 2 with reference to modern integrative pluralism (e.g. Mitchell, 2002, 2003). However, this relies on a certain amount of compatibility between the rival stands. In this case this is lacking, so much of what is important in Bartlett – what led to his influence being internalised – was lost. Pursuing both the Bartlett-influenced enactivist path, and the more representational brainbound computer metaphor, would have been a better strategy. This would have hopefully prevented the computer metaphor from becoming too dominant.

Part of the problem here may have been pressure towards unification rather than pluralism because it is seen as a virtue in its own right, something which I have already argued (see chapter 3) is an external factor in the cognitive and social sciences.

In his work on metaphor, Draaisma gives a more specific account of why unification may have been favoured in this case (although he does not put his discussion in terms of unification and pluralism). A metaphor causes research to focus on some aspects of a phenomenon while neglecting others; Draaisma talks about the meta-metaphor of a filter. If something is viewed through a filter, as with a metaphor, some parts are highlighted while others are obscured. This led Freud to recommend changing metaphors regularly to avoid this problem (Draaisma, 2000: 18–20).

The problem is that the computer metaphor for the mind (and other metaphors for the mind, e.g. the hologram) are so extensive, they cannot easily be alternated for other metaphors in this way. Draaisma suggests the term "metaphoric theme" rather than metaphor. He says of these metaphoric themes that '[t]hey not only furnish metaphorical terms for separate functions, they also provide a background against which all those separate metaphors have meaning. The interpretation of specific computational and holographic metaphors presupposes the metaphor theme of which they are a part.' (Draaisma, 2000: 20). Therefore although the computer metaphor was internal for Neisser, it is an internal factor of a kind that particularly resists the kind of pluralism I have suggested would have been rational to adopt here.

In fact, we can see that Neisser was becoming sensitive to something like this problem by 1976, when he began to express doubts about the dominance of the computer metaphor. In his *Cognition and Reality*, he says:

> If cognitive psychology commits itself too thoroughly to this model [the computer model], there may be trouble ahead. Lacking in ecological validity, indifferent to culture, even missing some of the main features of perception and memory as they occur in ordinary life, such a psychology could become a narrow and uninteresting specialized field. There are already indications that this may be happening. (Neisser, 1976: 7).

This was the beginning of Neisser advocating the ecological approach, largely under the influence of J. J. Gibson and his wife Eleanor. This influence loomed large by the time of the John Dean case study, to some extent working against the pressure from the computer metaphor, and the version of schemata based on computationalism (see Wagoner, 2013: 572, note 11). The Gibsons' influence will be the subject of the next subsection.

The Gibsons

James J. Gibson and his wife Eleanor Gibson were friends of Neisser's, and he acknowledged their influence by dedicating his 1976 book *Cognition and Reality* to them.

J. J. Gibson saw memory as tied up with perception, and for him, perception was direct. He advocated *direct realism* – the view that an organism can directly and veridically perceive the external world without intervening representations. The role of that world is of crucial importance for Gibson, and it is he who describes his approach as "ecological" (See e.g. "Ecological Optics" from 1961, in Reed and Jones, 1982: 61–75). This way of thinking culminated in his 1979 book *The Ecological Approach to Visual Perception*.

Of perception and memory, Gibson says:

> The objective operations do not distinguish memory from perception. Only our subjective feeling about them separates the two kinds of activity. We have the feeling that perception is confined to the present, whereas memory refers to the past. But this distinction, be it noted, is

> wholly introspective. Moreover, as will be evident later, it cannot be made with clarity. (Gibson, 1966, in Reed and Jones, 1982: 172).

The reason the distinction cannot be clearly made is that Gibson views time as equivalent to space. Looking around through space is therefore importantly similar to "looking around" through time, i.e. remembering. He says that '[i]t is no harder for a brain to integrate a temporal arrangement than a spatial arrangement' (Gibson, 1966, in Reed and Jones, 1982: 174). Much of this work is based on experiments on perceiving motion. The view is therefore dynamic, not like the static storage and retrieval view encouraged by the computer model.

In "The Useful Dimensions of Sensitivity" from 1963, Gibson says:

> Information about the world that has been obtained will continue to be obtained, and the information-pick-up will improve with practice. This is perceptual learning. But this does not imply that information is stored in memory. The information continues to be available outside the skin, i.e. the invariants that specify the world. Perception is a skill, not a constructing of the mental world out of psychic components. The observer has no *need* to store information. The fact that he can recall, recollect, imagine, and think about parts of the environment "not present to the senses" is a different matter entirely. This fact does *not* prove that memories are combined with sensations so as to yield perceptions. (Gibson, 1963, in Reed and Jones, 1982: 372).

Static storage of representations is replaced by skill at retrieving information from the world. This is the kind of "knowledge how" approach to cognition typical of the enactivist position (e.g. Noë, 2004: 11, 106).[59]

In her foreword to a collection of J. J. Gibson's papers, his wife Eleanor Gibson says:

> He gave up the notion of the "retinal image" as being the basis of visual perception and began to think of visual perception as an active process of "looking" (searching for change and invariants), which involved a much enlarged receptive system that is not just receptive.

[59] It should be noted that skill and knowledge-how are also emphasized by Wheeler (2005) in his proposals for a Heideggarian cognitive science that is based on the *embedded* cognition perspective. However, Wheeler still leaves room for representations, making his approach further from Gibson's picture than the enactivist is, as he himself notes (Wheeler, 2005: 301 note 9, 306 note 3).

> Adjustments of the eyes, head and trunk are all involved and the perceiver is not static. He moves around and as he does so the scene is continuously changing, and the flow of stimulation as well.' (Reed and Jones, 1982: xi).

This kind of approach, where movement of the body is an important part of perception, is strongly reminiscent of Alva Noë's enactivist work on perception (Noë, 2004). Although Noë does not endorse all of Gibson's work as it stands, he thinks it can be defended and reconstructed in a way that is amenable to his project (see 2004: 21, 105).

One of Gibson's key contributions is his introduction of the concept of "affordances" into psychology. Here is Eleanor Gibson again:

> Before the book [J. J. Gibson's *The Ecological Approach to Visual Perception*] was finished, a new concept was introduced – the concept of "affordances". A careful description of the information for perception, even as it approaches elegance in the form of a mathematical statement [a reference to classical psychophysics, which Gibson came to see as inadequate], does not convey sufficiently the reciprocity of a creature and the environment, especially its own niche or habitat. This mutuality of creature and environment is the basis of the need for an ecological optics, one that is meaningful for a living creature. The surfaces and substances of this environment provide opportunities of diverse kinds for the creature's activities, offering it support for living successfully in the world. These opportunities are its "affordance" (a made-up word)' (Reed and Jones, 1982: xiii).

Again, the environment is essential for cognition as we know it. The "mutuality of creature and environment" referred to here is at the root of another modern enactivist approach, that of Varela, Thompson and Rosch (1993). They note the compatibility of Gibson's notion of affordances with their approach (Varela, Thompson and Rosch, 1993: 203), however they do emphasize other differences, seeing their enactivism as a middle way between Gibson and representationalist approaches treating the world as pregiven (Varela, Thompson and Rosch, 1993: 202).

Although there are differences between Gibson's work and the modern enactivist perspectives of Noë, and Varela, Thompson and Rosch, it is the closest match among the perspectives on cognition considered here, and Gibson has influenced the enactivists more than any other group. For the purposes of my project, we can therefore see the influence of the Gibsons as applying pressure on Neisser's concept towards an enacted cognition perspective.

Neisser has Gibson's view of memory in mind when he undertakes the John Dean case study, as he explicitly acknowledges in the introduction to that paper in *Memory Observed*:

> J. J. Gibson insisted that the study of perception should begin with veridical seeing rather than illusion and error; maybe the study of memory can benefit from a similar approach. It was partly with this possibility in mind that I undertook the study of John Dean's testimony. What could psychology learn from a case where the witness was right? (Neisser, 1982: 139).

This notion of veridical memory is part of what Edwards and Potter criticise in Neisser's approach, and I will return to this issue below in the subsection on constructed memory. In the present subsection, the important question is whether the pressure from the Gibsons' work towards an enacted cognition concept, as set out above, was an internal factor in shaping Neisser's concept of memory. I have said that the broader framework of ecological cognitive psychology that Neisser was trying to set up was still too new to have a history of success, but the Gibsons' more specific ecological work on vision and learning had been ongoing since the 1950s, so the same need not be said of it.

It was in conflict with the dominant paradigm, which as we have seen was the information processing view of schemata – a paradigm within which Neisser had previously been working (Neisser, 1967). This representationalist, information processing view of schemata was successful, being responsible for most of the results psychology has obtained. J. J. Gibson's view was and is less popular, but still successful. Its influence can still be seen in embodied, enactive, extended and radical embodied cognition theories (e.g. Noë, 2004; Chemero, 2009; Varela, Thompson and Rosch, 1991), Human Computer Interaction, Dynamical Systems Theory, and design, where his concept of affordances (discussed above) is of particular relevance. These disciplines have produced novel practical applications, as well as explanations of phenomena.

At the time Neisser was writing, these successes were already beginning, although primarily in the guise of providing a fruitful new way of thinking about cognition, rather than experimental data (see Goldstein, 1981). Neisser described the Gibsons' work as a "flourishing line of research" in *Cognition and Reality* (Neisser, 1976: 8), and he takes many ideas from their work throughout that book, supporting his claim that much of cognitive psychology was missing out the effects of real contexts on cognition. The influence of the Gibsons therefore had a track-record of success, and so was an internal factor.

Both the influence of the computer metaphor and the influence of the Gibsons were internal, but there was tension between them, with the former applying pressure towards a brainbound concept, and the latter towards enacted cognition. This is the same kind of situation as we found between the influence of Bartlett and the computer metaphor. What Neisser did here, as there, was to combine the two influences, thus responding both to pressures for change, and for continuity with his former work. (See Neisser, 1976, chapter 4, "Schemata", particularly pp.52–53).

Again, as in the previous case, I would suggest that remaining pluralist would have been the better course. There are pressures from internal factors towards maintaining both views, and attempting to integrate them in this case means that something important is lost. Again, the external pressure towards unification for its own sake is probably at fault here (see above, and chapter 3).

1980s intellectual milieu

The next influence on Neisser's concept I want to consider is another set of pressures toward continuity with others. In the 1980s, arguably the changes in many surrounding disciplines were laying the foundations for a situated cognition approach (see Michaelian and Sutton, 2013). Continuity with these surrounding disciplines was an important pressure on Neisser's concept, if one that it is very hard to quantify. It is important to mention this more general intellectual milieu, as well as more specific influences like the Gibsons. I will consider several of his contemporary influences rather sketchily here, and make no pretence that it is a full survey.

The influences in question include the failure of classical AI and the rise of connectionism, alliances with dynamical systems theory within cognitive science, Bruner's notion of scaffolding learning with external resources, and the rediscovery of the ideas of Halbwachs and Vygotsky (Michaelian and Sutton, 2013: 4).

Halbwachs was a French social theorist who worked in the early 20th century. He was most famous for his work on collective memory, proposing that human memory can only function within a collective context (Halbwachs, trans. 1992). Vygotsky was a Soviet developmental psychologist who worked at around the same time. He is known for the influential idea that thought is an internalisation of language. Clearly both collective memory and the internalisation of previously external strategies are ideas related to modern situated cognition perspectives, and these figures have been an influence.

In addition, around the time that Neisser was writing, new disciplines became involved in cognitive science, including anthropology, education, media theory, design, and the emerging discipline of Human Computer Interaction. These influences also encourage a situated approach to cognition where the influence of the environment, other people, and technology are given an important role.

Were these influences internal factors in the epistemic niche for MEMORY? Some are pressures from other disciplines, so it may appear not. However, ideas were not being imported from these disciplines and relied on unquestioningly; they were just influences that were in the air at the time. The rediscovery of the work of Halbwachs and Vygotsky in particular may seem to be external because their influence had waned before this, suggesting their frameworks failed to result in success. However, rediscovery of old ideas can be successful; they may yet be internalised. Again, this is not about taking their ideas as authoritative and relying on them. They are merely being tried out, and old ideas are as reasonable a source for this as new ideas from neighbouring disciplines.

The pressure to take account of these new ideas is just the pressure to maintain a connection to neighbouring disciplines, which is an internal consideration. Note that this is not as strong a consideration as unification, which I argued was external to the cognitive sciences. It is just the weaker claim that it is good to take notice of what is happening in neighbouring disciplines. While taking account of neighbouring disciplines is an internal consideration, the influences themselves are untried hypotheses in this case.

Constructed memory

I now want to consider another idea which applied pressure toward a concept that was embedded in the real world. The idea is that of memory as constructed.[60] Again, Bartlett is a large influence here. This seems to be an aspect of Bartlett's work that Neisser took up right from the start.

Michaelian and Sutton in *Review of Philosophy and Psychology* say:

> 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). (Michaelian and Sutton, 2013: 3–4).

Bartlett's work on constructive remembering is still well-known and well respected, particularly his work on memory for stories, such as "The War of The Ghosts" (Bartlett, 1932: ch V). This story comes from Canadian Indian Folklore. In the experiment, English participants were asked to read the story, and then recall it after various time intervals. Bartlett found that participants tended to change elements of the story – to reconstruct it – to fit better with their own culture. He described this in terms of altering the story to fit with the participants' schemata. (Although as we have already seen, Bartlett's original notion of schema was not the same as Neisser's).

Another influence on Neisser's concept of memory as constructed was the work of Elizabeth Loftus on witness testimony in court. He includes some of her work in a section in *Memory Observed* on "Testifying" (Loftus and Palmer, "Reconstruction of

[60] Later in his career, Neisser became involved with the False Memory Syndrome Foundation. This suggests that he started to believe that the constructed memory notion was becoming too powerful (see also see Neisser and Winograd, 1988: 4), but at the time of the John Dean case study, it was very influential for him.

Automobile Destruction", first printed as Loftus and Palmer, 1974). This section of *Memory Observed* also includes his paper on John Dean.

Both Loftus' and Bartlett's work were widely influential, having practical success through changing how witness testimony is treated in court. Loftus first provided expert testimony on eyewitness identification for Washington state in June 1975. This was unprecedented at the time, but was something she continued to do in subsequent years (Zagorski, 2005). The idea of memory as constructed was internalised in psychology as a result of this kind of success.

There is some dispute over in exactly what sense memory should be seen as constructed. According to Edwards and Potter, the influence of Gibson's work on direct perception and memory changed Neisser's view of construction. Gibsonian direct perception leaves no room for any construction. Edwards and Potter say 'Neisser has been moving from his early cognitive constructivist position (Neisser, 1967) increasingly towards a position where memories are organized as reflections of the true facts, albeit after a process of repisodic synthesis.' (Edwards and Potter, 1992: 56). Recall that repisodic synthesis, an important idea in Neisser's analysis of John Dean, is when the subject has what appears to be an episodic memory of a particular event, but is really a synthesis of the common characteristics of a series of events. Although this analysis is still constructionist in a sense, in the Dean case study it is based on comparison of the memory with a correct version of events (the tape recordings of the meetings).

Here we see the influence of Gibson, which I argued was internal, affecting the notion of construction coming from Bartlett and Loftus, which is also internal. For my purposes here, this conflict is less important to adjudicate than those discussed in the last two subsections, because it is not really a dispute among the different situated cognition perspectives. Pressure towards constructed memory is pressure towards taking the environment seriously, and therefore towards an embedded cognition concept, regardless of exactly what kind of construction is endorsed. A full analysis would have to take into account the recommendations of pluralism I made in the last two subsections, as well as the compatibility or otherwise of Gibsonian direct perception and Bartlett and Loftus's kinds of construction. This would be a large diversion, and I will therefore not consider this dispute further, but as a clash of internal considerations, it may be another place that pluralism should be recommended.

Mnemonists

I want to consider one final influence on Neisser's concept of memory. This is a preoccupation with mnemonists: those who perform impressive feats of memory using special techniques. Neisser prefers the term "memorist" because this includes all those with particularly good memories, whether they use such special tricks or not (see his paper "Memorists" in Neisser, 1982: 377, 379). However, it is mnemonists that are of most interest from a situated cognition perspective. One common technique in particular, the method of loci, is important. In Vygotskian fashion, this technique is about internalising the external world. One memorises a list by imagining each item on the list at a particular location, e.g. along a familiar route. Imagining travelling this route then facilitates recall of the list.

Neisser's *Memory Observed* includes a section called "Special People" which includes his own chapter on "Memorists" as well as chapters from others on mnemonists. The book also includes a section on "Performing" which is about oral traditions of illiterate people, and also contains an impressive collection of memory feats. Neisser acknowledges that some chapters included in one section may have been in the other (e.g. "The Mnemonic Feat of the "Shass Pollak"" by George M. Stratton could have been included in the Special People section, while a chapter on Toscanini could have been included in the chapter on Performing, see Neisser's introductions to these chapters).

In his study of memorists, Neisser is interested in how different contexts affect memory. For example, he talks about what counts as remembering being different in different contexts: 'Literal, verbatim memory does exist…It makes its appearance whenever a performance is *defined* by fidelity to a particular text.' ("Literacy and Memory" in Neisser, 1982: 242). The book includes the extreme example of memorizing the Talmud (in "The Mnemonic Feat of the "Shass Pollak"" by George M. Stratton), where even typographical detail must be remembered. In his editor's introduction to this chapter, Neisser says that '[c]ulturally defined memory performances differ widely in the kinds of fidelity to the original that they require; individual skills of memory adapt themselves to those definitions.' (Neisser, 1982: 311). Here, the relevance of the embeddedness in context so important for the ecological perspective is clear. The context is causally relevant, but nothing external is constitutive of memory. The influence of mnemonists therefore applies pressure towards an embedded cognition perspective.

At first glance, this looks like an external influence. Looking to the experts at a particular cognitive task is not something that has been internalised – one would normally look for typical subjects. There is no reason to think memorists are typical (quite the opposite) and their study hardly seems to be about remembering in a natural context. However, looking more closely at Neisser's approach to studying these people, we can see that he doesn't take for granted the conclusions that should be drawn from their study, but treats his work on them more as an untried hypothesis.

He is very well aware that memorists may not be typical subjects, but suspects they are more common than we think. He does not take even that suspicion for granted, and advocates finding out just how many such people there are, and taxonomising them as a first step (Neisser, 1982: 379).

He points out that their study has been largely neglected by experimental psychologists. He says that '[t]here are many reasons for this neglect. The one on which I wish to focus arises from the prevailing conception of memory itself – a conception that suggests no way in which the study of outstanding individuals could contribute to scientific progress.' (Neisser, 1982: 377). However, to challenge this conception is not to problematically take for granted its opposite.

Neisser's suspicion is that the study of memorists risked falsifying the prevailing hypothesis that we all have more or less the same unalterable hardware, and that different mnemonic techniques are just software. This is why Neisser is interested in memorists who are not mnemonists. He is open to the possibility that these phenomena are not always just the result of special tricks. For Neisser, the prevailing hypothesis was what led people into the laboratory – because laboratory work is a good way to study a universal mechanism (Neisser, 1982: 378). Studying memorists is therefore part of the ecological approach that hopes to lead people out of the laboratory and into the world. It is therefore part of the rejection of the dominant paradigm that I argued was a hypothesis in the first subsection.

Summary

In summary of what I have said about the epistemic niche for Neisser:

Factor applying pressure to concept	**Concept encouraged by this pressure**	**Internal / Hypothesis / External**
Rejection of dominant paradigm	Embedded	Hypothesis
Bartlett	Enacted	Internal
Computer metaphor	Brainbound	Internal
Gibsons	Enacted	Internal
Unification	Embedded (in this case, because unifying brainbound and enacted)	External
1980s milieu	Situated (various)	Hypotheses
Constructed memory	Embedded	Internal
Mnemonists	Embedded	Hypothesis

Earlier in this chapter, I argued that Neisser's implicit concept of memory was an embedded cognition concept, and here I have considered several factors that applied pressure towards this embedded concept. These include reacting against the dominant paradigm (mainstream cognitive psychology) and setting up the new ecological version, the idea of memory as constructed, and the study of memorists or mnemonists. Other influences in the intellectual milieu of the time (failure of classical AI, involvement of new disciplines in cognitive science, rediscovery of Halbwachs and Vygotsky etc.) also applied pressure towards a situated cognition concept of some kind. All of these are untried hypotheses apart from the idea of constructed memory.

The number of hypotheses reflects the newness of ecological psychology, an approach that Neisser was still working to set up at the time of the Dean case study.

Constructed memory had been internalised at a coarser level of grain (into psychology in general) so was not a hypothesis. However, there is some possible tension between it and the Gibsonian perspective, which I mentioned above but largely set aside as outside the scope of this project. Recalling that a concept is legitimate to the extent it is shaped by internal factors, and rational to the extent it is shaped by internal factors and hypotheses, we can see that the embedded cognition concept is rational to the extent that it is shaped by all of the factors mentioned so far. It is legitimate to the extent that it is shaped by the idea of constructed memory, though non-legitimate to the extent it is shaped by the untried hypotheses. The evidence for the embedded cognition concept functioning as an investigative kind is therefore relatively weak when we only consider direct pressures towards an embedded concept. However, that is not the end of the story.

There are also pressures towards other types of concept. The influence of Bartlett applies pressure towards an enacted cognition concept, while the pressure to retain continuity with the dominant paradigm (in particular the computer metaphor) applies pressure towards a brainbound concept. Both of these pressures are internal, and Neisser tries to integrate them, reinterpreting Bartlett's schemata to fit in with computationalism, and resulting in an embedded cognition concept. I have argued that pluralism would have been a better option than integration here, because the two pressures were not compatible enough to integrate without something essential being lost, in this case from Bartlett's approach. The integration is likely to have come about under the influence of the external pressure towards unification for its own sake, and to the extent that this was the case, the resulting concept is non-legitimate and irrational.

The pressure from the computer metaphor was mitigated somewhat by the influence of the Gibsons, another internal pressure towards an enacted cognition concept. Again, Neisser tried to integrate the computational reinterpretation of schemata with the Gibsonian approach, probably as a result of external pressure towards unification for its own sake. Again, I argued that pluralism would have been a better approach. The embedded cognition concept that resulted was legitimate and rational to the extent that it was a result of the influence of the Gibsons, Bartlett, and the computer metaphor, but a single embedded concept was still not the best response to these pressures combined. This case study therefore provides some evidence of an embedded cognition concept functioning as an investigative kind in cognitive psychology, although that evidence is not very strong. It does however provide some evidence that multiple concepts should be in use at once, including another situated cognition concept (in this case enacted). Therefore there is still a reasonable amount of evidence in favour of the situated cognition approaches here.

Edwards and Potter criticise Neisser, and have a different concept of memory. I will now go on to consider the pressures on their concept in a similar manner.

Edwards and Potter

Setting up a new subdisipline

In a similar manner to Neisser, Edwards and Potter are partly motivated by reacting against mainstream cognitive psychology. However they are also reacting against its ecological version. Setting up a new subdiscipline in competition to the dominant approach is part of what their concepts (including MEMORY) are needed to do. For them, the new subdiscipline in question is of course discursive psychology. Setting up this new approach applies pressure towards a distributed cognition concept of memory because, as I said above, constructing a memory through discourse involves distribution across the members of the group involved in the discourse, coordination with material structures (e.g. written discourse), and sometimes also distribution through time because later pieces of discourse are responses to earlier ones.

Discursive psychologists saw their work as leading a “second cognitive revolution”, the first having been the overthrow of behaviourism by the cognitive paradigm (see e.g. Harré and Gillett, 1994: ch 2). Edwards and Potter are quite explicit about the role the case studies in their book *Discursive Psychology* play in this endeavour. When discussing their choice of cases, they say: ‘In each case, these materials provide an excellent opportunity for developing the main themes of the discursive action model which we are offering as a centrepiece of discursive psychology’ (Edwards and Potter, 1992: 6–7). Additionally, they spend a lot of the book differentiating themselves from both cognitive psychology and social psychology, showing what they have to offer that is different from, and an improvement on, these fields.

Although this discursive revolution doesn't seem to have swept all before it as intended, was the aim of setting it up internal to the science? Differentiating itself from other disciplines (e.g. social psychology) just to show that it has something different to offer is not a method that has been internalised. However, this is not all that is happening here. The discursive psychologists do offer real critiques based on what cognitive and social psychology leave out.

Even ecological cognitive psychology disregards important things because the experimenter has the "correct version", or knowledge of the input to memory. In the John Dean case study, this is the tape recordings of the conversations. As Edwards and Potter say, the presence of these recordings was the whole reason for Neisser's study (Edwards and Potter, 1992: 33–34). For the discursive psychologists on the other hand, 'we might say that everyday conversational remembering often has this as its primary concern – the attempt to construct an acceptable, agreed or communicatively successful version of what happened (Edwards and Middleton, 1986a, refer to this as the 'validation function' in conversational remembering).' (Edwards and Potter, 1992: 75). On their view, other goals may be more important than strict accuracy, and what strict accuracy is varies with the goals and criteria for remembering in that particular context.

There is nothing wrong with pointing out an aspect of the phenomena that current approaches do not consider, and this can result in success. Because it was new at this time, discursive psychology had not had time to acquire a track record of success. Setting up this new approach is therefore an untried hypothesis.

Influence of other disciplines

Another important pressure on the discursive approach, and therefore on MEMORY, is the influence of other disciplines. Here I will discuss three of these: other fields which make use of discourse analysis, later Wittgensteinian philosophy, and the sociology of scientific knowledge (SSK). All three of these emphasize discourse, so apply pressure towards a distributed cognition concept, for the reasons given in the last subsection.

Discursive psychology is a variety of discourse analysis, so we might think that pressure is being externally imposed from other disciplines that already employ discourse analysis. At the time of writing the book, Potter was a reader in discourse analysis in the department of social sciences, and Edwards was a social psychologist, and both were members of the interdisciplinary "Discourse and Rhetoric" group at Loughborough University. They were therefore both familiar with discourse analysis as it is more traditionally applied in the humanities. Discourse analysis has diverse influences, for example from semiotics, narratology, ethnographic approaches etc. It seems that there is a risk of taking for granted techniques that worked when discourse analysis was applied in other disciplines, and importing them into psychology as external considerations. Edwards and Potter were immersed in these techniques, so this seems like a real possibility.

However, when a new subdiscipline is first set up, it is often too early to see whether this is happening. Techniques, theories, etc. have not yet had time to establish a track record of success, and so it is here. The techniques of discourse analysis seem to be reasonable techniques to try here, given that what Edwards and Potter argue is missing from psychology is a study of the discursive construction of memories etc. This influence is therefore not problematically external, but is an untried hypothesis.

The second important influence I will consider here is the later work of Wittgenstein. The importance of this factor is clear in Edwards and Potter's focus on pragmatically occasioned accomplishments occurring in discourse. In their criticism of Neisser's treatment of his Dean case study, they say of their suggested improvements in approach:

> It is to shift our allegiance from the Wittgenstein of the Tractatus, looking for the rules of correspondence between propositions and the world, to the Wittgenstein of the Philosophical Investigations, looking for the uses of language for constructing truth inside the 'language games' that make up a 'form of life' (Wittgenstein, 1921, 1953). (Edwards and Potter, 1992: 40).

They are seeking an even more contextualised approach than that of ecological cognitive psychology. Part of the point of this is for psychology to make itself relevant to everyday life, and this concern is internal to the science. A science that did not do this to some extent would be of little use. However, the specifically Wittgensteinian approach is an approach that is only just being tried out in psychology, so it hasn't had time to be internalised, but it is not being relied on unquestioningly, so is an untried hypothesis.

Another powerful influence on discursive psychology, and on Potter in particular, is SSK. Sociologists of scientific knowledge treat science as a social practice, and use the techniques of sociology, including discourse analysis, to study it. Potter at one stage planned to do a PhD with prominent sociologist of scientific knowledge Harry Collins. In the end the offer was withdrawn due to funding cuts, but Potter continued to work in the philosophy and sociology of science, with an interest in many of its major figures, including Kuhn, Feyerabend, Lakatos, Collins, Mulkay, Latour and Woolgar.

The influence of SSK can be seen in Edwards and Potter's work on Lawson, and their criticism of Neisser. In particular, in their criticism of Neisser for using his own categories of accuracy or truth in remembering, they cite major figures from SSK and say 'it is vital to maintain a neutral position with respect to what the participants treat as facts, or else their own interests and purposes begin to contaminate the analytical conclusions (Bloor, 1976; Collins, 1981; Mulkay, 1979)…this is precisely the issue raised by Neisser's (1981) study' (Edwards and Potter, 1992: 57). This recommendation to use participants' categories rather than analysts' categories is part of their rejection of an objective truth independent of the experimenter, which is an important idea in SSK. In Neisser's case study, the problem is, as I said above, his focus on the "correct version" – the tapes of the conversations. In their words: 'In the case of cognitive studies of memory, truth is equivalent to the psychologist's direct access to the input.' (Edwards and Potter, 1992: 73).

Another of their criticisms of Neisser is also heavily influenced by SSK. They criticise Neisser for his appeal to Dean's vanity as a distorting influence to account for particular errors in remembering (Edwards and Potter, 1992: 72). According to the strong program in SSK, social factors must be appealed to in order to explain what happens when science gets it right, as much as when it gets it wrong. The same is true here; veridical memory needs to be explained as much as erroneous memory. As Edwards and Potter say, '[a]n infinity of falsities is possible, so we need to account for why particular ones are produced. (It is our own argument, of course, that truthful descriptions can also be indefinitely elaborated).' (Edwards and Potter, 1992: 72). Neisser's talk of a distorting influence in some cases is therefore the wrong approach from this perspective. They explicitly cite the sociologist of scientific knowledge Bloor in this discussion (Edwards and Potter, 1992: 73).

Is the influence of SSK an internal factor? The techniques of discourse analysis it employs have not yet been tried out in psychology, so as I said before, they have not had time to develop a track record of success and be internalised. These techniques seem to be good ones to try out here; SSK involved looking closely at discourse, among other social aspects of the scientific process, and this is part of what seems to the discursive psychologists to be lacking in cognitive psychology.

Although there are good reasons to try these techniques, there is some reason to be concerned that their application is not speculative enough, and the theories of SSK are being taken too much for granted as theories that will also apply in psychology. Figures from SSK are cited as authorities in Edwards and Potter's book in a way that perhaps suggests their ideas are not sufficiently open to question and may be being externally imposed. However, it is too early in the history of discursive psychology to say this for sure.

Edwards and Potter do make some remarks about their treatment of Latour, which suggest that they are aware of taking his work for granted, but that doing so is a methodological necessity. They use Latour's own term "black boxing", saying that they black box Latour, as opposed to their treatment of Neisser, where 'we have opened the lid to gaze at the workings of this piece of research' (Edwards and Potter, 1992: 71). They point out that not doing this black boxing 'would quickly make all accumulative talk and text impossible' (Edwards and Potter, 1992: 71). This is surely right, and their self-awareness on this issue suggests that the black boxed ideas could still be subject to question in other circumstances, and are therefore untried hypotheses rather than problematically external.

Bartlett

Another influence on Edwards and Potter's work and on their concept of memory was Bartlett. When I discussed Bartlett as an influence on Neisser, I considered Wagoner's claim that Neisser's early work distorted Bartlett's notion of a schema, in particular neglecting its more social aspects. Edwards wrote a paper in 1987 with David Middleton, reinterpreting Bartlett as a fruitful source of inspiration for discursive psychology. In that paper, they share Wagoner's suggestion that Neisser and others neglected important parts of Bartlett's work, thus distorting it. In particular, they say that the affective and contextual aspects of schemata have been ignored largely because of the influence of the information processing account. They say:

> For Bartlett, schemata were not static knowledge structures stored in the brains or minds of individuals for the interpretation of experience, but rather were functional properties of adaptation between persons and their physical and social environments. Their essential properties therefore were social, affective and purposive, the basis of actions and reactions in the contexts of living one's life. (Edwards and Middleton, 1987: 80).

Their reconsideration of the more social aspects of Bartlett's work applies pressure toward a situated cognition perspective. Edwards and Middleton complain that recent study of memory has been too individualist, and advocate instead a focus on discourse as 'directly instrumental in the realisation and constitution of both individual and collective remembering' (Edwards and Middleton, 1987: 89). Notice that discourse here is *constitutive*, not just something that has an effect on remembering. The position in question here therefore seems to be something stronger than Neisser's merely embedded cognition. They go on to say:

> Bartlett was truly concerned with social-cognitive issues. He was concerned not with the ways in which social factors affect individual cognitions (e.g. Stephen, Brandstatter and Wagner, 1983), where two heads are seen to be more effective than one, but rather with the inherently social basis of mentality itself. (Edwards and Middleton, 1987: 89).

Pressure toward a distributed cognition perspective specifically can be seen in this understanding of Bartlett. [61] The focus is not individual cognition extended through discourse, but the discourse constructed by multiple individuals acting in concert.

Edwards and Middleton argue that the focus on social context leads to a loss of focus on accuracy of memory, and that Bartlett saw this. They quote him as saying: '"There is ordinarily no directed and laborious effort to secure accuracy"…(Bartlett, 1932, p. 96)' and '"literal recall is extraordinarily unimportant"… (*ibid.*, p. 204)' (Edwards and Middleton, 1987: 85). Bartlett's influence is therefore also partly responsible for the idea of shifting focus away from comparison with a correct version of events.

Another way in which an acquaintance with Bartlett's work may have applied pressure toward a distributed cognition concept of memory is through his work on what Edwards and Middleton call "symbolic remembering", i.e. 'putting things into words, into conventional and communicable symbols.' (Edwards and Middleton, 1987: 82). Bartlett was interested in what he called "conventionalization", i.e. the process by which material (e.g. techniques, customs, institutions) are changed when they are adopted into an alien culture (Bartlett, 1932: ch. XVI). He says of this material: 'The new material is assimilated to the persistent past of the group to which it comes' (Bartlett, 1932: 280). The part of this persistent past that most interested Edwards and Middleton is that embodied in symbolic form that can then be transmitted via text and pictures (Edwards and Middleton, 1987: 78–79). This recalls Hutchins' work in distributed cognition in the navigation of a naval vessel, in particular the role of maps and charts. For Hutchins, the symbols in maps and charts show the distribution of cognition, not just across people and artefacts, but also through time, as they are handed down and sometimes changed from one user to another (Hutchins, 1995).

[61] When looking at Bartlett's influence on Neisser, I claimed that his work applied pressure towards an enacted cognition concept. This is compatible with it applying pressure towards a distributed cognition concept for Edwards and Potter, because they are responding to different aspects of Bartlett's work, interpreted differently. The aim here is not to find out Bartlett's implicit concept and claim that reading him applies pressure in that direction, but to see how his influence appeared to those he was influencing.

Bartlett's symbolic remembering is also linked to the irreducibly social nature of cognition for discursive psychology, because the symbols exist only as part of a joint activity in some sense. As Edwards and Middleton say, '[t]he study of discourse in the functional contexts of everyday life offers the bridge between the individual and the social that Bartlett sought throughout his work, and attempted to capture through the conventionalization of successively remembered symbolic materials.' (Edwards and Middleton, 1987: 89). In terms of the modern distributed cognition framework, the transmission of memory through symbolic materials can be seen as the distribution of memory across groups of people, the symbolic materials themselves, and through time (compare Hutchins, 1995).

This use of Bartlett's work in the new field of discursive psychology has not yet had time to acquire a history of success and be internalised, but is it being problematically imposed from an external perspective, or treated as an hypothesis?

The role Bartlett is playing here is not that of an authority figure whose words should be accepted without question. Edwards and Middleton are not entirely uncritical. For example, they note that Bartlett was still looking *through* discourse, rather than *at* it as they advocate (Edwards and Middleton, 1987: 87). It is also important that Bartlett is not being set up as anything like a discursive psychologist from before discursive psychology was invented. Rather he is being treated as a potential source of inspiration. Edwards and Middleton say explicitly that they are not reinvestigating Bartlett to set the historical record straight or to provide a definitive account of his work, but '[r]ather the prime concern is with the value of such a re-examination for certain issues which are currently significant in cognitive psychology and in the study of social cognition' (Edwards and Middleton, 1987: 77). His views are not being imposed on the field from outside, but are just being explored speculatively. His influence is therefore an untried hypothesis.

1980s intellectual milieu

Having considered some specific influences on Edwards and Potter, I now want to mention the more general intellectual milieu, as I did for Neisser. The influences here are much the same as they were for Neisser, given that they were only writing a few years later. Recall that these included the rediscovery of the work of Halbwachs and Vygotsky, the failure of classical A.I. and the rise of connectionism, the influence of dynamical systems theory, and new disciplines becoming involved in cognitive science, including anthropology, education, media theory, design, and Human Computer Interaction.

The discursive psychologists are in a similar position to Neisser here, in that reacting to what is happening in neighbouring disciplines helps them to stay relevant to those disciplines. These are not influences that are being treated as authoritative, merely possible sources of inspiration for new methods and explanations to try out in a developing discipline. They are therefore untried hypotheses.

The Gulf War

I want to consider one final influence on Edwards and Potter's concept of memory, and this is the Gulf War, and particularly its portrayal in the media. Edwards and Potter's book was written at the time of the Gulf War, and they specifically discuss it in both the introduction and the conclusion. In fact, in the very first paragraph, they say that while writing the book: 'We were particularly struck by the way that what we were writing about psychology could be applied to the versions of the Gulf War that were made available in a torrent of media coverage.' (Edwards and Potter, 1992: 1).

At the end of the book, they give some suggestions as to how their ideas might be applied, in particular with relation to some of the words used in the media reports, what those words were meant to accomplish, and what inferences from the reports were licensed by them that may not have been licensed if other words had been used (Edwards and Potter, 1992: 170–175).

Looking at the media construction of accounts of the war as an important illustration of their approach is part of the epistemic niche for the discursive psychologists' concepts. Significant events like wars are important examples of memories which can be seen as distributed across a number of people and texts, and through time. This factor therefore applies some pressure toward a distributed cognition concept of memory.

Looking for ways to apply techniques to current issues is something that is internal to all sciences. That is not to say that there can't be research that cannot be so applied, only that when it can be, the idea of doing so is an internal consideration. Such relevance to important contemporary issues is part of what *constitutes* success for a science, so it is something like an analytic truth that it has a track record of success and is therefore an internal consideration. In this instance, despite discursive psychology's immaturity, the fact that this factor is internal at a coarser level of grain makes it internal here.

Summary

In summary of what I have said about the epistemic niche for Edwards and Potter:

Factor applying pressure to concept	**Concept encouraged by this pressure**	**Internal / hypothesis / external**
Setting up a new subdiscipline	Distributed	Hypothesis
Influence of other disciplines	Distributed	Hypothesis
Bartlett	Distributed	Hypothesis
1980s milieu	Situated (various)	Hypothesis
Gulf War	Distributed	Internal

All of the factors in the epistemic niche considered here apply pressure towards a situated cognition concept of memory, in most cases a distributed cognition concept. I have considered the effect of reacting against cognitive psychology and setting up the new subdiscipline of discursive psychology, the influence of the ways in which discourse analysis was applied in other disciplines, the work of later Wittgenstein, SSK, the reconsideration of Bartlett's work, other influences in the intellectual milieu of the time (failure of classical AI, involvement of new disciplines in cognitive science, rediscovery of Halbwachs and Vygotsky etc.), and the application of the discursive approach to Gulf War narratives in the media at the time. I have argued that none of these influences are problematically external, so the distributed cognition concept is rational to the extent that it is shaped by the factors considered here.

The application to Gulf War narratives is an internal factor, so the distributed cognition concept is legitimate to the extent that it is shaped by this factor. It is non-legitimate to the extent that it is shaped by the other factors considered here, but this is because discursive psychology is not old enough at the time of this case study for any of the other factors to have had a history of success and been internalised. There is therefore relatively weak evidence for a distributed concept functioning as an investigative kind in discursive psychology, but that is due to the newness of the subdiscipline, and it is at least suggestive that a situated cognition concept is rationally being tried out in a newly emerging field.

Conclusion

This case study has revealed that Neisser (an ecological cognitive psychologist) and Edwards and Potter (discursive psychologists) are working with situated cognition concepts of memory. In Neisser's case the concept is embedded, and in Edwards and Potter's case it is distributed. In each case I have considered some of the main factors in the epistemic niches of the concepts of memory which are responsible for them taking the shapes they do.

Neisser's concept is being shaped by a variety of competing factors, some applying pressure towards an embedded concept, some enacted, and some brainbound. There are internal considerations in all three of these categories. There is some evidence that the embedded cognition concept is functioning as an investigative kind, but I have argued that the pressures on Neisser's concept should result in pluralism, where multiple kinds of concept and their accompanying frameworks are in use side by side. Therefore there is evidence here of a situated (in this case embedded) cognition concept functioning as an investigative kind, *and* scope for more kinds of concept to do so too.

Edwards and Potter's distributed cognition concept is rational to the extent that it is shaped by all the factors considered here, and legitimate to the extent that it is shaped by the application of the discursive approach to Gulf War narratives in the media. It is non-legitimate to the extent that it is shaped by the other factors considered here, but this is because discursive psychology is not old enough at the time of this case study for any of the other factors to have had a history of success and been internalised. This is not surprising, given what I have said about immature sciences in previous chapters.

It is too early to tell whether the discursive psychologists' distributed cognition concept will pick out an investigative kind, but it is suggestive that a situated cognition concept is rationally being tried out in a newly emerging discipline. There is no way of knowing whether it will be a successful concept, but there is no reason to think that it will not be. There is therefore no reason to stipulate that the science would be best served by sticking to brainbound concepts, as some critics of the situated cognition approaches might recommend.

8. Transactive Memory Systems: A Case Study

Introduction

This case study focuses on a social variety of situated cognition, where memory is putatively extended (or distributed etc.) over other people rather than over artefacts or tools. In particular, the topic is *Transactive Memory Systems*, proposed in 1985 by Wegner, Guiliano and Hertel (hereafter WGH) in their chapter "Cognitive interdependence in close relationships" in W. J. Ickes (Ed.), *Compatible and incompatible relationships*, a book in the *Springer Series in Social Psychology*. A Transactive Memory System (TMS) is defined as 'a set of individual memory systems in combination with the communication that takes place between individuals' (Wegner, 1986: 186; see also WGH, 1985).

The idea was enlarged upon in Wegner's paper "Transactive Memory: A Contemporary Analysis of the Group Mind" (Wegner, 1986). A TMS is something like a post-cognitive-revolution group mind. Wegner and his collaborators are keen to distance themselves from older discredited group mind notions, and to use the tools and ideas developed in the cognitive revolution (an encoding-storage-retrieval model of memory based on the computer metaphor etc.) This interesting relationship to their own history will be important in what follows.

The case study will compare the social psychological work of WGH with work in Communication Studies, a complex interdisciplinary hybrid of psychology, sociology, anthropology, journalism, political science and English, among others. In particular I will focus on a discussion of TMS in the journal *Human Communication Research* in 2003. Here Pavitt (a professor of communication) criticises the notion of a TMS (Pavitt, 2003a), others in his field respond (Hollingshead and Brandon, 2003; Propp, 2003; Wittenbaum, 2003), and Pavitt replies in light of their responses (Pavitt, 2003b).

The main point of contention is what the relevant baseline against which a putative TMS is tested should be. There are at least two important questions that comparison with a baseline can be used to answer: (a) whether a TMS exists, and (b) whether the

TMS is functioning well. These questions are often not well distinguished, and I will discuss this further below.

Other important issues are the role of communication between the individuals in a TMS, and the relationship between the concept of a TMS and older notions of a group mind. The latter is something that has received considerable attention from participants in the debate, and which my historical approach is ideally placed to investigate. All of these issues turn on *what it is to remember* for the different groups, i.e. on MEMORY.

The case study

WGH's initial research proposing the idea of the TMS concerned remembering in long-term heterosexual couples. WGH argued that such dyads were able to remember things that neither individual could remember alone, using processes such as interactive cuing (WGH, 1985: 256–257). Essential to this achievement is meta-memory, in particular remembering one another's areas of expertise (WGH, 1985: 264). These meta-memories may be based on stereotypes (e.g. that the female remembers things to do with the children), or specialisms (e.g. one partner typically takes responsibility for remembering social engagements), or who acquired a particular piece of information first (e.g. in taking a telephone message). (See WGH, 1985: 265–266; Wegner, 1986: 191–192).

The relevance to social versions of extended cognition, and to distributed cognition, should be obvious. I will consider below whether the participants in the debate actually seem to be making use of such concepts of memory, or whether this link is only apparent. More recently, philosophers have noticed this work and connected it explicitly with distributed cognition, for example Theiner (2013) and Tollefsen et al. (2013), and I will briefly mention their papers in the next section.

This case study is important for my overall project because there may be important differences between social versions of situated cognition and the more artefact-based versions I have considered so far. Tollefsen et al. explicitly suggest that social extended cognition is plausible whereas artefact extended cognition is not (Tollefsen et al., 2013: 58, note 7). Although my last case study concerned memory as constructed between participants in discourse, written records of the discourse (tape recordings and newspaper reports) still had an important role. Although the idea of a TMS including artefact memory has been discussed (e.g. Wegner, 1986: 187–189), it is usually considered to be purely social. This is important if Tollefsen et al. are correct about the difference in plausibility between artefact and social varieties of situated cognition.

There has been considerable interest in the notion of a TMS from business, management and communication studies because the ideas can be applied to work groups and teams, and possibly to larger organisations (although this presents more challenges). The results of such research may help groups to work more efficiently, thus benefitting businesses.

For the communication studies side of my analysis, I will look at criticisms of the TMS from Pavitt (a professor of communication) in *Human Communication Research* in 2003, and some responses to his work.

Pavitt is concerned that TMS research is rejuvenating the old abandoned idea of the "group mind", even though its proponents go to great lengths to explain how their proposal is different from the work of Jung, Durkheim, Hegel, Rousseau, Wundt, McDougall, etc. Pavitt acknowledges that it can be valuable to apply terms usually used about individual cognition to the group level. However, he goes on to say that '[t]his use is, however, inherently metaphorical, and there is a risk that the idea of "group information processing" will become reified.' (Pavitt, 2003a: 593).

Pavitt's main dispute is over the baseline to which a putative TMS should be compared. His claim is that TMS supporters have taken the wrong baseline in their studies. He argues that a group such as those studied in TMS research performs no better than an aggregate of individuals (a nominal group). He uses a mathematical model (Lorge and Solomon model B) to calculate the performance such a nominal group should have, and compares this to the results of TMS research.

Lorge and Solomon (1955) propose a model (Model A) which predicts group performance under a series of assumptions, most importantly the assumption that communication between group members serves as an accurate conduit for communication of the results of individual problem solving, but no more. The model therefore gives the expected result of a nominal group – a group whose performance is that of the aggregate of its individual members – rather than the result of a group whose interactions allow new properties to emerge at the group level. Model B is an extension of Model A to allow it to handle problems with multiple stages.

In every case Pavitt considers, the group remembers more than an individual, but less than the prediction of Lorge-Solomon model B (Pavitt, 2003a: 597, table 1). According to Pavitt, this shows *inhibition* caused by the communication between group members. This is because, given that Lorge-Solomon model B's prediction should give the result expected for the aggregate of the individuals, if the group underperforms this, the group is less than the sum of its parts. To use Pavitt's term, we should be "reductionists" about group memory. Because there is nothing special emerging at the group level, for Pavitt the term "Transactive Memory System" should only ever be used as a metaphor; it does not describe a real new phenomenon, hence his worries about "reifiying", quoted above.

Three responses to Pavitt in the same journal issue are more optimistic about the usefulness of the idea of a TMS. The first of these is Hollingshead and Brandon (2003). They argue that a group that has a TMS outperforms a group that does not. They say: 'When group members divide responsibility for different knowledge areas and are aware of one another's expertise (i.e., when they have a transactive memory system), groups perform better than when they do not (Stasser, Stewart & Wittenbaum, 1995).' (Hollingshead and Brandon, 2003: 608). This disagreement with Pavitt could either be due to a difference in appropriate baseline, or in what groups are classed as having a TMS. In fact it seems to be both, although they only explicitly argue the latter.

As a baseline, much of the research looked at by Hollingshead and Brandon took actual results from testing of nominal groups rather than using a mathematical model as Pavitt does. Although they do not explicitly argue that this is a problem for Pavitt, it does seem to be an important difference between them.

What Hollingshead and Brandon do explicitly argue is that most of the research looked at by Pavitt does not even concern groups that have developed a TMS. They say of the cases Pavitt refers to:

> Group members were strangers to one another prior to the study (with the exception of the dating couples in Wegner et al., 1991), and did not have an opportunity to learn about each member's expertise and abilities. Participants learned the information individually, and then were asked to retrieve the information collaboratively, so the ability to communicate and to collaborate with others on the memory task varied during the learning and the recall phases of the studies. Many of the tasks used in the studies involved episodic rather than semantic memory, and specific knowledge areas were not clearly defined, which may have made them less conducive to a division of labor. Based on these features, we believe that it is unlikely that the groups in the studies reviewed by Pavitt (2002, with the exception of Wegner et al., 1991) had a developed transactive memory system. (Hollingshead and Brandon, 2003: 613).

There is disagreement over both the baseline and the putative TMS that is being compared to it. There is therefore some debate over what counts as a TMS, i.e. what counts as remembering for the group. Because this is a disagreement over what counts as remembering, it is a disagreement over MEMORY. I will consider this issue in more detail below, as part of a discussion of the pressure particular notions of "success" (what it is to succeed at remembering) exert on MEMORY.

The second paper from the communication group that is favourable to TMS is Propp (2003). She argues that Pavitt's critique should be taken as a spur to further work, particularly on the communication between individuals in a TMS. Pavitt appears to gloss over this issue, and again the problem is with his use of a mathematical model as baseline.

Pavitt's favoured Lorge-Solomon model B makes the following assumptions:

> (a) communication serves as a flawless conduit for information, such that if any group member solves a given problem the group will accept the member's solution as correct, and (b) the odds of different members solving that given problem are independent of one another. (Pavitt, 2003a: 593).

To assume that communication is flawless is to ignore the importance of variations in communication. Propp therefore disputes assumption (a) as follows:

> It is important to note that while metaphors of collective information processing and TMS posit the centrality of communication, they do not presume, as Pavitt (2003) suggests, that interaction will be flawless (or even effective) in the creation of a group memory. Actually, quite the opposite is true. To extend the collective information processing metaphor, just as individuals are prone to processing limitations and errors, the group also will experience processing limitations and errors, and these problems are by definition communicative phenomena when examined at the group level (see Wegner, 1986; Propp, 1999, for discussion of the ways interaction can inhibit group memory). Moreover, even as we seek to substantiate assembly bonus effects [beneficial effects of working in a group beyond the summing of individual group members' contributions] we should also identify the ways in which interaction inhibits group memory, because it may be through the understanding of process loss that methods for enhancing process gains are identified. (Propp, 2003: 602).

As can be seen from this quotation, Propp is arguing that a group need not communicate completely effectively in order to have a TMS. Her disagreement with Pavitt concerns whether a group that remembers less well than a particular baseline (itself contentious) should be classed as not having a TMS, or as having one that is functioning poorly. In other words, it is a question of whether poor recall should not be classed as memory at all, or whether it should just be classed as bad memory.

Making a related point about successful remembering, Propp also says:

> It is possible that the definition of improved performance might need to be conceptualized more broadly to account for the presumption of groups' superiority in organizational theory and practice. This definition of performance might include other group outcomes that have been found to make groups superior, such as representation of diverse goals and values, distribution of responsibility for decisions, and increased commitment to an action (Gigone & Hastie, 1996). (Propp, 2003: 604–605).

In other words, there may be more to the idea of a group performing better than an aggregate of individuals than exceeding a baseline like Lorge-Solomon model B. Like Hollingshead and Brandon, Propp is disputing Pavitt's assumptions about what counts as having a TMS, and therefore what counts as remembering for a group (see the discussion of success below).

The final response to Pavitt is Wittenbaum (2003). Her main point appears to be similar to Propp's although it is expressed as a suggested expansion of Pavitt's work rather than a criticism. I will not consider her response further here, as much of the paper is a summary of Pavitt's work citing data in support of it, and this does not give enough information to infer what her concept of memory might be.

Pavitt's response to these three papers is to clarify his point about relevant baselines as follows:

> If the relevant question is whether groups remember items of information better than do individuals, the correct baseline for evaluating group memory is individual memory. If the relevant question is whether group interaction enhances, retards, or has no impact on group memory, the correct baseline is whether groups of a given number of people remember items of information better or worse than either equal-sized combinations of analogously skilled individuals ("nominal groups") or, as I did using Lorge-Solomon Model B (1955), arithmetically derived facsimiles of such combinations. The researchers I critiqued in the opening essay used the former baseline when they should have used the latter. (Pavitt, 2003a: 626).

Here he sticks to the claim that Lorge-Solomon Model B is an accurate reflection of the performance of nominal groups, but is more open about what counts as successful remembering; it need not always be performing better than this particular baseline.

Pavitt also says in his response that he agrees with Propp and Wittenbaum that communication researchers have an important role to play in studying the communication between group members, and that more research in this area would be welcome. I will return to this point in my discussion of the pressure placed on MEMORY by the aim to consolidate the subdiscipline.

In the next section, I will discuss the implicit concepts of memory employed in this debate, first for Wegner and his collaborators, then the communication researchers.

Concepts of memory

Social psychology

Wegner is a social psychologist, but one of his collaborators, Hertel, is a cognitive psychologist. Cognitive psychology is typically individualist, and WGH are very keen to incorporate the insights of the cognitive revolution so that they are not seen to be going back to older group mind ideas. As they say, 'references to the group mind in contemporary literature has dwindled to a smattering of wisecracks' (WGH, 1985: 253). However, it seems that social psychology must assume that groups (or at any rate society as a group) exist in some interesting sense. The resolution of this tension is often to see groups in terms of the individuals that comprise them. This is reflected in the definition of a TMS as 'a set of *individual* memory systems in combination with the communication that takes place between *individuals*' (Wegner, 1986: 186, my emphasis).

Does this relatively individualist stance indicate a brainbound perspective on cognition rather than a situated one? I don't think it can do so without giving up on the very idea of a TMS. In TMS research, it is the *group* that is typically said to remember. Wegner says quite explicitly that '[t]ransactive memory is…not traceable to any of the individuals alone, nor can it be found somewhere "between" individuals. Rather, it is a property of a group.' (Wegner, 1986: 911). The only candidates for a perspective on cognition that captures this social element of the TMS are distributed cognition (e.g. Hutchins, 1995), and social versions of extended cognition (e.g. Clark and Chalmers, 1998; Clark 2003, 2008).

There is an important difference between these two views, namely that extended cognition is *organism centred* (see especially Clark, 2007) – *someone's* memory extends over or into another person or people. On a distributed cognition perspective, the memory may be a property of the group as a whole (see Hutchins' 2011 review of Clark's *Supersizing the Mind* for more on this difference). A TMS is therefore a distributed cognition notion, rather than an extended one, because it is *the group* that is said to remember. In fact, given Wegner's definition, a distributed cognition concept is essential to the idea of a TMS; anyone who accepts TMS as true group memory is employing some kind of distributed cognition concept.

Some indications of this implicit concept can be found in the writing of TMS researchers, for example WGH say that '[t]o build a transactive memory is to acquire a set of communication processes whereby *two minds can work as one*' (WGH, 1985: 263, my emphasis). WGH also talk about an analogy between the functions of group memory and individual memory, saying of the functional equivalence of individual and transactive memory:

> Both kinds of memory can be characterized as systems that, according to general system theory (von Bertalanffy, 1968), may show rough parallels in their modes of operation. Our interest is in processes that occur when the transactive memory system is called upon to perform some function for the group – a function that the individual memory system might reasonably be called upon to perform for the person (WGH, 1985: 256).

This is reminiscent of the Parity Principle (the argument that there is functional parity between brainbound cognition and extended cognition (Clark and Chalmers 1998)), but here it is being used to argue for something more like distributed cognition. The equivalence is not between a person's brain and their brain plus external elements (perhaps other people), but between an individual and a group as a whole.

The attitude that it is the group that remembers is also revealed in the experiments carried out. It is the *group's* answer to a question or task that WGH were interested in. In the main piece of research they discuss (Guiliano and Wegner, 1983), negotiation to produce a single answer from the dyad was important. The research compared "close" and "distant" couples, where the close couples were those who had carried out together a task designed to induce cohesion before testing. They say:

> The hypothesis was that "distant" couples, when faced with disagreement would opt for the personal choice of one or the other partner; "close" couples, in contrast, were expected to use such conflict as a stimulus to invent a new, group-generated possibility (WHG, 1985: 268).

From this interest in negotiation to produce a single answer, it is clear that memory was seen as something attributed to the dyad, as we would expect from the distributed cognition perspective, not to one member but including the other as a part, as we would expect from the extended cognition perspective.

It is also suggestive that some philosophers interested in distributed cognition have recently taken an interest in TMS research. Theiner (2013) and Tollefsen et al. (2013) write papers concerning TMSs in a special issue of *Review of Philosophy and Psychology*. The purpose of the special issue is to look at the theoretical arguments and experimental evidence for distributed cognition in memory research.[62]

All these things point to the fact that the TMS as conceived by WGH essentially, if implicitly, involves a distributed cognition concept of memory.

Communication studies

Does the same apply to Pavitt and the other communication researchers? Must their concept of memory be distributed simply because they are discussing TMSs? Pavitt's for one need not, since he is arguing against the notion; in fact, it seems that having a different concept of memory is a large part of Pavitt's rejection of TMS. A consequence of this is that the participants in the debate are talking about different (although of course overlapping) things when they talk about memory.

I quoted above Pavitt's fear 'that the idea of "group information processing" will become reified' and his belief that all talk of cognition at the group level is "inherently metaphorical" (Pavitt, 2003a: 593). Given that he admits that such talk can be valuable, his argument against it cannot be based on what is fruitful for science, but must be based on what he takes memory to be. This interpretation is backed up by his choice of the phrase "*inherently* metaphorical". It also receives some support from the fact that he often (although not always) uses scare quotes when talking about putative group mental states (e.g. what the group "knows", group "memory", etc.). This suggests that he has a concept of memory according to which "memory" cannot be *literally* applied at the group level.

[62] A caveat here is that the focus of the special issue is on what I have called "situated cognition" perspectives more broadly, despite their use of the term "distributed". In their editors' introduction, Michaelian and Sutton talk about "distributed/extended" ideas (2013: 1). Neither Theiner nor Tollefsen et al. are particularly interested in the taxonomy of the different situated cognition positions, but their focus is primarily on versions where cognition is distributed in the narrower sense of Hutchins (1995).

If this reading is correct, then Pavitt certainly does not have a distributed cognition concept of memory; but what kind of concept does he have? It is individualist, but that does not mean that it isn't a situated cognition concept of some kind. All the situated cognition perspectives apart from distributed cognition are individualist, even the social variety of extended cognition, because it is organism-centred, as I said above.

However, extended cognition does not seem to be the perspective that Pavitt has in mind. In the experiments he cites (see Pavitt, 2003a: 595–596), individual recall is compared to group recall, but in the group recall condition it is the answer of the *group* to each memory task that is recorded. If he was working with an extended cognition concept, these experiments would not be relevant, and instead he would be interested in the answers of a particular individual after consulting with the group.

We can also rule out embedded and enacted cognition by looking at the experiments Pavitt cites. They are laboratory-based experiments, and he makes no mention of possible problems with this fact, or of the role that the wider environment could be expected to play. In both the embedded and enacted cognition approaches, this would be essential. There is also no reason to think he has an embodied cognition concept since there is no mention of the role of the body in his paper or in any of the experiments he cites. That leaves us by elimination with the conclusion that his concept of memory is a brainbound one.

Pavitt is at least implicitly aware that accepting WGH's notion of a TMS would involve accepting a distributed cognition concept of memory. In fact, I argue that his rejection of the notion of a TMS is partly because he implicitly rejects this concept in favour of a more individualist, brainbound concept of memory. Do the other communication researchers share Pavitt's concept?

As I have said, the TMS as construed by WGH essentially involves distributed cognition. Given that Hollingshead and Brandon and Propp are critical of Pavitt and favourable to the notion of a TMS, these researchers must either have a distributed cognition concept of memory, or they have a different idea of the TMS. In the latter case, their talk of "group memory" would be merely metaphorical.

Hollingshead and Brandon introduce the idea of the TMS by analogy with memories being stored in an external artificial device like a personal digital assistant (Hollingshead and Brandon, 2003: 608). This makes it sound as though they are thinking in terms of extended cognition, acknowledging that other people can function as an external memory store, much like artificial devices (compare Clark and Chalmers, 1998). However, as discussed above, this way of thinking is not a very good fit with the TMS because memories belong to the group, not to an individual using the rest of the group as a cognitive tool. It therefore seems unlikely that their implicit concept matches this explicit avowal.

Later in the paper, they talk about ‘the development of an effective group memory system’ (Hollingshead and Brandon, 2003: 608). They also talk about other mental states at the group level:

> Communication can give groups the opportunity to develop a shared representation or shared mental model of their task (e.g., how they will work together to learn and later remember the words, who will assume responsibility for which categories). Shared mental models provide group members with a set of organized expectations for collective performance (Mohammed & Dumville, 2001). Creating a shared representation may be useful when groups know they will have the opportunity to interact when they need to recall the information. (Hollingshead and Brandon, 2003: 612).

The repeated use of the word “shared” in this quotation could either mean that each individual has the same representation, model etc., or that that these things can actually be found at the group level. The former option cannot be right because differentiation of tasks is an important part of the TMS (strategies for who will remember what etc.) The best interpretation of “shared” is therefore that representations, models etc., and memory too, are actually found at the group level. Therefore their uses of “memory” in the phrase “group memory” cannot be metaphorical, which indicates that they have a distributed cognition concept of memory, true to the original notion of the TMS.

Propp on the other hand talks about ‘*metaphors* of collective information processing and TMS’ (Propp, 2003: 602, my emphasis). It seems that she does not take the TMS to be one kind of description of what memory *is*, but rather as a fruitful metaphor for research. She has a different idea of the TMS to WGH for whom, as I argued above, it is a genuinely distributed memory. Although this indicates that she has an individualist brainbound concept like Pavitt's, there is some evidence that her concept is more flexible than his. She claims that 'it is premature to halt attempts to identify assembly bonus effects' (Propp, 2003: 601), arguing for further study of the communication between group members instead.

She also shows some flexibility in terms of what should count as performing well for a group, as I said above (Propp, 2003: 604–605). Changing this criterion would effectively be a redefining of what counts as remembering, i.e. of MEMORY, although not necessarily with respect to taking a brainbound or situated cognition approach. This indicates some flexibility in her concept, much like that seen in the philosophers' concepts in the first case study. What the concept should be is part of what is at issue in the research for Propp.

There is therefore some variation in concepts among the communication researchers, with Hollingshead and Brandon having a distributed cognition concept, and Pavitt and Propp retaining a more individualist brainbound concept with references to "group memory" being only metaphorical. However, Propp's concept is more flexible than Pavitt's. In the next section, I will consider possible reasons for this variation.

Summary

The social psychologists have a distributed cognition concept. Of the communication researchers, Pavitt has a brainbound concept, Hollingshead and Brandon have a distributed cognition concept, and Propp has a concept that is brainbound, but with some degree of flexibility.

Identifying and assessing the epistemic niches

Social psychology

The epistemic niche of a concept is constituted by the features of the context which apply pressure to the concept. Here I will therefore consider the influences on WGH, and what they needed their concept of memory to do. The first influence I will consider is older ideas about a group mind.

Rejection of old group mind ideas

Getting away from old group mind ideas and their bad associations is a major pressure on MEMORY for WGH. It is a pressure away from a distributed cognition concept, where memory is viewed as something that a group can possess, and towards an individualist concept. As I said in the last section, any of the other positions I am considering here could be individualist. However, in this case, pressure away from the distributed cognition concept is pressure towards a brainbound concept, because the kind of work being done is laboratory-based experimental work that does not concern itself with the environment or context in the way that any of the other situated perspectives would do.

WGH begin their paper with a lengthy section about what was bad about older group memory research, and how different their work is from this discreditable history. They say that they are ‘hoping to establish a more verifiable (and falsifiable) analysis by means of the idea of transactive memory’ (WGH, 1985: 256). The group mind notion has overtones of the supernatural, and they are particularly concerned to distance themselves from this. However, there is more to it than this supernatural element – it was a dominant research program for a long time – and there therefore seems to be a risk of throwing out something important.

Whether or not distancing themselves from their own history in this way is internal to the science will depend on why the old group mind notions were discredited, and whether WGH are taking these reasons into account, or for example just trying not to associate themselves with a tainted research paradigm for reputation’s sake. This latter situation could have the bad consequence of throwing out something important along with that which should be discarded.

In the paper itself, WGH do discuss how the old group mind idea came to grief. They say that being forced to see the group mind as nothing more than a group of individuals was in danger of giving the group mind proponents nothing interesting to say, so they retreated to the supernatural in order to retain something distinctive (see

WGH, 1985: 255). Here they cite some works on the history of social psychology (Allport, 1968; Knowles, 1982), suggesting some regard for their own history beyond the stereotype of the supernatural group mind. They also consider what might have been positive in the old idea: 'Is there anything in the idea worth preserving? Along with the early theorists, we believe that an emphasis on the difference between group and individual mental processes is an indispensable part of the definition of each' (WGH, 1985: 255).

This doesn't grant much to the old ideas, but it does show some consideration of advantages as well as disadvantages. Of course, this could be mere lip service to history, so it is important to look at the wider story beyond WGH's work to see the origin of this pressure to avoid the group mind idea.

The notion of a group mind enjoyed wide acclaim for a time in both psychology (e.g. Jung, 1922; Wundt 1910/1916) and sociology (e.g. Durkheim 1915). It is therefore perhaps unsurprising that it was an important idea in the foundation of social psychology, as a hybrid of the two disciplines. According to *The Blackwell Handbook of Social Psychology: Group Processes*:

> Two of the earliest texts in social psychology were Le Bon's (1896/1960) *Psychologie des Foules (Psychology of Crowds)* and McDougall's (1920) *The Group Mind*. Both espoused as a central tenet the view that behaviour in social aggregates was not simply a function of some combination of individual acts. Rather, they saw social behaviour as being guided by forces defined by the aggregate – a "collective consciousness" or "group mind" – that could not be understood fully by simply understanding individual behaviour or individual minds. Such ideas were not unusual for the times. Durkheim (1893/1984, 1965), Mead (1934) and other sociologists and social philosophers also saw collective or shared meaning as an integral component for understanding social behaviour (see Farr, 1996). (Hogg and Tindale, 2001: 1).

Since then, the fortunes of the group mind notion have changed somewhat. The *Blackwell Handbook* goes on to say:

> However, with the onset of Behaviourism, psychology's focus moved almost exclusively onto the individual, and the notion of collective thought and meaning fell out of favour (Allport, 1924). In mainstream social psychology, focus on aggregates versus individuals has

> waxed and waned (see Steiner, 1974, 1986; Moreland, Hogg & Hains, 1994 for reviews), but the key explanatory variables have remained mainly at the individual level. Thus in recent social psychology textbooks, the early ideas concerning "collective cognition" are rarely mentioned except for historical context, if they are mentioned at all (e.g., Baron & Byrne, 2000). (Hogg and Tindale, 2001: 1).

By the time WGH were writing, behaviourism had largely been replaced by cognitivism, but the individualism remained. I will discuss the influence of cognitivism in a later subsection, but we can see already that the move towards individualism was part of a wider story in which the dominant paradigm for the whole discipline of psychology changed.

More specifically, there were important critiques of the older group mind notion as part of this shift. One of these came from Bartlett, a major figure in both early cognitive psychology and early social psychology, who I have already introduced in this thesis (see chapter 7). Bartlett criticised the notion of a collective unconscious, in particular Jung's work, in his (1932), relating his criticism particularly to memory research.

According to Bartlett, the collective unconscious in social psychology is an analogue of memory traces in individual or general psychology. He had already argued against the static trace theory of memory in the first half of the book, so this comment sets up his criticism of the collective unconscious. More precisely, he says that Jung confuses a storehouse of knowledge with capacities or functions. He quotes Jung as saying:

> Every man is born with a markedly differentiated brain which makes him capable of very varied mental functions, whose acquisition and development are not ontogenetic in origin. Now, in proportion as the brains of all human beings are equally differentiated, the mental functions rendered possible by this level of differentiation are collective and universal. (Jung, 1916, quoted in Bartlett, 1932: 285).

Bartlett says that this is both vague and confused. The confusion is between what brain differentiations *are* and what they *allow* people to do. He points out that '[n]obody knows precisely what brain differentiations are capable of producing in the way of human reaction; but it is fairly certain that they are a basis for the subsequent development of functions, and not a storehouse of a mass of detailed acquired knowledge.' (Bartlett, 1932: 285).

Bartlett also considers the evidence in favour of the collective unconscious (intelligence tests on institutionalised children from varying social classes), and shows it to be wanting (Bartlett, 1932: 282–283). Such evidence is supposed to show "biological continuity" that is not just due to the children's surroundings. Innate continuity of this kind would suggest a collective unconscious. Bartlett allows that there may be some positive evidence, but shows that this evidence falls well short of proof (Bartlett, 1932: 283, 291).

The beginning of the backlash against the collective unconscious and the group mind can be seen here as part of a wider story, in this case the move from a static trace theory of memory to a broadly functionalist theory, in the sense of "functionalist" introduced in the last chapter (concerning function in a particular environment). Early rejections of the group mind idea, like Bartlett's, were motivated by consideration of the evidence, and their alternative functionalist paradigm was successful. Many of these successes were surveyed in the last chapter, including Bartlett's own work, that of the Gibsons, and Neisser's ecological psychology. Modern cognitive psychology is a descendant of this kind of functionalism, and this had also had much success by the time WGH were writing. I will consider this fact and its implications for social psychology in more detail in the subsection on cognitivism below.

WGH's rejection of the group mind is part of this bigger picture, and fitting in with the narrative in which older group mind ideas (e.g. the collective unconscious) were rejected is therefore an internal factor in the epistemic niche of MEMORY for WGH.

It should be noted here however that this older discredited kind of view was not the only path available at the time WGH were writing that a science of the group mind could proceed along. In the next subsection, I will consider the pressure towards another kind of group mind idea.

New group mind ideas

Wegner says of the old group mind ideas: ‘Obviously, these ideas do not represent the only direction in which group mind theories may develop (cf. Bartlett. 1932), and this chapter presents a fresh start toward a more useful formulation’ (Wegner, 1986: 185). In this subsection, I want to look at the pressure towards an alternative group mind thesis, and where this pressure came from. Like the older group mind idea, this is a pressure towards a distributed cognition concept of memory.

Wegner mentions Bartlett’s 1932 work, which I have already discussed in some detail in the last case study. Recall from that discussion that Wagoner (2013) accused Neisser of misinterpreting Bartlett, leaving out the more social and culturally-embedded side of his work. Wagoner goes on to say that ‘[b]y the late 1980s, psychologists were beginning to again acknowledge the social and cultural side of Bartlett’s thinking’ (Wagoner, 2013: 567). He goes on to talk about discursive psychology and ecological psychology. These movements, as discussed in the last case study, were part of a broader move in the 1980s, including the rise of connectionism due to the failure of classical AI, alliances with dynamical systems theory within cognitive science, Bruner’s notion of scaffolding learning with external resources, and the rediscovery of the ideas of Halbwachs and Vygotsky (Michaelian and Sutton, 2013: 4).

However, this resurgence in more social ideas was not yet really underway when WHG were writing. The *Blackwell Handbook of Social Psychology* notes that cognitivism led to a decline in research on groups in social psychology in the 1970s and ‘80s so that ‘by the mid-1980s, the notion of groups as a focus of the field had all but evaporated.’ (Hogg and Tindale, 2001: ix). WGH were writing at exactly this low point for work on groups, so their work can be seen as part of the beginning of a resurgence, rather than as joining a trend that was already dominant. In this respect, their use of the new group mind ideas was an untried hypothesis. Although it was derived from some older ideas such as Bartlett’s, some newer influences such as connectionism and Dynamical Systems Theory were also important, and its application to group memory in modern social psychology was new.

There is an interesting tension between this new kind of group mind theorizing and the dominant cognitivist paradigm, and the next subsection will begin to explore this and its effects on MEMORY.

Cognitivism

Cognitivism was the dominant framework in psychology as a whole at the time that WGH were writing. It applies pressure towards a brainbound concept of memory[63] and in this sense is a contradictory pressure to the new group mind ideas just discussed. Its emphasis on individualism has already been mentioned (see above), as has its dominant model of the computer (see chapter 7). The mind is conceived in terms of software running on the hardware of the brain, and memory is described in terms of encoding, storage and retrieval. (This framework is still dominant today, although its form, particularly with respect to memory, has changed somewhat).

We can see its influence on WGH in the way they structure their work on TMS in terms of transactive encoding, transactive storage and modification, and transactive retrieval (WGH, 1985: 258–263). They describe this as an "analogical leap" from individual to transactive memory systems, but claim that it is reasonable 'as long as we restrict ourselves to considering the *functional equivalence* of individual and transactive memory' (WGH, 1985: 256, emphasis in original). Here we see the influence of the computer metaphor and modern computational functionalism: as long as we stay at the software level the analogy holds, so we can talk about the same processes (encoding, storage and retrieval) in both individual and transactive systems.

At the time WGH were writing, the dominance of cognitivism meant that it was natural to make use of it, and it was helpful for them in showing their move away from the discredited old group mind. The computer model seems as far from the supernatural overtones of the collective unconscious as could be. This is one motivation for working with it, and in the same passage of the paper that the analogy with individual memory is set up, we see WGH emphasise that there is nothing to

[63] Unless you are doing the kind of ecological cognitive psychology discussed in the last case study. Here I am talking about what I there referred to as "traditional cognitive psychology" because this was the dominant framework.

their version of the group mind apart from *individual* memory systems and *observable communicative processes between individuals*. Because of these communicative processes, the result is a system at the group level that is greater than the sum of the individual systems, but one which is nevertheless still marked by the individualism of cognitivism.

Moving away from the old group mind is not the only motivation for adopting the insights of cognitivism however. The cognitive paradigm, and the work of its major early figures (e.g. Neisser in the early part of his career, see chapter 7), resulted in fruitful research in cognitive psychology, as I discussed in the last chapter. However, what is of interest here is social psychology because, although one author (Hertel) is a cognitive psychologist, WGH's work on TMSs is a work in social psychology (the book in which it is included is part of the *Springer Series in Social Psychology*). In particular, what is of interest is the social psychology of groups. At the time WGH were writing, because work on groups was only just making a return, the application of cognitivism to such work was new. Like the new group mind ideas, cognitivism was an untried hypothesis for WGH's social psychology of group memory.

We might speculate that because of the overall dominance of cognitivism, there is some risk of it being retained as an external factor if it fails to acquire a history of success in this kind of social psychology of groups. However, there are some pressures in place working against this. One is the new group mind ideas already discussed; cognitivism and the new group mind are two hypotheses applying pressure in different directions (the former towards an individualist, brainbound concept, and the latter towards a more social, distributed cognition concept). Here we see the factors in the epistemic niche applying pressure to one another as well as to MEMORY, resulting in tensions in the niche which I will explore further below.

Success

I now want to discuss an aspect of MEMORY that is particularly important for this case study, and is importantly linked to cognitivism, namely what counts as success. The standard for successful remembering is crucial because it is the main point on which Pavitt criticises TMS research; recall that he argues that TMS researchers were using the wrong baseline to assess the memory capacities of the groups. Before looking more closely at Pavitt's favoured baseline and notion of success in the next section, it is important to look at the view he was criticising. We therefore need to consider what success was for WGH, and where that notion of success came from, to determine whether it was an internal factor in shaping MEMORY for them.

The original notion of success in WGH was a dyad remembering something that neither individual could remember alone (e.g. by interactive cueing). The baseline here then is an individual's memory. This is the baseline that was used in what is often considered to be one of the first papers on group memory, namely Shaw, 1932. But '[t]he classic (e.g., Shaw, 1932) question, "which is more productive, individuals or groups?" has (appropriately) been supplanted in social psychology with the question suggested by Steiner (1972), "do groups do as well as they could, and when they don't, to what can we attribute their suboptimality?" (Hogg and Tindale, 2001: 109).

I will return to this changing standard in the next section when I discuss Pavitt's criticisms of the baseline used in TMS research, but in this section I want to focus on whether accuracy of memory of any kind is an appropriate measure of success.

The research cited by WGH (Giuliano and Wegner, 1983) is more about the processes involved in the TMS than success, in particular the strategies adopted when dyads are faced with a task. WGH say:

> Because the emphasis of this research was on transactive processes, and not on the accuracy of group retrieval of presented information, it was not necessary to develop these findings in the context of a standard memory paradigm. Until further inquiry is made into the production of unique, integrative knowledge in close relationships, the generality of the observed phenomenon – and its impact on the accuracy of transactive memory – can be anticipated only in broad outline (WGH, 1985: 271).

It seems from this quotation that WGH would be in favour of more work that considers accuracy rather than transactive processes. While they recognise other possibilities than a focus on success, such a focus would be their ideal.

A likely reason for this is the ready measurability of success. Although exactly what should be measured is much disputed, the idea of being able to produce numerical results is a seductive one. Laboratory-based cognitive psychology in the tradition of Ebbinghaus has been a big influence on the field in this respect; on this approach memory, often of simple words or nonsense syllables, is tested by comparison to a correct answer held by the experimenter.

The philosophers whose work I introduced above mention another possible notion of success: consolidating group identity (Theiner, 2013: 78). This aim can be met even when there is collaborative inhibition, i.e. performance could be terrible but the group identity could benefit. Propp also mentions the possibility of a greater diversity of standards for success, as I mentioned when introducing her work, and as I will return to below when considering the epistemic niche for her concept.

If it lacks accuracy, something meeting these other aims would just not count as memory for the laboratory-based cognitive psychologists, therefore if this factor exerted maximal pressure on WGH’s concept, memory would have to be accurate in order for a TMS to be said to exist. Stipulating the accuracy-based notion of success rather than being open to various possibilities of what might constitute success therefore makes it less likely that a TMS can be said to exist. Because the idea of the TMS essentially involves distributed cognition, the accuracy-based standard for success is a pressure away from a distributed cognition concept, and towards a brainbound one.

Although WGH were influenced by the cognitive psychologists’ notion, they did not go so far as to say that anything failing to meet this standard couldn’t be a TMS, as can be seen from the last quotation. This fact can be explained with reference to interaction with other factors in the epistemic niche that apply pressure towards a distributed cognition concept, particularly the new group mind ideas. As part of the new group mind ideas, there was some pressure towards an identity-consolidation-based notion of success when WGH were writing. The social identity perspective in social psychology arose in the 1970s and 1980s. It is cited as one of the things that brought groups back into the limelight in social psychology (see Hogg and Tindale, 2001: 6–7). As we have seen, WGH can be seen as early participants in this return.

I argued that both the new group mind ideas and cognitivism were hypotheses, but that there was tension between them. The easily testable standard of success fits naturally with cognitivism, and the more social new group mind ideas such as Bartlett's earlier work give rise to the alternative constructivist notion of success. Bartlett says of accurate memory:

> [S]o-called "literal," or accurate, recall is an artificial construction of the armchair, or of the laboratory. Even if it could be secured, in the enormous majority of instances it would be biologically detrimental. Life is a continuous play of adaptation between changing response and varying environment. Only in a relatively few cases—and those mostly the production of an elaborately guarded civilization—could the retention unchanged of the effects of experience be anything but a hindrance. (Bartlett, 1932: 15–16).

Bartlett prefers a more constructivist notion of memory that is flexible according to context. This kind of notion has been taken up by many who are arguably moving toward situated cognition frameworks for memory, including the ecological and discursive psychologists discussed in the last chapter.

Arguably the constructivist notion would fit better with WGH's discussion of the TMS in intimate relationships, and I will return to this issue in the subsection on aims below. However, it seems that WGH were somewhat more in favour of an easily testable notion of success. Why was this, and was it an internal consideration? I have already suggested the influence of Ebbinghausian laboratory-based psychology on the field, largely due to its presentation of phenomena such as memory as measurable. I discussed this feature of psychology in chapter 3 of this thesis; mathematization is an external factor imposed on psychology due to its track record of success in physics.

In the first case study, I also discussed easy testability of memory in some detail. I argued that this factor is internal to disciplines that are under pressure to produce results fast (in that case neuropsychologists and neurologists improving the rate of diagnosis to improve patient outcomes). However, for disciplines that primarily aim to explain and predict the world, it is external. Social psychology is a research-based discipline, so easy testability is an external factor here.

Overall therefore, the idea of a standard for accurate memory is an external factor in the epistemic niche of MEMORY for WGH. Mathematization and easy testability were external considerations for social psychology. Although the accuracy-based standard fits well with their cognitivism, another promising alternative was available at the time, in the form of a standard where success was about consolidating group identity, not strict accuracy. The more constructivist concept of memory associated with this alternative standard fits well with the new group mind ideas. Again, tensions within the niche are important here.

Aims

Before moving on to discuss the communication researchers' concepts, I want to look at the aims of TMS research in social psychology. What the research is intended to do is an important pressure on MEMORY in its own right, as well as exerting pressure on other factors we have already considered, creating tensions within the epistemic niche. In the next subsection, I will talk about the aim of bringing respect and a distinct identity to the subdiscipline of social psychology, but here I want to talk about the specific aim of WGH's research, namely to say something about intimate relationships. I will also mention the important aim of aiding work groups and organizations (see Wegner 1986 for other possible applications). Consideration of these applications applies pressure towards taking for granted that there is such a thing as a TMS, and therefore towards a distributed cognition concept of memory, because the TMS essentially involves such a concept.

WGH are writing in a book called *Compatible and Incompatible Relationships* and are trying to explain compatibility via the successful development of a TMS. They say that 'much could be learned about intimacy' by researching cognitive interdependence in close relationships (WGH, 1985: 253). They go on to make the bold claim that '[a] transactive memory is a fundamental component of all close relationships…[T]he potential for transactive memory makes intimacy possible…' (WGH, 1985: 271).

I have already mentioned that the constructivist group-consolidation notion of success might well meet the aim of saying something about intimate relationships better than the accuracy-based notion of success. We can see this particularly in the final major section of WGH's paper, "Transactive Memory and Intimate Life" (WGH, 1985: 271–274). Here, they talk about the influence of a TMS on building and maintaining closeness, and about the impact on an individual after the breakdown of a relationship and its accompanying TMS. In both of these instances, the identity of the group would seem to be important, and in the former case in particular, it seems more important than strict accuracy of memory. If the goal of having a TMS is primarily to maintain the relationship, then constructed memories which contribute to consolidation of the dyad meet this aim better than strictly accurate memories which do not, and which may even contribute to its breakdown.

I have also suggested that the new group mind ideas would fit more naturally with the group-consolidation notion of success than individualist cognitivism, and now I want to add that these parts of the epistemic niche fit well with the practical applications that TMS research aims at. This is brought out best in relation to another aim of TMS research, namely that of aiding groups and teams for businesses and other organisations. The aim of aiding work groups and organizations is cited as one of the reasons for the resurgence in research on groups in the *Blackwell Handbook of Social Psychology* (Hogg and Tindale, 2001: 2) that I discussed under the "new group mind" heading above. The first work on organizations that the Blackwell handbook cites is Weik, 1979, although most of the citations are from the 1990s. As I have said, the resurgence in work on groups was only just beginning with WGH's work in 1985, and the aim of aiding organisations was also new.

Although these specific practical applications were being newly considered by group-focussed social psychology, I have argued that working towards practical applications that are of immediate social importance is an internal factor for all sciences (see chapter 7). Because this consideration has been internalised at a coarser level of grain, the more fine-grained applications considered here are internal considerations.

Consolidating the subdiscipline

A more general and implicit aim of WGH's research is that of consolidating social psychology, by trying to win it both respect, and identity as a distinct subdiscipline.

As an uneasy hybrid of sociology and more individual psychology, it has always been a contested discipline struggling to reconcile competing methods and worldviews from its parent disciplines, and to carve out a distinctive niche of its own.

The pressure created by this goal is important because it in part explains WGH's hastiness to divorce themselves from earlier discredited theories of the group mind that gave social theory a bad name, to latch on to highly respectable cognitivism to a greater extent than to the new group mind ideas, and to favour testable measures of accuracy over a more constructivist notion of memory. Because of its fit with all of these factors, consolidating the subdiscipline can be described as applying pressure towards a brainbound concept of memory. It does not apply this pressure directly, but via its influence on other factors in the epistemic niche that favour a brainbound concept.

It is worth noting that this pressure is not totally unambiguous because, in some respects, responding to the new group mind notion and adopting a distributed cognition concept work towards meeting the aim of forging a distinct identity, for example by making social psychology more distinct from individual psychology. However, there is a danger that in better distinguishing itself from individual psychology, it could risk being less distinct from sociology. More importantly, the prestige and respect attached to cognitivism and easy testability outweigh this possible consideration in favour of distributed cognition with respect to this factor.

The desire to demonstrate that your discipline is respectable and engaged in research that other disciplines have considered worthwhile is an external factor. It is something that is imposed from those other disciplines, rather than having had a successful track record within the discipline in question. However, making your discipline coherent by finding what it can offer that is distinct from neighbouring disciplines can be construed as an internal factor. This is because it is part of establishing any kind of track record of success that a discipline or subdiscipline has an identity of some sort. In order to be an internal factor, this attempt to create and maintain an identity must be carried out with an awareness of its own history. If it is not done in this way, then the discipline would not be basing itself on a history of success, developing according to a rational narrative of its own. Instead, it would have to rely on considerations chosen at random, or those that have proved successful in other subdisciplines, i.e. external considerations. Of course, it is fine to try out things that have been successful in other subdisciplines, but only with an awareness of where they have come from, and therefore the tentative attitude I have described as characterising the use of untried hypotheses as opposed to external factors.

Social psychology does not display sufficient awareness of its own history for this to be an internal factor here. Allport ends his history of social psychology with some words about social psychologists' awareness of their own history. He says that 'journals and textbooks are filled with specific and particular investigations, with a minimum of theorizing' but that this may change:

> Interest in broad theory may again have its day. If so, investigators who are familiar with the history of social psychology will be able to strike out with firm assurance. They will be able to distinguish what is significant from what is trivial, to progress from platitude, and to borrow selectively from the past in order to create a cumulative and coherent science of the future (Allport, 1968: last paragraph of paper).

This is very much in keeping with what I am saying about the role of understanding of their own history in maintaining a successful subdiscipline with its own identity (Allport's words "cumulative and coherent" here are at least analogous in spirit to "successful…with its own identity"). Until this awareness of history is in place, consolidating the subdiscipline is an external factor.

Summary

In summary of what I have said about the epistemic niche for the social psychologists:

Factor applying pressure to concept	**Concept encouraged by this pressure**	**Internal / hypothesis / external**
Rejection of old group mind	Brainbound	Internal
New group mind	Distributed	Hypothesis
Cognitivism	Brainbound	Hypothesis
Success (easy testability)	Brainbound	External
Aims (practical applications)	Distributed	Internal
Consolidating the subdiscipline	Brainbound	External

In this case study, tensions within the niche are of great importance. Certain considerations fit best with certain others, so the pressure they apply to one another ends up being important in determining which apply most pressure to MEMORY. In particular, cognitivism, easy testability, and the aim of consolidating the subdiscipline fit together and all apply pressure towards a brainbound concept. The new group mind (which involves a more constructivist concept of memory with possible broader notions of success) and the aim of providing certain practical applications apply pressure towards a distributed cognition concept. I have argued that a distributed cognition concept is essential to the TMS as WGH define it, so a move in response to the pressures towards a brainbound concept of memory would involve a redefinition of the TMS.

The aim towards practical applications is internal, while the aim to consolidate the subdiscipline is external, suggesting that the former is the better option. This is backed up by the claim that the consideration of easy testability that would help meet the external aim is itself external. Both the new group mind ideas, and the application of cognitivism to group social psychology are untried hypotheses, so they do not help to decide the issue. The concept is rational to the extent that it is shaped by either of them, but non-legitimate because neither has had the opportunity to acquire a history of success.

However, despite this apparent superiority of the distributed cognition concept, there is also pressure towards a brainbound concept from the internal consideration of moving away from the older group mind ideas. A brainbound concept would be rational and legitimate to the extent that it was shaped by this pressure. We therefore cannot reach an unambiguous conclusion here. I argued that WGH employ a distributed cognition concept, and we can see that there was some pressure towards this from internal factors. To the extent that it was shaped by these pressures, their concept was legitimate and rational. There is therefore some evidence of it functioning as an investigative kind. However, we cannot conclude that it was the best or only concept they could have employed under the circumstances.

In the next section, I will apply a similar treatment to the communication researchers.

Communication studies

Many of the factors that apply pressure to the communication researchers' concepts are the same as those affecting the social psychologists' concepts, but some are new, and some of those that are the same affect them differently because they are from a different subdiscipline and were writing several years later (2003).

Previous TMS research

An important pressure on the communication researchers' concept is previous research on the TMS. They are writing some years after WGH first proposed the notion (2003, compared to 1985), so there has been plenty of time for such research to accumulate, and this is a major respect in which they differ from WGH.

This previous TMS research is important because it justifies Propp and Hollingshead and Brandon's assumption that memory can be something possessed by a group. This is part of what allows them to advocate going ahead and investigating the workings of the TMS rather than recommending a baseline to test for its very presence like Pavitt does. They take it for granted as something to have its properties investigated, rather than a hypothesis having its existence investigated. This was perhaps Pavitt's fear when he talked about the dangers of reifying the TMS. Ren and Argote say:

> Interest in transactive memory has accelerated in the past 10 years. At the same time, the proportion of papers that refer to transactive memory systems but in which TMS is not the core topic of the paper has also grown significantly. These are signs that transactive memory systems are being reified, as with other well-known concepts such as absorptive capacity (Lane, Koka, & Pathak, 2006). In other words, as more people become familiar with the concept, they begin to "take it for granted" and use it conveniently to meet individual research needs, with little consideration of the original assumptions and propositions that define the relationships of the construct with other constructs. (Ren and Argote, 2011: 194).

If Ren and Argote are correct in this assessment, then in 2003 when the papers I am analysing were written, this process of reifying would already have been happening. It seems it was something that Pavitt was trying to resist, while the other communication researchers were already beginning to take the concept for granted. This gloss on the debate would explain their comfortable talk of group memory, despite it being what was supposed to be at issue. We can explain Pavitt's resistance in terms of his responding more to pressures towards a brainbound concept, for example cognitivism and rejection of the old group mind ideas.

Taking the TMS for granted is something that came about from previous fruitful research using it. As Ren and Argote say:

> Compared with the antecedents to transactive memory systems, findings are much more consistent about the impact of transactive memory systems on group outcomes. Numerous studies have shown the positive effects of transactive memory systems on a variety of group outcomes such as group learning, team creativity, effectiveness, and member satisfaction (e.g., Austin, 2003; Faraj & Sproull, 2000; Lewis, 2003; Liang et al., 1995; Michinov et al., 2008) (Ren and Argote, 2011: 205).

Research showing the TMS having such positive outcomes translates into success for the TMS framework, because it is the framework that is allowing the positive outcomes to be demonstrated, resulting in useful descriptions and recommendations for practical improvements in team behaviour. The studies cited in this quotation demonstrate successes across a broad range of subdisciplines, most of which date from before the papers I am looking at here (2003). The breadth suggests that responding to the pressure exerted by previous TMS research is not just internal to social psychology, where the concept and accompanying framework was first proposed, but that it has been internalised at a coarser level of grain by this point. It is therefore an internal consideration for communication studies.

This does not mean that the TMS framework and distributed cognition concept that go along with it are necessarily appropriate for the papers I am looking at here however. Previous work on TMS is only one consideration in the epistemic niche for their concept, and it might be that others are internal and apply competing pressure. I will now go on to look at some other considerations to investigate whether this is the case.

Rejection of old group mind ideas

Another important influence, as for the social psychologists, is the pressure to get away from earlier discredited group mind ideas. Pavitt is particularly affected by this, and he thinks that WGH don't go far enough in this respect. The other communication researchers do not mention the old group mind ideas, so it is hard to say whether it exerts any pressure on their concepts. This subsection therefore mainly concerns Pavitt.

Rejecting the old group mind ideas applies pressure away from a distributed cognition concept and towards a more individualist one. As I said when discussing the social psychologists, in this case, pressure away from the distributed cognition concept is pressure towards a brainbound concept, because the kind of work being done is laboratory-based experimental work that does not concern itself with the environment or context in the way that any of the other situated cognition frameworks would do.

Pavitt says that Wegner 1986 was 'purposely attempting to reintroduce the "group mind" notion' (Pavitt, 2003a: 593). As I said above, Pavitt acknowledges that it can be valuable to apply terms usually used about individual cognition to the group level. However, he goes on to say that '[t]his use is, however, inherently metaphorical, and there is a risk that the idea of "group information processing" will become reified' (Pavitt, 2003a: 593).

When I considered this factor for the social psychologists, I examined the reasons for the rejection of older group mind ideas within the subdiscipline, and looked to see whether WGH were responding to this tradition, or trying to distance themselves from a tainted research program for reputation's sake. There is no corresponding history of proposal and rejection of these group mind ideas in communication studies as such. Communication studies is an interdisciplinary hybrid subject, and the rejection of the group mind has a history of success only in some of the relevant subdisciplines. It is therefore not something that communication studies as a whole can take for granted and rely upon.

It seems that this is exactly what they are doing, however. Pavitt does not make any mention of Wegner et al.'s disclaimers about and discussion of earlier group mind notions. Although in a sense he is right that Wegner (1986) seeks to reintroduce the group mind notion in some form, he does not distinguish the influence of what I have called the "new group mind" ideas (reintroduction of the work of Bartlett in particular, see the discussion above) from ideas like Jung's collective unconscious. Wegner himself distinguishes them somewhat more clearly. For example, he shows his awareness of what he is taking from the old ideas and what has changed, when he says:

> Like early theories of the group mind, transactive memory draws deeply on the analogy between the mental operations of the individual and the processes of the group. Unlike early theories of group mind, the new notion of transactive memory benefits from recent advances in the study of the thinking processes of the individual (Wegner, 1986: 185).

Pavitt may disagree that this is a worthwhile research program, but he does not discuss this, instead seeming to assume that anything tainted by group mind notions is to be avoided, despite his avowal that the concepts of "collective information processing" and "transactive memory systems" can have some value (Pavitt, 2003a: 597–598).

Rejection of the group mind is an external consideration imposed on communication studies from social psychology. Although some parts of psychology are part of communication studies, making this seem an odd use of the term "external", the proposal and rejection of the older group mind ideas do not have a history within all of the participating subdisciplines, or the interdisciplinary hybrid. It therefore should not be taken for granted as an established fact, as Pavitt seems to do, but should be treated more in the manner of a hypothesis.

Cognitivism

Another factor which affects the communication researchers as much as it affects WGH is cognitivism. It was still the dominant paradigm when Pavitt, Propp and Hollingshead and Brandon were writing (and it still is). It is likely that Hollingshead and Brandon were particularly influenced by it because Hollingshead has a background in psychology (she has a B.A. in Psychology and an M.A. and Ph.D. in Social Psychology), but the impact of cognitivism has been so great that it is a big influence on everybody. As I said in the section on the social psychologists, it applies pressure towards an individualist, brainbound concept of memory.

I argued above that for WGH, cognitivism was an untried hypothesis. At that stage it did not have a history in social psychological research on groups, because such research had not been taking place since cognitivism became the dominant paradigm, and was only just beginning a resurgence. As I said in the first subsection concerning the communication researchers, by the point they were writing, TMS research had a history, so such group work was more established. TMS researchers had continued to make use of cognitivist ideas (for example speaking of transactive encoding, storage and retrieval, see Ren and Argote, 2011), and using this framework had proved successful (see subsection above on previous TMS research). Cognitivism was therefore internal to TMS research.

To be internal to a particular research program is to be internal at a finer level of grain than the subdiscipline. Communication studies as a subdiscipline is composed of elements of many different other subdisciplines, not all of which make use of cognitivist ideas; cognitivism is therefore not internal to communication studies. It seems that as TMS researchers, Pavitt et al. are entitled to take cognitivism for granted, but as communication researchers, they are not. What can we make of this?

The relevant framework here is the coarser-grained framework of communications studies, rather than the finer-grained one of TMS research. This is because the point of the papers in question here is not to carry out research into the TMS, but to talk about whether and how communication researchers should contribute to such research. Given that cognitivism is internal to TMS research, ignoring it entirely would be irrational, but when engaged in the kind of project being undertaken here, the researchers should not be taking for granted things that are external to their subdiscipline, like the cognitivist paradigm with its computer metaphor.

However, the evidence suggests that they do take it for granted. In particular, they seem to have forgotten that much of the cognitivist paradigm is metaphorical (the computer model, talk of "storage" etc.). I quoted Pavitt above as saying of the extension of individual mental terms to groups that '[t]his use is, however, inherently metaphorical, and there is a risk that the idea of "group information processing" will become reified.' (Pavitt, 2003a: 593). He does not seem to recognise here that "information processing" is already a metaphor (part of the computer metaphor). In addition, all the papers concerned here use the cognitivist encoding, storage, and retrieval model of memory without question.

It is one of the advantages of interdisciplinary collaborations such as communication studies that multiple frameworks can be employed to look at the same problem from different perspectives. This is one way in which things that are taken for granted in one subdiscipline (internal considerations in their framework) can be brought into question in a productive way. (Some things must be taken for granted for research to proceed, so we need internal factors, but nothing in science is ever put beyond question for ever). This productive questioning is something that communication studies researchers from subdisciplines other than social psychology could offer with respect to TMS research. The communication researchers I am discussing here seem to recognise this to some extent; they consider what distinctive elements communication studies could bring to the table. However, they do not sufficiently recognise that cognitivism is a presupposition they should raise questions about.

A better awareness of their own history would help the communication researchers here, by making them more explicitly aware of the metaphorical nature of cognitivism. There is some evidence suggesting that communication studies as a discipline has a good awareness of its own history, and I will talk about this below in the subsection “consolidating the subdiscipline”. However, the history of cognitivism is not part of the history of communication studies as such, so this historical self-awareness is not of much help to them here.

The communication researchers are taking cognitivism for granted and relying on it, but not because of a history of success within their own subdiscipline’s framework, so it is an external factor for them.

I argued above that Propp shares Pavitt’s conviction that the group mind should remain a metaphor, and thus construed the TMS differently to WGH. Individualist cognitivism is a major factor applying pressure in the direction of doing this. However, Propp’s concept does have some flexibility, and this can be explained with reference to other factors in the epistemic niche, as I will explain below. Hollingshead and Brandon have a distributed cognition concept of memory, and cognitivism would apply pressure away from this. Therefore, they have resisted this external factor, despite Hollingshead’s background in psychology. This can be explained in terms of competing pressures in the epistemic niche (see summary below).

Aims

As I did for the social psychologists, I want to look at the aims of the research being carried out. They are part of what the concept of memory is needed to do. One major aim, as for the social psychologists, is to help businesses by researching work groups and teams. This factor applies pressure towards taking for granted that there is such a thing as a TMS, and therefore towards a distributed cognition concept of memory, because the TMS essentially involves such a concept.

I have argued previously that providing practical applications that are important for society is an internal consideration because providing such applications is part of what sciences exist to do. This consideration has therefore been internalised at the relatively coarse level of grain that is science in general. However, the situation is not quite so simple for communication studies, because it is a hybrid subject consisting of humanities subdisciplines as well as sciences. I therefore cannot rely on internalisation at a coarser level of grain because communication studies does not fall neatly under a single category at such a level. More needs to be said about the history of the discipline itself.

Communication studies had pragmatic aims from its inception, or at least from the period of its institutionalisation as a fully-fledged discipline with specialised departments, journals and conferences. Although there are some disagreements about the origins of the discipline, many agree that the Second World War (and later the Cold War) were very important in determining the interests and available funding for research. The media and mass persuasion were of particular interest (see Wahl-Jorgensen, 2004: 551).

> As Simpson [1994] argued, the psychological warfare effort of the American government during the Second World War shaped the communication research community and determined "which of the competing scientific paradigms of communication would be funded, elaborated and encouraged to prosper" (p.3). (Wahl-Jorgensen, 2004: 552).

Wahl-Jorgensen argues that the origins of the discipline are to be found even earlier, in Chicago from the university's inception in 1892. He says:

> Chicago is known as the birthplace of pragmatism and, more generally, of an American social science (cf. Bulmer, 1984; Abbott, 1999). This was a social science engaged with tangible problems of urban life and determined to make a difference (cf. Bulmer, 1984, p.23). To the turn-of-the-century pragmatists, newspapers and other media were central to this process, as they went about studying how the ethnic groups who shared Chicago sought to survive and make sense of life there.
>
> As such, mass communication and its role in society had been on the research agenda at the University of Chicago since its inception (Wahl-Jorgensen, 2004: 550).

Concern with practical applications was fruitful for Communication Studies, allowing it to get off the ground as a discipline, acquire funding and become institutionalised. It was thus internalised. The application of helping work groups and teams is a part of this narrative, and so it is an internal factor.

Consolidating the subdiscipline

Another important aim of the research is, as for the social psychologists, consolidating their own discipline. This is arguably even more pressing for communication studies because it is an even newer discipline, and an even more complex interdisciplinary hybrid (of psychology, sociology, anthropology, journalism, political science and English, among others).

The communication researchers' response to this consideration can be seen in their insistence on more focus on the communication taking place between individuals in a TMS. This focus would give them something particular to contribute to the research that is distinct from psychologists' contributions. This applies pressure towards taking the TMS seriously as something that really exists, and therefore towards a distributed cognition notion of memory.

I said above that the desire to acquire and maintain a distinctive identity can be an internal factor if the discipline has sufficient awareness of its own history. In general, communication studies has a good awareness of its own history, so this factor is internal for them. Since its inception, the discipline has been concerned with its own history and this has proved fruitful in terms of allowing them to get funding and become institutionalised, as Wahl-Jorgensen explains in this quotation:

> The roots of mass communication scholarship in the U.S. are by no means a new research topic. Ever since the 1980s, disciplinary genealogies have become both fashionable and necessary. In the face of budget cuts, it is increasingly important for each discipline to justify its own existence by defining its core, its boundaries, its unified and identifiable projects, and its myth of origin. For a young and untenured discipline such as ours, this process has been particularly crucial. As Carey (1996) pointed out:
>
>> The history of mass communication research is a recent literary genre, albeit a minor one: a self-conscious creation (and now an endless recreation)…the [historical] narrative that emerges serves ultimately a variety of purposes: principally to focus, justify, and legitimate a 20th-century invention, the mass media, and to give direction and intellectual status to professional teaching and research concerning these same institutions (p.21).

> Histories, in imposing a unified narrative of our origins, organize and structure our activity by locating our roots (cf. Wahl-Jorgensen, 2000). As Peters (1986a) has put it, historical accounts have as their subtext the "transformation of communication research from an intellectual to an institutional entity" (p.537). (Wahl-Jorgensen, 2004: 548).

The fact that historical papers such as Wahl-Jorgensen's are published in the *Journal of Communication* shows the continuing relevance of their history for the discipline.

Attempting to consolidate the discipline in an historically informed way is internal, even if it is a factor that researchers do not always respond appropriately to. Not every communication researcher will be appropriately aware of their history on every occasion (for example I have argued above that Pavitt fails to be with respect to the older group mind ideas, and all the communication researchers fail to be with respect to cognitivism). This may be explained with reference to competing factors in the epistemic niche (particularly in this case, see summary below), and with reference to lack of fit between concept and niche.

Success

Another factor affecting the communication researchers as well as the social psychologists is concern with a particular notion of success. This is a major point of contention in the debate I am considering. Pavitt has a different standard for success compared to the other communication researchers, so I will deal with his paper first, then move on to talk about Hollingshead and Brandon and Propp.

Recall that Pavitt's favoured measure of success is comparison with a baseline calculated using Lorge and Solomon's Model B. The model is a prediction of the performance of a nominal group (an aggregate of the performances of the individuals in the group assuming effective pooling of answers but no other interaction).

Above, I quoted the *Blackwell Handbook of Social Psychology* as saying that '[t]he classic (e.g., Shaw, 1932) question, "which is more productive, individuals or groups?" has (appropriately) been supplanted in social psychology with the question suggested by Steiner (1972), "do groups do as well as they could, and when they don't, to what can we attribute their suboptimality?" (Hogg and Tindale, 2001: 109).

Pavitt is effectively suggesting a more specific version of Steiner's question, cashing out "as well as they could" in terms of Lorge-Solomon Model B. This stringent standard that must be met before a TMS can be said to exist makes it less likely that a group can be said to remember, i.e. less likely that a distributed cognition concept of memory will be appropriate. It therefore applies pressure towards a brainbound concept of memory.

Lorge and Solomon's paper was written in 1955 in *Psychometrika*. Lorge was a psychologist and Solomon a statistician. Their model has been used in social psychology, including to analyse the results of Shaw (1932), usually cited as the classic experiment on group problem solving (Lindzey and Aronson, 1954/1968: 236). Their model therefore has a history in social psychology, so Pavitt's suggestion that WGH should have used it is not unprecedented in *their* field (although this does not make it internal to his).

As I argued was the case for WGH, part of the reason for Pavitt's choice of measure of success is easy testability. Pavitt takes this even further than the social psychologists; a mathematical model doesn't even need an experimenter, unlike comparison of performance of putative TMSs with individuals or nominal groups in the laboratory.

For the social psychologists, I argued that this concern with easy testability was an external factor, as was a desire for mathemetization brought about by the success of quantitative methods in physics. For the communication researchers, easy testability is similarly an external factor because, like the social psychologists, they are a research discipline not under particular time pressure to produce applications (unlike e.g. medicine or law). Mathematization on the other hand has a different history in communication studies than in social psychology, so a little more needs to be said about this.

Communication studies has its origins in so many disciplines that there are many different methodological traditions in its history. Mathematical methods were an important part of the mix right from the start however. Fiske says in his *Introduction to Communication Studies*, 'Shannon and Weaver's *Mathematical Theory of*

Communication (1949; Weaver, 1949[…]) is widely accepted as one of the main seeds out of which Communication Studies has grown.' (Fiske, 1990: 6). Shannon and Weaver's work sets out a quantitative model in information theory, developed during the Second World War in Bell Telephone Laboratories. Mathematical methods have been fruitful in the past, so they are internal to communication studies.

The combination of what I have said about easy testability on the one hand, and mathematization on the other, means that it is legitimate to use a mathematical standard, but not just because it is easily testable. In this case, looking at tensions within the epistemic niche is instructive. The question is whether a mathematical standard could meet the aims of the research. The aims I have looked at here include practical applications such as helping work groups and teams, and consolidating communication studies as a subdiscipline. I argued that both are internal, and both are pressures towards a distributed cognition concept. As part of working towards both of these aims, the communication researchers are working to say something about the communication between the individuals in a TMS. A measure of success like Lorge-Solomon Model B is not the best way to meet this aim because it is a measure of remembering *well*, not just of remembering and the mechanisms by which it takes place in a group.

The philosophers' discussion of TMS is helpful here. Tollefsen et al. say in a footnote:

> Incidentally, it occurs to us that it may not be the best way to test the existence of a system by whether it performs better or worse than some other system; the question is whether two people indeed act as a transactive system, and this seems to be separate from whether they are performing well. (Tollefsen et al., 2013: 56).

This is a helpful critique to apply to Pavitt's standard for success, and highlights the tension between it and the aims of the research. There could still be much to say about how the TMS functions in terms of the communication between individuals, even if there is collaborative inhibition (the group performs worse than the aggregate of its members). The use of a baseline like Lorge-Solomon model B would insist that in this case there could not even be a TMS. Therefore, although mathematization is internal to communication studies, Pavitt's standard for success does not fit well with other parts of the epistemic niche, in particular the aims of the research, and is likely to have come about as a response to the external consideration of easy testability.

This gets to the major difference between Pavitt and the other communication researchers. For Pavitt something must meet a minimum level of success to count as a memory; he argues that the systems considered do not do so, so they are not group memories. For Propp and Hollingshead and Brandon, the groups are assumed to have a TMS that they then advocate a particular way of investigating. If they are right, an appropriate baseline may not even be needed; it might be that another notion (e.g. memory as consolidating group identity as discussed above) is more appropriate.

Propp is relatively explicit about this. As I quoted above, she says:

> Based on previous research, it is possible that even the study of specific interaction patterns under optimal conditions may fail to substantiate the existence of the assembly bonus effect as it is currently defined. It is possible that the definition of improved performance might need to be conceptualized more broadly to account for the presumption of groups' superiority in organizational theory and practice. This definition of performance might include other group outcomes that have been found to make groups superior, such as representation of diverse goals and values, distribution of responsibility for decisions, and increased commitment to an action (Gigone & Hastie, 1996). (Propp, 2003: 604–605).

Such a redefinition would change the standard for what it is for a group to be said to remember, and so would represent a change in MEMORY. This is a change that makes it easier to say that a TMS exists, i.e. that a group can be said to remember, and so is a pressure towards a distributed cognition concept.

Neither Propp nor Hollingshead and Brandon stipulate a particular standard that must be met for success. The main point of their papers is to argue that Pavitt has been too hasty, that we should not have such stringent standards as to what constitutes a TMS, and to encourage further research, particularly from communication studies. In their openness to continuing to look for positive effects of groups remembering together, they are treating any standards for success they suggest, such as those mentioned in the last quotation from Propp, as untried hypotheses.

Summary

In summary of what I have said about the epistemic niche for communication researchers:

Factor applying pressure to concept	**Concept encouraged by this pressure**	**Internal / hypothesis / external**
Previous TMS research	Distributed	Internal
Moving away from group mind (mainly Pavitt)	Brainbound	External
Cognitivism	Brainbound	External
Aims (practical applications)	Distributed	Internal
Consolidating the subdiscipline	Distributed	Internal
Success	Brainbound for Pavitt; Distributed for H&B and Propp	External for Pavitt; Hypothesis for H&B and Propp

Of the factors in the epistemic niche looked at here, those that apply pressure towards a brainbound concept of memory have all been found to be external. These include cognitivism, Pavitt's standard for success, and moving away from the old group mind ideas, which is mainly a factor for Pavitt.

The factors applying pressure towards a distributed cognition concept are previous TMS research, the aim to help work groups and teams, the aim of consolidating communication studies as a discipline, and Hollingshead and Brandon and Propp's broad notions of success. All of these are internal apart from the last, which is an untried hypothesis. I argued that only Hollingshead and Brandon have a distributed cognition concept. We can now see that this concept is both legitimate and rational.

There is therefore evidence of a situated cognition concept (in this case distributed) functioning as an investigative kind in communication studies. Pavitt and Propp both have brainbound concepts, and this can be seen to be a result of responding to pressure from external considerations. Their concepts are therefore non-legitimate and irrational to the extent they are shaped by the factors considered here. Propp's concept has some degree of flexibility, indicating that she is more open to the influence of the internal factors than Pavitt. Her openness to more flexible notions of success is particularly interesting. Although this is only an untried hypothesis at this stage, it may indicate a future move toward a more distributed cognition concept. Pavitt does not share that openness, and there are particular tensions in the epistemic niche for him, in particular between his notion of success and the aims of the research. He is also particularly affected by the external pressure to move away from the old group mind ideas, which he does not sufficiently separate from the new group mind ideas I discussed in relation to the social psychologists.

Both Pavitt and Propp would do better to move in the direction of a distributed cognition concept of memory if they want to talk about TMSs, and this is not surprising, given that I argued that a distributed cognition concept is essential to the TMS framework as it has grown out of the work of Wegner et al. It is Wegner et al.'s framework for the TMS that participants in the debate cite, but the internalisation of cognitivism into TMS research suggests that they cannot really be using it, because cognitivism applies pressure towards a brainbound concept of memory. What a TMS is, and therefore what it means for a group to remember, has changed since the work of WGH. TMS has come to mean something different for researchers such as Pavitt and Propp, but according to what I have argued here, there are reasons to think that the concept of memory involved in this change is non-legitimate and irrational.

Conclusion

I have argued that the social psychologists working on TMS have a distributed cognition concept of memory. Of the communication researchers Pavitt, who criticises the social psychologists, has a brainbound concept. Of the communication researchers who criticise Pavitt, Propp shares his brainbound concept but with more openness to changing it based on the results of future research, while Hollingshead and Brandon share the social psychologists' distributed cognition concept.

Examination of some of the main factors in the epistemic niches shaping these concepts has revealed that, for the social psychologists, there was some pressure from internal factors towards their distributed cognition concept. To the extent that it was shaped by these pressures, their concept was legitimate and rational. There is therefore some evidence of it functioning as an investigative kind. However, we cannot conclude that it was the best or only concept they could have employed under the circumstances because there is also some pressure from internal factors towards a brainbound concept. Cognitivism also applies pressure towards a brainbound concept, and was a hypothesis at the time WGH were writing. If it were to go on to be successful (which I argued in the section on communication studies is exactly what happened in TMS research), then there would come to be another internally generated pressure towards a brainbound concept. This would create further tensions within the epistemic niche because cognitivism fits poorly with some other internal factors, and it is too early to say by the date I am looking at here whether this would play out in favour of a distributed cognition concept.

For the communication researchers, there was pressure from internal factors towards a distributed cognition concept, and all the pressures towards a brainbound concept were from external factors. Hollingshead and Brandon’s distributed cognition concept was therefore legitimate – i.e. functioning as an investigative kind – and rational. Pavitt and Propp’s brainbound concepts were found to be largely non-legitimate and irrational, although there is some evidence of Propp moving towards a distributed cognition concept. This case study has therefore not only provided evidence of a situated cognition (in this case distributed cognition) concept functioning as an investigative kind, but also provided evidence suggesting that situated cognition should be employed to a greater extent.

9. Conclusions

Specific conclusions

The question addressed in this thesis has been: “Are there any situated cognition concepts of memory functioning as investigative kinds in the sciences of memory?” – what I have been referring to as “the situated cognition question”. The short answer to this question is “yes”. I will give a slightly more qualified and fleshed out answer below, after briefly reviewing how we got there.

Whether or not a concept is functioning as an investigative kind is the epistemological question of whether it can support fruitful science. I have argued that how concepts are individuated depends on the level of analysis, and that the level of grain at which the framework is viewed depends on the analyst’s aims. I have focussed primarily at the subdiscipline level of grain, because MEMORY often varies between subdisciplines, and the communication difficulties stemming from this variation are part of the problem I have been addressing.

I have used a case study-based conceptual ecology to determine what concepts were in play in particular projects in various subdisciplines of memory science, and what the epistemic niches for those concepts were. An historical study of the epistemic niches then revealed whether the major factors in those niches were internal or external, or untried hypotheses. Many of the sciences of memory are immature, so in some of my case studies, the concepts of memory were found to be weakly constrained by internal factors, with many untried hypotheses. However, I claimed before embarking on the case studies that it is important to try to begin to answer the situated cognition question because of the importance of memory research, and in fact I have found enough internal and external factors to draw some substantial conclusions about the legitimacy and rationality of particular concepts. Recall that to the extent that a niche is made up of internal factors, the concept it shapes is legitimate, and to the extent that it is made up of internal factors and untried hypotheses, the concept is rational. Concepts that are functioning as investigative kinds are legitimate concepts.

In line with Shapere's rejection of imposing metascientific standards, I have tried to treat the scientists' work as much on its own terms as possible, looking at how different aspects of their frameworks fit together and apply pressure to MEMORY. I made use of Khalidi's interpretationist strategy (proposed as a solution for the concept individuation problem for the Theory Theory), which explains how we can analyse concepts in the framework despite standing outside the framework ourselves. This strategy involves treating the concept users as intentional agents, and assuming they have at least a minimal level of rationality. Despite this standard for what it is to be rational being shared at the coarse level of grain that is epistemic enquiry in general (see chapter 4), I did not assume that this is enough of a constraint to rule out the possibility of pluralism of frameworks at finer levels of grain. In particular, I left open the option that multiple different concepts of memory may be functioning as investigative kinds in the case studies I investigated.

This kind of pluralism is in fact what has been found. Although I have only looked at three case studies, and the analysis has only focussed on a few important features in the epistemic niches for each, the case is already more than suggestive. Further work improving the richness of the historical analysis of these case studies, and addressing others, could paint a somewhat different overall picture, but it would be surprising from what has been revealed here if that picture did not include multiple different situated cognition concepts functioning as investigative kinds in the sciences of memory.

In the case study on locked-in syndrome and brain computer interfaces, I found that there was some variation even *within* the subdisciplines I looked at, with some neuropsychologists and neurologists having a brainbound concept, and others an embedded cognition concept. Most of the philosophers I looked at had flexible concepts that are open to interpretation from different perspectives including the situated cognition perspectives, but one had an extended cognition concept. I found that when just the internal considerations were taken into account, the neuropsychologists and neurologists had some pressure towards a brainbound concept, and some pressure towards an embedded cognition concept, particularly when carrying out work that aims to take patients' own views into account. The philosophers had pressure towards having flexible concepts. They also had some pressure towards accepting a situated cognition concept outright (particularly extended cognition) to the extent that their work tries to have an impact in fields outside philosophy. Therefore all of the various concepts in play here were found to be functioning as investigative kinds in the subdisciplines considered, although the evidence was considerably stronger for the brainbound and flexible concepts, and more tentative for the embedded and extended concepts. I argued that therefore these subdisciplines should remain pluralist with respect to their concept MEMORY for the time being, although it is too early to say whether any of them should eventually adopt a unified situated or brainbound perspective on memory, or retain their plurality of concepts indefinitely.

The case study on constructing memory in political scandals revealed that Neisser (an ecological cognitive psychologist) was working with an embedded cognition concept, and Edwards and Potter (discursive psychologists) were working with a distributed cognition concept. Neisser's concept was shaped by a variety of competing factors, some applying pressure towards an embedded cognition concept, some enacted and some brainbound, with internal considerations in all three of these categories. There is evidence that his embedded cognition concept was functioning as an investigative kind, but I argued that the competing pressures taken together should result in pluralism, with multiple kinds of concept and their accompanying frameworks in use side by side. Discursive psychology was a new subdiscipline with many untried hypotheses, so it was too early to tell whether the discursive psychologists' distributed cognition concept would pick out an investigative kind. However, it was rational to use such a concept, and this is suggestive. Although there was no way of knowing whether it would go on to be a successful concept, there was no reason to think that it would not. I therefore found no reason to stipulate that the science would be best served by sticking to brainbound concepts, as some critics of the situated cognition approaches might recommend.

The case study on Transactive Memory Systems revealed that the social psychologists had a distributed cognition concept, but there was some variation among the researchers in communication studies. Pavitt had a brainbound concept, Hollingshead and Brandon had a distributed cognition concept, and Propp had a brainbound concept, but with some degree of flexibility. For the social psychologists, there was some pressure from internal factors towards their distributed cognition concept, and to the extent that it was shaped by these pressures, their concept was legitimate and rational. There was therefore some evidence of it functioning as an investigative kind. However, I argued that we could not conclude that it was the best or only concept they could have employed under the circumstances because there was also some pressure from internal factors towards a brainbound concept. Cognitivism also applied pressure towards a brainbound concept, and was a hypothesis at the time WGH were writing. If it were to go on to be successful (which I argued later in that chapter is exactly what happened in TMS research), then there would come to be another internally generated pressure towards a brainbound concept. I discussed tensions within the epistemic niche at the time of the case study, and further tensions that would be created by this development, concluding that it was too early to say by the date of the case study whether the situation would play out in favour of a distributed cognition concept. For the communication researchers, there was pressure from internal factors towards a distributed cognition concept, and all the pressures towards a brainbound concept were from external factors. Hollingshead and Brandon's distributed cognition concept was therefore rational and legitimate, i.e. functioning as an investigative kind. Pavitt and Propp's brainbound concepts were found to be largely non-legitimate and irrational, although there was some evidence of Propp moving towards a distributed cognition concept. This case therefore not only provided evidence of a situated cognition (in this case distributed cognition) concept functioning as an investigative kind, but also provided evidence suggesting that situated cognition concepts should be employed to a greater extent.

Overall then, I have found evidence in recent research in the sciences of memory of brainbound, embedded cognition, extended cognition, and distributed cognition concepts of memory, as well as flexible concepts that are open to interpretation from different perspectives, all functioning as investigative kinds. I also found internally generated pressures towards concepts that were not in play in the cases where those pressures were found, and these included pressures towards brainbound and enacted cognition concepts. I also found a distributed cognition concept in play in a new subdiscipline (discursive psychology) that was potentially set to develop into an investigative kind in the future of that subdiscipline, and a brainbound concept that the evidence suggested would have been better replaced by a situated cognition concept (for the communication researchers in the TMS case study).

At the level of the sciences of memory, this suggests that pluralism with respect to MEMORY is the best option, at least for now. There are multiple different kinds of concepts of memory, many of them situated cognition concepts, functioning as investigative kinds in these sciences. This is not just a brand new development; the case studies span dates from the early 1980s to today, and the roots of their use of situated cognition concepts go back much further (e.g. Bartlett, 1932; Gibson's work from the 1960s and 1970s). This offers considerable support to the case for situated cognition frameworks.

Although some variation was found within subdisciplines, the greatest variation in MEMORY is found between subdisciplines, and this creates many of the communication difficulties I outlined as part of the problem. The recommendation of pluralism would seem to exacerbate this problem and associated worries about the fragmentation of the disciplines involved. So how are we to deal with pluralism? Here I can only make some speculative comments in answer to this question, but they are important speculations which help to situate the more specific conclusions drawn here in current and future scientific practice.

Broader context of these conclusions

In chapter 1, I quoted Clark's discussion of the debate between HEC (the Hypothesis of Extended Cognition), HEMC (the Hypothesis of EMbedded Cognition), and his solution HOC (the Hypothesis of Organism Centred cognition). Clark says:

> HEC, HEMC, HOC? We should not feel locked into some pale zero-sum game. As philosophers and as cognitive scientists, we can and should practice the art of flipping among these different perspectives, treating each as a lens apt to draw attention to certain features, regularities, and contributions while making it harder to spot others or to give them their problem-solving due. (Clark, 2008: 139).

This recommendation of flipping between different perspectives seems apt for the situation I have uncovered in this thesis. Pluralism does not necessarily mean fragmentation, but there is a risk of fragmentation occurring because of the communication problems pluralism creates. My concept-centred approach in particular highlights the difficulties scientists face in communicating about memory – they do not even share a concept of the target phenomenon. If individuals and the subdisciplines of which they are a part were better at this flipping between perspectives, this problem could be alleviated. Doing so would involve fluency in multiple frameworks, which is a challenge, but which might be achieved with some changes in how scientists are educated and how their enterprise is thought of. There are already some relevant suggestions in the literature, for example Sternberg and Grigorenko's advice to individuate fields by the phenomena they study, rather than the methods used (memory or emotion, rather than social psychology or clinical psychology) and to therefore educate psychologists in multiple methodologies (Sternberg and Grigorenko, 2001).

Why should we undertake this difficult task? I said in chapter 3 that pluralism is becoming increasingly popular, both with respect to cognitive science in particular, and science in general. I went on to argue that the cognitive and social sciences should not use unification for its own sake as a criterion, because it is an external consideration for them. This does not rule out the future internalisation of unification, but with such a pluralist turn beginning to take place in the literature, the prospects for this look poor.

There are perhaps reasons to be optimistic about this. I also said in chapter 3 that pluralism is not just an increasingly popular position, but also a fruitful one. I mentioned there the argument in the feminist philosophy of science that theories and models are currently partial and goal-directed according to the interests of particular groups, and that pluralism could allow traditionally marginalised voices to be heard (Longino, 1996: 275–277).

As we have seen, pluralism about frameworks also involves pluralism about the concepts embedded in those frameworks. This entails a pluralism about kinds or categories, which also has implications for feminist theory. Dupré describes how the anti-essentialism of a view like his promiscuous realism about kinds has positive implications for how we think about sex and gender (Dupré, 2002: 175–195), homosexuality (Dupré, 2002: 156), and disability (Dupré, 2002: 67). Pluralism about natural kinds, or kinds of kinds, also allows us to study, in terms of kinds, some things that cannot be assimilated to the essentialist natural kinds program. For example Haslam makes a pluralist argument that different mental illnesses can best be described as falling under different kinds of kind (Haslam, 2002). Machery notes that according to a pluralist conception of natural kinds (in this case the pluralist interpretation of homeostatic property cluster theory), artefacts can be classed as natural kinds, and therefore a proper object of scientific research, as they are in Paleoanthropology (Machery, 2009: 234). Such examples suggest that pluralism is not only increasingly popular, but advantageous.

I have speculated that what is needed is a change in the way we think about science in favour of both individuals and subdisciplines flipping between different perspectives. I also suggested in chapter 3 that the road to maturity for the social and cognitive sciences may involve internalising pluralism, if such an approach was to prove successful. Clark also links the idea of flipping between different perspectives to maturity in the conclusion of his *Supersizing the Mind*: 'The appeal to embodiment, if this [combining new perspectives on the mind with old representational and computational ones] is correct, signals not a radical shift as much as a natural progression in the maturing of the sciences of the mind.' (Clark, 2008: 219).

The cognitive and social sciences appear to be maturing in a pluralist manner in the middle of a pluralist turn in the philosophy of science. The situation with respect to situated cognition perspectives, and the work I have done regarding it in this thesis, is a good example of this. Even if the physical sciences were to make a pluralist turn in the future, they have recently undergone a long monist phase, and they reached maturity as monistic sciences.[64] The situation is different for the cognitive and social sciences, and this gives us a perhaps unique opportunity to see a science mature in an intellectual climate of pluralism.

[64] It might be that the physical sciences would have matured somewhat differently in a more pluralist intellectual climate like that developing today. Alternatively their subject matter may be sufficiently different that this would not have been the case.

However, philosophy of science is not just about observing science. As well as watching this narrative develop, we can contribute to its smoother unfolding. The kind of flipping between different perspectives I have talked about here is something that much of science is still ill-equipped to handle, and in the future we can make recommendations to change this. These are large projects. There is much to be said about the future of subdisciplines, interdisciplinarity, the training of scientists, the role of integration as a kind of unification within pluralism,[65] and the relationship of a pluralist science to society and culture more broadly. These are tasks with which the philosophy of science should concern itself in the immediate future.

Sources of reflexivity

One interesting and inescapable aspect of work in any human or social sciences or humanities discipline is its inherent reflexivity. I want to end by briefly considering this aspect of my project. The broad question is whether my methods can be applied to my own thesis. Thus far, I have only talked about how they might apply to the *sciences* of memory. Even the philosophers I considered in chapter 6 were doing a type of philosophy that would be widely classed as cognitive science. The same could not be said of my own work.

It seems clear that the main themes of the approach could be applied to philosophy, but not necessarily the specifics. For example, the idea of research frameworks would generalise, but the kinds of theories and practices the frameworks consist of would differ. Trying to apply a framework account here would be a considerable undertaking in the philosophy of philosophy, but I will try to say some more specific things below by addressing questions about the epistemic niches for my concepts, pluralism in philosophy, and the consequences of my work for interdisciplinarity in philosophy.

It is worth noting that if my methods are applicable outside the sciences, they may also be applicable to other philosophy of memory, for example in traditional analytic philosophy of mind, and in continental philosophy, and to other humanities studies

[65] For example, Brigandt suggests that we don’t need unity for its own sake, but may need it for certain specific problems, and there is much more to be said about this (Brigandt, 2010a; 2011: 17; 2012: 82, footnote 5).

of memory more broadly (for example in history or literature). This would offer considerable scope for expanding the kind of project I have undertaken here.

- What is the epistemic niche for my concept of memory (and for other concepts I employ)? What sorts of factors does the dynamic framework for my research consist of?

  If the idea of a dynamic framework can be applied beyond the sciences, it should be applicable to my own research. It would seem fair to say the same of my concept of memory as I said of the philosophers in chapter 6: My concept is flexible in the sense of being open to interpretation from different perspectives, because which perspective to employ is what is in question.

  What kinds of things would be in the epistemic niche shaping this concept? Openness to questioning concepts is an important factor (compare chapter 6). My aim was to say something useful about how scientists and philosophers of science should proceed with respect to the situated cognition question, so this aim would also be part of the framework. Looking to my framework more broadly, the very situated cognition perspectives I am investigating are also important. The dynamic framework viewed at the subdiscipline level of grain could be seen as a shared cognitive system, and therefore analysed in terms of distributed cognition; alternatively, because of the role played by methods, apparatus etc. as well as theory, my methods could be seen from an embedded, extended or enacted cognition perspective.

  Deciding which of these approaches is best for my work, or whether remaining flexible between them is the best option would involve tracing the history of the factors in my framework. I cannot undertake this task here; however, it is clear that my work, like any in philosophy, is embedded in various literatures that have their own histories (in this case some of the most obvious examples are the literatures on situated cognition, concepts, the role of history of science, pluralism, and the post-Kuhnian literature on paradigms and their kin). As philosophers, we do tend to concern ourselves very much with our relationships to these literatures, and the work I have done here would suggest that we are correct to do so, but not to establish the prestige of our pedigree, rather to demonstrate the history of success attached to factors in our frameworks, thus showing that they are internal to our discipline, subdiscipline, or research project.

- Is philosophy a monistic or pluralistic discipline, and which should it be?

  I argued in chapter 6 that philosophy has internalised plurality because over time it has proved more successful than unity. In particular, flexible concepts have allowed us to see things from others' perspectives and therefore engage in debate more effectively. While I think this is true of our implicit framework, many philosophers have remained more explicitly committed to a kind of monism. We seek *the* correct or best account, not just *an* account that works well for our particular purposes. If what I have said about the success and internalisation of plurality is right, it might be that that ought to change. A genuinely pluralist philosophy would harness argument to eliminate frameworks that do not work, and to work out which frameworks are best for which purposes, but it would no longer consider such arguments a fight to the death for one framework or the other.

  A pluralist approach to the project carried out in this thesis would allow that other ways of construing science than the dynamic framework account could be useful for other purposes, and this is something I have allowed here. I hope the work I have done shows that the framework approach is fruitful for my purposes, but this does not rule out other kinds of account.

- What can my approach say about communication within philosophy, and interdisciplinarity, for example between philosophy and science?

  If we apply the framework account to philosophy, we can see that we should expect communication problems, both within the discipline, and between philosophy and other disciplines we would like to work with (for example science, history). Ways to alleviate this might include the kind of perspective-flipping approach I suggested above for science. In fact, if what I have said about philosophers tending to have flexible concepts is right, we have a head start in this respect. Philosophers are used to thinking in terms of others' frameworks in order to argue against them. If the suggestion that philosophy should be more self-consciously pluralist is correct, this skill could be put to good use finding the best framework for a particular purpose. The same is true at a coarser level of grain for interdisciplinary work. The more differences there are between frameworks, the greater the challenge of the perspective flip, but because fine-grained sharing is not needed for every project, the task is not as formidable as it might seem.

The key here is that there is something we share with everybody if you look at a coarse enough level of grain. I argued in chapter 4 that a concept of rationality emerged and was internalised at the very coarse-grained level of epistemic enquiry in general. I said that the interpretationist strategy only works in virtue of this fact because it is what entitles us to treat anyone we are studying as at least minimally rational, and allows a principle of charity to get off the ground. Based on what I have said here, we can now see that this is true not only of philosophers of science studying scientists, but of philosophers working with one another, and with people from other disciplines. The way into another's framework is via what we already share at a coarser level of grain; this allows us to build to the point where seeing things from the other's perspective is possible.

The suggestion of this last section has been that many of the conclusions we draw about best practice in science can teach us something about our own best practice. Far from being damagingly circular, this implication gives us a way to improve our methods based on our own conclusions. This is what we should expect from a Shaperean point of view; we learn how to learn as we learn (Shapere, 1977b: 185; 1986a: 7).

Bibliography

Abbott, A. (1999), *Department & discipline: Chicago sociology at one hundred*. Chicago: University of Chicago Press.

Adams, F. and Aizawa, K. (2001), “The Bounds of Cognition”, *Philosophical Psychology*, 14, pp. 43–64.

---------- (2008), *The Bounds of Cognition*. MA, Oxford, Victoria: Blackwell.

Allain, P., Joseph, P. A., Isambert, J. L., Le Gall, D., Emile, J. (1998), “Cognitive functions in chronic locked-in syndrome: A report of two cases”, *Cortex*, 34, pp. 629–634.

Allen-Hermanson, S. (2013), “Superdupersizing the mind: extended cognition and the persistence of cognitive bloat”, *Philosophical Studies*, 164 (3), pp. 791–806.

Allport, F. (1924), *Social Psychology*. New York: Houghton Mifflin.

Allport, G. W. (1968), “The Historical Background of Modern Social Psychology”, in G. Lindzey and E. Aronson (eds.), *The Handbook of Social Psychology*, 2nd ed. Reading, MA: Addison-Wesley Publishing Company, pp. 1–80.

Angelucci, M. E. M., Cesário, C., Hiroi, R.H., Rosalen, P.L., Da Cunha, C. (2002), “Effects of caffeine on learning and memory in rats tested in the Morris water maze”, *Brazilian Journal of Medical and Biological Research*, 35, (10), online version: http://dx.doi.org/10.1590/S0100-879X2002001000013, accessed 16/04/15.

Austin, J. (2003), “Transactive memory in organizational groups: The effects of content, consensus, specialization, and accuracy in group performance”, *Journal of Applied Psychology*, 88 (5), pp. 866–878.

Barnier, A. J., Sutton, J. (eds.) (2008), “From individual to collective memory: Theoretical and empirical perspectives” [Special issue], *Memory*, 16 (3), pp. 177–326.

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), pp. 33–51.

Baron, R. A. and Byrne, D. (2000), *Social Psychology*, 9th ed. Needhan, MA: Alleyn & Bacon.

Bartlett, F. C. (1932), *Remembering: A study in experimental and social psychology*. Cambridge: Cambridge University Press.

Bechtel, W. and Hamilton, A. (2007), "Reductionism, integration, and the unity of the Sciences", in T. Kuipers (ed.), *Philosophy of science: focal issues,* (Volume 1 of the handbook of the philosophy of science). New York: Elsevier.

Becker, K. (1998), "On the perfectly general nature of instability in meaning holism", *Journal of Philosophy*, 95, pp. 635–640.

Bennett, M. R. and Hacker, P. M. S. (2003), *Philosophical Foundations of Neuroscience*. Oxford: Blackwell.

Benton, A. (2000), *Exploring the History of Neuropsychology: Selected Papers*. New York: Oxford University Press.

Beurton, P., Falk, H., Rheinerger, H–J. (2000), *The Concept of the Gene in Development and Evolution: Historical and Epistemological Perspectives*. Cambridge: Cambridge University Press.

Bloor, D. (1976), *Knowledge and Social Imagery*. London: Routledge.

Boyd, R. (1991), "Realism, Anti-Foundationalism and the Enthusiasm for Natural Kinds", *Philosophical Studies*, 61, (1–2), pp. 127–148.

---------- (1999a), "Homeostasis, species, and higher taxa", in R. A. Wilson (ed.), *Species*. Cambridge: MIT Press.

---------- (1999b), "Kinds, Complexity and Multiple Realization: Comments on Millikan's 'Historical Kinds and the Special Sciences'", *Philosophical Studies*, 95, pp. 67–98.

Brandom, R. (1994), *Making it Explicit: Reasoning, Representing, & Discursive Commitment*. Cambridge, MA: Harvard University Press.

---------- (2000), *Articulating Reasons: An Introduction to Inferentialism*. Cambridge, MA: Harvard University Press.

---------- (2007), "Inferentialism and Some of Its Challenges", *Philosophy and Phenomenological Research*, 74 (3), pp. 651–676.

Brigandt, I. (2003), "Species Pluralism Does Not Imply Species Eliminitivism", *Philosophy of Science*, 70 (Proceedings), pp. 1305–1316.

---------- (2004a), "Conceptual Role Semantics, the Theory Theory, and Conceptual Change", unpublished manuscript, available at http://www.ualberta.ca/~brigandt/presentations.html, accessed 20/03/14.

---------- (2004b), “Holism, Concept Individuation, and Conceptual Change”, in M. Hernandez Iglesias (ed.), *Proceedings of the 4th Congress of the Spanish Society for Analytic Philosophy*, pp. 30–34. [Pagination from http://www.ualberta.ca/~brigandt/SEFA_04.pdf accessed 31/12/14.]

---------- (2004c), “An Alternative to Kitcher’s Theory of Conceptual Progress and His Account of the Change of the Gene Concept”, unpublished manuscript, available at http://philpapers.org/rec/BRIAAT-2, accessed 31/12/14.

---------- (2010a), “Beyond Reduction and Pluralism: Toward an Epistemology of Explanatory Integration in Biology”, *Erkenntnis*, 73, pp. 295–311.

---------- (2010b), “The epistemic goal of a concept: accounting for the rationality of semantic change and variation” *Synthese*, 177, pp. 19–40.

---------- (2011), “Natural Kinds and Concepts: A Pragmatist and Methodologically Naturalistic Account”, in J. Knowles and H. Rydenfelt (eds.), *Pragmatism, Science and Naturalism*. Berlin: Peter Lang Publishing, pp. 171–196.

---------- (2012), “The dynamics of scientific concepts: The relevance of epistemic aims and values”, *Scientific concepts and investigative practice*, 3, pp. 75–103.

Brigandt, I., and Love, A. C. (2010), “Evolutionary Novelty and the Evo-Devo Synthesis: Field Notes”, *Evolutionary Biology*, 37, pp. 93–99.

Bulmer, M. (1984), *The Chicago School of Sociology: Institutionalization, diversity, and the rise of sociological research*. Chicago: University of Chicago Press.

Carey, J. W. (1996), “The Chicago School and Mass Communication Research”, in E. E. Dennis and E. Wartella (eds.), *American communication research: The remembered history*. Mahwah, N.J.: Erlbaum, pp. 21–38.

Cartwright, N. (1999), *The dappled world: A study of the boundaries of science*. Cambridge: Cambridge University Press.

Casper, S. T. (2010), “Book Review: A Revisionist History of American Neurology”, *Brain*, 133 (2), pp. 638–642, available at http://brain.oxfordjournals.org/content/133/2/638.full, accessed 10/07/13.

Chang, H. (2004), *Inventing Temperature: Measurement and Scientific Progress*. New York: Oxford University Press.

---------- (2012), *Is Water H_2O?: Evidence, Realism and Pluralism*. Dordrecht, London: Springer.

Chemero, A. (2009), *Radical Embodied Cognitive Science*. Cambridge, MA: MIT Press.

Clark, A. and Chalmers, D. (1998), "The Extended Mind", *Analysis*, 58, pp.7–19.

Clark, A. (2003), *Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence*. New York: Oxford University Press.

---------- (2007), "Curing Cognitive Hiccups: A Defence of the Extended Mind", *The Journal of Philosophy*, 104 (4), pp. 163–192.

---------- (2008), *Supersizing The Mind: Embodiment, Action, and Cognitive Extension*. New York: Oxford University Press.

---------- (2010), "*Memento*'s Revenge" in R. Menary (ed.), *The Extended Mind*. Cambridge Massachusetts: The MIT Press, pp. 43–66.

Clark, J. I., Brooksbank, C., Lomax, J. (2005), "It's all GO for Plant Scientists", *Plant Physiology*, 138, pp. 1268–1279.

Collins, H. M. (1981), "What is TRASP? The radical programme as a methodological imperative", *Philosophy of the Social Sciences*, 11, pp. 215–224.

Collins, H. M. and Evans, R. (2002), "The Third Wave of Science Studies: Studies of Expertise and Experience", in E. Selinger and R. P. Crease (eds.) (2006), *The Philosophy of Expertise*. New York: Columbia University Press.

Craver, C. F. (2002), "Interlevel Experiments and Multilevel Mechanisms in the Neuroscience of Memory", *Philosophy of Science*, 69, pp. S83–S97.

---------- (2009), "Mechanisms and natural kinds", *Philosophical Psychology*, 22 (5), pp. 575–594.

Dahlbäck, N., Kristiansson, M., Stjernberg, F. (2013), "Distributed Remembering Through Active Structuring of Activities and Environments", *Review of Philosophy and Psychology*, 4, pp. 153–165.

Dale, R. (2008), "The possibility of a pluralist cognitive science", *Journal of Experimental & Theoretical Artificial Intelligence*, 20 (3), pp. 155–179.

Darden, L. and Maull, N. (1977), "Interfield Theories", *Philosophy of Science*, 44 (1), pp. 43–64.

Davidson, D. (1987), 'Knowing One's Own Mind', *Proceedings and Addresses of the American Philosophical Association*, 61, pp. 441–58.

DeJong, R. N. (1982), *A history of American neurology*. New York: Raven Press.

Delia, J. G. (1987), “Communication research: A history”, in C. Berger and S. Chaffee (eds.), *Handbook of Communication Science*. Newbury Park, CA: Sage, pp. 20–98.

Djerassi, C. and Hoffmann, R. (2001), *Oxygen*. Weinheim: Wiley-VCH.

Draaisma, D. (2000), *Metaphors of Memory: A History of Ideas about the Mind*, Trans. P. Vincent. Cambridge: Cambridge University Press. Originally published in Dutch as *De Metaforenmachine – een geschiedenis van het geheugen*, (1995), Historiche Uitgeverij.

Dudai, Y. (2002), *Memory from A to Z. Keywords, Concepts, and Beyond*. Oxford: Oxford University Press.

Dupré, J. (1993), *The Disorder of Things: Metaphysical Foundations of the Disunity Of Science*. Cambridge, MA: Harvard University Press.

Durkheim, E. (1893), *The Division of Labour in Society*, trans. (1984), W.D. Halls. New York: The Free Press.

---------- (1915), *Elementary forms of the religious life*. New York: Macmillan.

---------- (1965), “A Durkheim Fragment: the Conjugal Family”, *American Journal of Sociology*, 70 (5), pp. 527–536.

Edwards, D. and Middleton, D. (1986), “Joint remembering: constructing an account of shared experience through conversational discourse, *Discourse Processes*, 9, pp. 423–459.

---------- (1987), “Conversation and Remembering: Bartlett Revisited”, *Applied cognitive psychology*, 1, pp. 77–92.

Edwards, D. and Potter, J. (1992), *Discursive Psychology*. London: SAGE Publications.

Faraj, S. and Sproull, L. (2000), “Coordinating expertise in software development teams”, *Management Science*, 46 (12), pp. 1554–1568.

Farr, R. M. (1996), *The roots of modern social psychology: 1872-1954*. Oxford, UK: Blackwell.

Fenton, A. and Alpert, S. (2008), “Extending Our View on Using BCIs for Locked-in Syndrome”, *Neuroethics*, 1, pp. 119–132.

Figdor, C. (2013), “What is the ‘Cognitive’ in Cognitive Neuroscience?”, *Neuroethics*, 6, pp. 105–114.

Fisher, J. (2006), *Pragmatic Conceptual Analysis*. PhD thesis, University of Arizona, USA. Available from ProQuest Dissertations & Theses (PQDT), 2007.

Fiske, J. (1990), *Introduction to Communication Studies*. London/New York: Routledge.

Fodor, J. (1994), “Concepts: a potboiler”, *Concepts*, 50, pp. 95–113.

Fodor, J. (1998), *Concepts: Where Cognitive Science Went Wrong*. Oxford: Oxford University Press.

Fodor, J. and Lepore, E. (2007), “Brandom beleaguered”, *Philosophy and Phenomenological Research*, 74 (3), pp. 677–691.

Fournier, G. (2009), “Schema”, *PsychCentral encyclopaedia*, available at http://psychcentral.com/encyclopedia/2009/schema/, accessed 24/01/14.

Frege, G. (1948), “Sense and Reference”, *The Philosophical Review*, 57 (3), pp. 209–230.

Galison, P. (1997), *Image and Logic*. Chicago: University of Chicago Press.

Gardner-Thorpe, C. (2000), “Book Review: A Short History of Neurology”, *Brain*, 123 (12), pp. 2573–2575, available at http://brain.oxfordjournals.org/content/123/12/2573.full, accessed 10/07/13.

Gene Ontology Consortium (2000), “Gene Ontology: tool for the unification of biology”, *Nature Genetics*, 25 (1), pp. 25–29.

Gene Ontology, (1999), “Documentation”, available at http://geneontology.org/page/documentation, accessed 20/04/2015.

Gibson, J. J. (1963), “The Useful Dimensions of Sensitivity”, in E. Reed, and R. Jones (eds.), (1982), *Reasons for Realism: Selected Essays of James J. Gibson*. New Jersey: Lawrence Erlbaum Associates, pp. 350–374.

---------- (1966), “The Problem of Temporal Order in Stimulation and Perception”, in E. Reed, and R. Jones (eds.), (1982), *Reasons for Realism: Selected Essays of James J. Gibson*. New Jersey: Lawrence Erlbaum Associates, pp. 171–179.

---------- (1979), *The Ecological Approach to Visual Perception*. Boston: Houghton-Miffin.

Gigone, D. and Hastie, R. (1996), “The impact of information on group judgment: A model and computer simulation”, in E. H. Witte and J. H. Davis (eds.), *Understanding group behaviour*. Mahwah, NJ: Erlbaum, pp. 221–251.

Goertzen, J. R. (2010), “Dialectical pluralism: A theoretical conceptualization of pluralism in psychology”, *New Ideas in Psychology*, 28, pp. 201–209.

Goertzen, J. R. and Smythe, W. E. (2010), “Theorizing pluralism: An introduction”, *New Ideas in Psychology*, 28, pp. 199–200.

Goldstein, B. (1981), "The Ecology of J. J. Gibson's Perception", *Leonardo*, 14 (3), pp. 191–195.

Goldstone, R. L. and Rogosky, B. J. (2002), "Using relations within conceptual systems to translate across conceptual systems", *Cognition*, 84, pp. 295–320.

Goodman, N. (1955), *Fact, Fiction, and Forecast*. Cambridge, MA: Harvard University Press.

Gray, W. D. and Fu, W.-T. (2004), "Soft constraints in interactive behavior: The case of ignoring perfect knowledge in the world for imperfect knowledge in the head", *Cognitive Science*, 28 (3), pp. 359–382.

Gray, W. D. and Veksler, V. D. (2005), "The acquisition and asymmetric transfer of interactive routines", in B. G. Bara, L. Barsalou and M. Bucciarelli (eds.), *27th annual meeting of the Cognitive Science Society*. Austin, Texas: Cognitive Science Society.

Green, C. D. (1992), "Is unified positivism the answer to psychology's disunity?", *American Psychologist*, 47 (8), pp. 1057–1058.

Griffiths, P. E. (2002), "Lost: One Gene Concept. Reward to Finder", *Biology and Philosophy*, 17, pp. 271–283.

---------- (2004), "Emotions as Natural and Normative Kinds", *Philosophy of Science*, 71, pp. 901–911.

Griffiths, P. E. and Stotz, K. (2006), "Genes in the Postgenomic Era", *Theoretical Medicine and Bioethics*, 27 (6), pp. 499–521.

---------- (2008), "Experimental Philosophy of Science", *Philosophy Compass*, 3, pp. 507–521.

Gruber, T. (1993), "A translation approach to portable ontology specifications", *Knowedge Acquisition*, 5, pp. 199–220.

Hacking, I. (1995a), "The Looping Effects of Human Kinds", in D. Sperber, D. Premack, A.J. Premack (eds.), *Causal Cognition: a multi-disciplinary debate*. New York: Oxford University Press, ch. 12, pp. 351–383.

---------- (1995b), *Rewriting the Soul: Multiple Personality and the Sciences of Memory*. Princeton: Princeton University Press.

---------- (1999), *The Social Construction of What*. Cambridge MA: Harvard University Press.

Halbwachs, M. (1992), *On Collective Memory*, trans. and ed. L. A. Coser. Chicago: University of Chicago Press.

Hanson, N. R. (1958), *Patterns of Discovery*. Cambridge: Cambridge University Press.

Hardin, C. L. and Rosenberg, A. (1982), ‘In Defence of Convergent Realism’, *Philosophy of Science*, 49, pp. 604–615.

Harré, R. and Gillett, G., (1994), *The Discursive Mind*. California: SAGE Publications.

Haslam, N. (2002), “Kinds of Kinds: A Conceptual Taxonomy of Psychiatric Categories”, *Philosophy, Psychiatry, & Psychology*, 9 (3), pp. 203–217.

Haslanger, S. (1999), “What Knowledge Is and What It Ought to Be: Feminist Values and Normative Epistemology”, *Noûs*, 33, Supplement: Philosophical Perspectives, 13, Epistemology, pp. 459–480.

Heaton-Armstrong, A., Shepherd, E., Wolchover, D. (eds.) (1999), *Analysing Witness Testimony*. London: Blackstone Press.

Heersmink, R. (2013), “Embodied Tools, Cognitive Tools and Brain-Computer Interfaces”, *Neuroethics*, 6 (1), pp. 207–219.

Hendry, R. F. (2008), “Chemistry”, in S. Psillos and M. Curd (eds.), *The Routledge companion to the philosophy of science*. London: Routledge, pp. 520–530.

Henriques, G. (2004), “Psychology defined”, *Journal of Clinical Psychology*, 60, pp. 1207–1221.

Hirst, W. and Manier, D. (2008), “Towards a psychology of collective memory”, *Memory*, 16 (3), pp. 183–200.

Hogg, M. A. and Tindale, R. S. (2001), *Blackwell Handbook of Social Psychology: Group Processes*. Massachusetts/Oxford: Blackwell.

Hollingshead, A. B. and Brandon, D. P. (2003), “Potential Benefits of Communication in Transactive Memory Systems”, *Human Communication Research*, 29 (4), pp. 607–615.

Hooke, R. (1682), *An Hypothetical Explication of Memory: how the Organs made use if by the mind in its Operation may be Mechanically understood*, in R. Waller (1705), *The Posthumous Works of Robert Hooke*, London.

Hurley, S. (2010), “Varieties of externalism” in R. Menary (ed.), *The Extended Mind*. Cambridge Massachusetts: The MIT Press, pp. 101–154.

Hutchins, E. (1995), *Cognition in the Wild*. Cambridge, MA: MIT Press.

---------- (2000), “Distributed Cognition”, http://www.artmap-research.com/wp-content/uploads/2009/11/Hutchins_DistributedCognition.pdf, accessed 27/08/12.

---------- (2011), “Enculturating the Supersized Mind”, *Philosophical Studies*, 152 (3), pp. 437–446.

Hyman, I. (2012), “Remembering the Father of Cognitive Psychology, Ulric Neisser (1928-2012)”, *APS Observer*, available at http://www.psychologicalscience.org/index.php/publications/observer/2012/may-june-12/remembering-the-father-of-cognitive-psychology.html, accessed 24/01/14.

Ilardi, S. S. and Feldman, D. (2001), “The cognitive neuroscience paradigm: A unifying metatheoretical framework for the science and practice of clinical psychology”, *Journal of Clinical Psychology*, 57, pp. 1067–1088.

Jackson, F. (1998), *From Metaphysics to Ethics: A defence of conceptual analysis*. Oxford: Clarendon Press.

Jung, K. (1916), “The Structure of the Unconscious” in G. Adler and R. F. C. Hull (1953, 1956) (eds. and trans.), *Two Essays on Analytical Psychology, The Collected Works of C. G. Jung, Vol. 7*. Princeton: Princeton University Press.

---------- (1922), *Collected papers on analytical psychology* (2nd ed.). London: Bailliere, Tindall and Cox.

Kaplan, D. M. (2012), “How to demarcate the boundaries of cognition”, *Biology and Philosophy*, 27 (4), pp. 545–570.

Keil, F. C. (1989), “Spiders in the Web of Belief: The Tangled Relations Between Concepts and Theories”, *Mind and Language*, 4, pp. 43–50.

---------- (2005), “Knowledge, Categorization and the Bliss of Ignorance”, in L. Gershkoff-Stowe, D. H. Rakison (eds.), *Building Object Categories in Developmental Time*. London: Routledge.

Kellert, S. H., Longino, H. E., Waters, C. K. (eds.) (2006), *Scientific Pluralism, Minnesota Studies in the Philosophy of Science, Volume XIX*. Minneapolis: University of Minnesota Press.

Khalidi, M. A. (1995), “Two Concepts of Concept”, *Mind and Language*, 10 (4), pp. 402–422.

Kirsh, D. (1995), “The intelligent use of space”, *Artificial Intelligence*, 73, pp. 31–68.

Kitcher, P. (2003), *Science, Truth and Democracy*. New York: Oxford University Press.

Knobe, J. and Samuels, R. (2013), "Thinking like a scientist: Innateness as a case study", *Cognition*, 126, pp. 72–86.

Knowles, E. S. (1982), "From individuals to group members: A dialectic for the social sciences", in W. J. Ickes and E. S. Knowles (eds.), *Personality, roles, and social behaviour*. New York: Springer-Verlag, pp. 1–32.

Koch, S. (1993), "'Psychology' or 'the psychological studies?'", *American Psychologist*, 48, pp. 902–904.

Kripke, S. (1972/1981), *Naming and Necessity*. Oxford: Blackwell.

Kuhn, T. S. (1970, 2nd ed.; 1996, 3rd ed.), *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

---------- (1977), *The Essential Tension: Selected Studies in Scientific Tradition and Change*. Chicago: University of Chicago Press.

Kukla, A. (1992), "Unification as a goal for psychology", *American Psychologist*, 47 (8), pp. 1054–1055.

Kunkel, J. H. (1992), "The units of unification: Theories or propositions?", *American Psychologist*, 47 (8), pp. 1058–1059.

Kusch, M. (1997), "The Sociophilosophy of Folk Psychology", *Studies in the History and Philosophy of Science*, 28 (1), pp. 1–25.

---------- (1999), *Psychological Knowledge: A Social History and Philosophy*. London: Routledge.

Kyselo, M. (2013), "Locked-in Syndrome and BCI - Towards an Enactive Approach to the Self", *Neuroethics*, 6, pp. 579–591.

Lakatos, I. (1968), "Criticism and the methodology of scientific research programmes", *Proceedings of the Aristotelian Society*, 69, pp. 149–186.

Lamiell, J. T., Laux, L., Goertzen, J. R., Smythe, W. E. (eds.) (2010), "Personalistic Thinking Theorizing Pluralism" [Special issue], *New Ideas in Psychology*, 28, pp. 105–262.

Lewis, K. (2003), "Measuring transactive memory systems in the field: Scale development and validation", *Journal of Applied Psychology*, 88 (4), pp. 587–604.

Lane, P. J., Koka, B. R., Pathak, S. (2006), “The reification of absorptive capacity: A critical review and rejuvenation of the construct”, *Academy of Management Review*, 31 (4), pp. 833–863.

Laudan, L. (1977), *Progress and Its Problems*. California: University of California Press.

---------- (1984), *Science and Values: The Aims of Science and Their Role in Scientific Debate*. Berkeley and Los Angeles, California: University of California Press.

Le Bon, G. (1896/1960), *The Crowd* (translation of *Psychologie des Foules*). New York: The Viking Press.

León-Carrión, J., van Eeckhoot, P., del Rosario Domínguez-Morales, M. (2002, review), “The locked-in syndrome: a syndrome looking for a therapy”, Review of subject, *Brain Injury*, 16 (7), pp. 555–569.

---------- (2002, survey), “The locked-in syndrome: a syndrome looking for a therapy”, Survey, *Brain Injury*, 16 (7), pp. 571–582.

Levy, N. (2007), *Neuroethics: challenges for the 21st Century*. New York: Cambridge University Press.

Liang, D. W., Moreland, R., Argote, L. (1995), “Group versus individual training and group performance: The mediating role of transactive memory”, *Personality and Social Psychology Bulletin*, 21 (4), pp. 384–393.

Lindzey, G. and Aronson, E. (eds.) (1954/1968), *Handbook of social psychology*. Reading, MA: Addison-Wesley.

Lipton, P. (1991), *Inference to the Best Explanation*. London and New York: Routledge.

Loftus, E. and Palmer, J. (1974), “Reconstruction of Automobile Destruction: An example of the interaction between language and memory”, *Journal of Verbal Learning and Verbal Behaviour*, 13, pp. 585–589.

Loftus, E. F. and Pickrell, J. E. (1995), “The formation of false memories”, *Psychiatric Annals*, 25 (12), pp. 720–725.

Lorge, I. and Solomon, H. (1955), “Two models of group behavior in the solution of eureka-type problems”, *Psychometrika*, 20, pp. 139–148.

Love, A. C. (2008), “Explaining Evolutionary Innovations and Novelties: Criteria of Explanatory Adequacy and Epistemological Prerequisites”, *Philosophy of Science*, 75, pp. 874–886.

Machery, E. (2009), *Doing without Concepts*. New York: Oxford University Press.
Machery, E. and Cohen, K. (2012) "An Evidence-Based Study of the Evolutionary Behavioral Sciences", *British Journal for the Philosophy of Science,* 63, pp. 177–226.
Magoun, H. W. (1975), "The role of research institutes in the advancement of neuroscience: Ranson's Institute of Neurology, 1928–1942", in F. G. Worden, J. P. Swazey, G. Adelman, (eds.), *The neurosciences: paths of discovery*. Cambridge, MA, and London: The MIT Press, pp. 515–527.
Mandler, J. M. and Johnson, N. S. (1977), "Rememberance of things parsed: Story structure and recall", *Cognitive Psychology,* 9, pp. 111–151.
Martin, G. N., Carlson, N. R., Buskist, W. (2009), *Psychology*. Harlow: Pearson Education Limited.
Matarazzo, J. D. (1987), "There is only one psychology, no specialties, but many applications", *American Psychologist*, 42, pp. 893–903.
McDougall, W. (1920), *The group mind: A sketch of the principles of collective psychology with some attempt to apply them to the interpretation of national life and character*. Cambridge, UK: Cambridge University Press.
McHenry, L. C. (1969), *History of neurology*. Springfield, IL: Charles C. Thomas.
McNally, R. J. (1992), "Disunity in psychology: Chaos or speciation?", *American Psychologist*, 47 (8), p. 1054.
Michaelian, K. (2010), "Is memory a natural kind?", *Memory Studies*, 4 (2), pp. 170–189.
Mead, G. H. (1934), *Mind, self and society from the standpoint of a social behaviourist*. Chicago, IL: Chicago University Press.
Meehl, P. E. (1978), "Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology", *Journal of Consulting and Clinical Psychology*, 46, pp. 806–834.
Michaelian, K. and Sutton, J. (2013), "Distributed Cognition and Memory Research: History and Current Directions", *Review of Philosophy and Psychology,* 4, pp.1–24.
Michinov, E., Olivier-Chiron, E., Rusch, E., Chiron, B. (2008), "Influence of transactive memory on perceived performance, job satisfaction and identification in anesthesia teams", *British Journal of Anaesthesia*, 100 (3), pp. 327–332.

Miller, R. B. (1992), "Introduction to the philosophy of clinical psychology", in R. B. Miller (ed.), *The restoration of dialogue: Readings in the philosophy of clinical psychology*. Washington, DC: American Psychological Association, pp. 1–28.

Minsky, M. (1975), "A framework for representing knowledge", in P. H. Winston (ed.), *The psychology of computer vision*. New York, NY: McGraw-Hill, pp. 211–277.

Mitchell, S. D. (2002), "Integrative Pluralism", *Biology and Philosophy*, 17, pp. 55–70.

---------- (2003), *Biological Complexity and Integrative Pluralism*. Cambridge: Cambridge University Press.

Mohammed, S. and Dumville, B. C. (2001), "Team mental models in a team knowledge framework: Expanding theory and measurement across disciplinary boundaries", *Journal of Organizational Behavior*, 22, pp. 89–106.

Moreland, R. L., Hogg, M. A., Haines, S. C. (1994), "Back to the future: Social psychological research on groups", *Journal of Experimental Social Psychology*, 30 (6), pp. 527–555.

Morris, R. (1984), "Developments of a water-maze procedure for studying spatial learning in the rat", *Journal of Neuroscience Methods*, 11 (1), pp. 47–60.

Mulkay, M. (1979), *Science and the Sociology of Knowledge*. London: Allen and Unwin.

Murphey, G. L. and Medin, D. L. (1985), "The Role of Theories in Conceptual Coherence", *Psychological Review*, 92, pp. 289–316.

National Human Genome Research Institute (2013), Online Education Kit, http://www.genome.gov/25520244, accessed 20/04/15.

Neisser, U. (1967), *Cognitive Psychology*. New York: Appleton-Century-Crofts.

---------- (1976), *Cognition and Reality*. San Francisco: Freeman.

---------- (1981), "John Dean's Memory: A Case Study", *Cognition*, 9 (1), pp. 1–22. Reprinted as chapter 11 in Neisser (1982).

---------- (1982), *Memory Observed: Remembering in Natural Contexts*. San Francisco: W. H. Freeman and Company.

---------- (1997), "The ecological study of memory", *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 352 (1362), pp. 1697–1701.

Neisser, U. and Winograd, E. (1988), *Remembering Reconsidered*. Cambridge: Cambridge University Press.

New, P. W. and Thomas, S. J. (2005), "Cognitive impairments in the locked-in syndrome: a case report", *Archives of Physical Medicine and Rehabilitation*, 86, pp. 338–343.

Noë, A. (2004), *Action in Perception*. Cambridge, MA: MIT Press.

Norris, S. L., Lau, J., Smith, S. J., Schmid, C. H., Engelgau, M. M. (2002), "Self-management education for adults with type 2 diabetes: a meta-analysis of the effect on glycemic control", *Diabetes Care*, 25, pp. 1159–1171.

Oldfield, R. C. (1954), "Memory mechanisms and the theory of schemata", *British Journal of Psychology*, 45 (1), pp. 14–23.

Onofri, M., Thomas, A., Paci,C. et al. (1997), "Event related potentials recorded in patients with locked in syndrome", *Journal of Neurology, Neurosurgery and Psychiatry*, 63, pp. 759–764.

Parker, S. (1666), *A Free and Impartial Censure of the Platonick Philosophie*. Oxford.

Passmore, J. (1967), "Logical Positivism", in P. Edwards (ed.), *The Encyclopedia of Philosophy (Volume 5)*. New York: Macmillan, pp. 52–57.

Pavitt, C. (2003a), "Why We Have to Be Reductionists About Group Memory" Colloquy: Do Interacting Groups Perform Better Than Aggregates of Individuals? *Human Communication Research*, 29 (4), pp. 592–599.

---------- (2003b), "Why We Still Have to Be Reductionists About Group Memory", *Human Communication Research*, 29 (4), pp. 624–629.

Pearson, H. (2006), "What is a gene?", *Nature*, 441, pp. 399–401.

Peters, J. D. (1986a), "Institutional sources of intellectual poverty in communication research", *Communication Research*, 13, pp. 527–559.

Poldrack, (n.d.), http://www.cognitiveatlas.org/, accessed 20/04/15.

Poldrack, R. A., Kittur, A., Kalar, D., Miller, E., Seppa, C., Gil, Y., Stott Parker, D., Sabb, F. W., Bilde, R. M. (2011), "The cognitive atlas: toward a knowledge foundation for cognitive neuroscience", *Frontiers in Neuroinformatics*, 5, article 17, pp. 1–11.

Propp, K. M. (1999), "Collective information processing in groups", in L. Frey, D. Gouran amd M. S. Poole (eds.), *Handbook of group communication theory and research*. Thousand Oaks, CA: Sage, pp. 225–250.

---------- (2003), "In Search of the Assembly Bonus Effect: Continued Exploration of Communication's Role in Group Memory", *Human Communication Research*, 29 (4), pp. 600–606.

Putnam, H. (1982), "Three Kinds of Scientific Realism", *Philosophical Quarterly*, 32, pp. 195–200.

Quine, W. V. O. (1976), *Two dogmas of empiricism*. Springer Netherlands.

Rand, K. L. and Ilardi, S. S. (2005), "Toward a Consilient Science of Psychology", *Journal of Clinical Psychology*, 61 (1), pp. 7–20.

Reed, E. and Jones, R. (eds.) (1982), *Reasons for Realism: Selected Essays of James J. Gibson*. New Jersey: Lawrence Erlbaum Associates.

Ren, Y. and Argote, L. (2011), "Transactive Memory Systems 1985–2010: An Integrative Framework of Key Dimensions, Antecedents, and Consequences", *The Academy of Management Annals*, 5 (1), pp. 189–229.

Rey, A. (1958), "Mémorisation d'une Série de 15 Mots en 5 Répétitions", *L'Examen Clinique en Psychologie*, Paris: PUF.

Roediger III, H. L., Dudai, Y., Fitzpatrick, S. M. (eds.) (2007), *Science of Memory: Concepts*. USA: Oxford University Press.

Rogers, E. N. (1994), *A history of communication study*. New York: Free Press.

Roll-Hansen, N. (1985), "A New Perspective on Lysenko?", *Annals of Science*, 42, pp. 261–278.

Rosch, E. (1978), "Principles of Categorization" in E. Rosch and B. Lloyd (eds.) *Cognition and Categorization*. Oxford: Lawrence Erlbaum.

Rosch, E. and Mervis, C. B. (1975), "Family Resemblances: Studies in the Internal Structure of Categories", *Cognitive Psychology*, **7**, pp. 573–605.

Ross, D. and Ladyman, J. (2010), "The Alleged Coupling-Constitution Fallacy and the Mature Sciences", in R. Menary (ed.), *The Extended Mind*. Cambridge Massachusetts: The MIT Press, pp. 155–166.

Rupert, R. (2009), "Keeping HEC in CHEC: One the Priority of Cognitive Systems", manuscript, available at http://spot.colorado.edu/~rupertr/Papers.htm, accessed 24/01/14.

Saito, A. (2000), *Bartlett, culture, and cognition*. Psychology Press.

Salomon, G. (1993), *Distributed cognitions: Psychological and educational considerations*. New York: Cambridge University Press.

Sampson, E. E. (1988), “The deconstruction of self”, in J. Shotter and K. Gergen (eds.), *Texts of Identity*. London: Sage.

Schacter, D. L. and Tulving, E. (1994), “What Are the Memory Systems of 1994?”, in D. L. Schacter and E. Tulving (eds.), *Memory Systems*. Cambridge, MA: MIT Press, pp. 1–38.

Schnakers, C., Majerus, S., Goldman, S., Boly, M., Van Eeckhout, P., Gay, S., Pellas, F., Bartsch, V., Peigneux, P., Moonen, G., Laureys, S. (2008), “Cognitive function in the locked-in syndrome”, *Journal of Neurology*, 255, pp. 323–330.

Schneider, S. M. (1992), “Can this marriage be saved?”, *American Psychologist*, 47 (8), pp. 1055–1057.

Shank, R., and Abelson, R. (1977), *Scripts, plans, goals and understanding*. Hillsdale, NJ: Erlbaum.

Shannon, C. and Weaver, W. (1949), *The Mathematical Theory of Communication*. Illinois: University of Illinois Press.

Shapere, D. (1977a), “Scientific Theories and Their Domains”, in F. Suppe (ed.), *The Structure of Scientific Theories*, 2nd ed.. Urbana, Chicago, London: University of Illinois Press, pp. 518–565.

---------- (1977b), “What can the theory of knowledge learn from the history of knowledge?” in Shapere (ed.) (1984), *Reason and The Search for Knowledge*. Dordrecht: D. Reidel Publishing Company, pp. 182–202.

---------- (1984a), “Objectivity, Rationality and Scientific Change”, *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association , vol. 2: Symposia and Invited Papers*, pp. 637–663.

---------- (1984b), “Alternation of Goals and Language in the Development of Science” in D. Shapere (ed.), *Reason and The Search for Knowledge*, Dordrecht: D. Reidel Publishing Company, pp. 325–341.

---------- (1986a), “External and Internal Factors in the Development of Science”, *Science and Technology Studies*, 4 (1), pp. 1–9.

---------- (1986b), “Replies to Carroll and Turner”, *Science & Technology Studies*, 4 (1), pp. 19–23.

---------- (1987), “Method in the Philosophy of Science and Epistemology: How to Inquire about Inquiry and Knowledge”, in N. J. Nersessian (ed.), *The Process* *of Science: Contemporary Philosophical Approaches to Understanding Scientific Practice*. Dordrecht: Martinus Nijhoff Publishers, pp. 1–39.

Shapin, S. (1992), “Discipline and Bounding: The History and Sociology of Science as Seen Through the Externalism-Internalism Debate”, *History of Science*, 30 (90), pp. 333–369.

Shaw, M. E. (1932), “A Comparison of Individuals and Small Groups in the Rational Solution of Complex Problems”, *The American Journal of Psychology*, 44 (3), pp. 491–504.

Signoret, J. L. (1991), *Batterie d’Efficience Mnésique (BEM 144)*. Paris: Editions Scientifiques Elsevier.

Simpson, C. (1994), *Science of coercion*. New York: Oxford University Press.

Smith, E. and Delargy, M. (2005), “Locked-in Syndrome”, Clinical Review, *British Medical Journal*, 330, pp. 406–409.

Smith, J. A., Harré, R. and Langenhove, L. V. (eds.) (1995), *Rethinking Psychology*. London: SAGE Publications.

Smythe, W. E. and McKenzie, S. A. (2010), “A vision of dialogical pluralism in psychology”, *New Ideas in Psychology*, 28, pp. 227–234.

Sparrow, B., Liu, J., Wegner, D. M. (2011), “Google Effects on Memory: Cognitive Consequences of Having Information at Our Fingertips”, *Science*, 333 (6043), pp. 776–778.

Staats, A. W. (1983), *Psychology’s crisis of disunity: Philosophy and methods for a unified* science. New York: Praeger.

---------- (1991), “Unified Positivism and Unification Psychology: Fad or New Field”, *American Psychologist*, 46 (9), pp. 899–912.

---------- (1999), “Unifying psychology requires new infrastructure, theory, method, and a research agenda”, *Review of General Psychology*, 3, pp. 3–13.

---------- (2004), “The Disunity-Unity Dimension”, *American Psychologist*, May–June, p. 273.

Stasser, G., Stewart, D. D., Wittenbaum, G. M. (1995), “Expert roles and information exchange during discussion: The importance of knowing who knows what”, *Journal of Experimental Social Psychology,* 31, pp. 244–265.

Steiner, I. (1972), *Group processes and productivity*. New York: Academic Press.

---------- (1974), “Whatever happened to the group in social psychology?” *Journal of Experimental Social Psychology*, 10, pp. 1467–1478.

---------- (1986), "Paradigms and groups", *Advances in Experimental Social Psychology*, 19, pp. 539–548.

Sternberg, R. J. and Grigorenko, E. L. (2001), "Unified Psychology", *American Psychologist*, 56 (12), pp. 1069–1079.

Stotz, K., Griffiths, P. E., Knight, R. (2004), "How biologists conceptualize genes: an empirical study", *Studies in History and Philosophy of Biological and Biomedical Sciences*, 35, pp. 647–673.

Suppes, P. (1978), "The Plurality of Science", *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association, Vol. Two: Symposia and Invited Papers*, pp. 3–16.

Sutton, J. (1998), *Philosophy and memory traces*. Cambridge: Cambridge University Press.

---------- (2004), "Representation, reduction and interdisciplinarity in the sciences of memory", in H. Clapin, P. Staines and P. Slezak (eds.), *Representation in mind: New approaches to mental representation*. Oxford: Elsevier.

---------- (2007), "Integrating the Philosophy and Psychology of Memory: Two Case Studies", in M. Marraffa, M. De Caro and F. Ferretti (eds.), *Cartographies of the Mind: Philosophy and Psychology in Intersection, Volume 4 of Studies in brain and mind*. Dordrecht: Springer, pp. 82–92.

Sutton, J., Harris, C. B., Keil, P. G., Barnier, A. J. (2010), "The psychology of memory, extended cognition, and socially distributed remembering", *Phenomenology and the Cognitive Sciences*, 9 (4), pp. 521–560.

Tanney, J. (2013), *Rules, Reason and Self-Knowledge*. Harvard: Harvard University Press.

Theiner, G. (2013), "Transactive Memory Systems: A Mechanistic Analysis of Emergent Group Memory", *Review of Philosophy and Psychology*, 4, pp. 65–89.

Toates, F. M. (2007), *Biological Psychology: An Integrative Approach, 2nd ed.*. Essex: Pearson Education Limited.

Tollefsen, D. P., Dale, R., Paxton, A. (2013), "Alignment, Transactive Memory, and Collective Cognitive Systems" *Review of Philosophy and Psychology*, 4, pp. 49–64.

Toulmin, S. E. (1972), *Human understanding (Vol. 1)*. Princeton, NJ: Princeton University Press.

Tulving, E. (2000), “Concepts of Memory” in E. Tulving and F. I. M Craik (eds.), *The Oxford Handbook of Memory*. New York: Oxford University Press, pp. 33–44.

Uttal, W. R. (2003), *The New Phrenology: The Limits of Localizing Cognitive Processes in the Brain*. Cambridge MA: MIT Press.

Varela, F. J., Thompson, E., Rosch, E. (1991), *The embodied mind: Cognitive science and human experience*. Cambridge MA: MIT Press.

Von Eckhardt, B. (1993), *What is Cognitive Science?* Cambridge MA: MIT Press.

Wagner, E. H. (2002), “Care for chronic diseases: The efficacy of coordinated and patient centred care is established, but now is the time to test its effectiveness”, *British Medical Journal*, 325 (7370), pp. 913–914.

Wagoner, B. (2013), “Bartlett's concept of schema in reconstruction”, *Theory & Psychology*, 23 (5), pp. 553–575.

Wahl-Jorgensen, K. (2000), “Rebellion and ritual in disciplinary histories of U.S. mass communication study: Looking for ‘the reflexive turn’”, *Mass Communication & Society*, 3 (1), pp. 87–115.

---------- (2004), “How Not to Found a Field: New Evidence on the Origins of Mass Communication Research”, *Journal of Communication*, 54 (3), pp. 547–564.

Walter, S. (2010), “Locked-in Syndrome, BCI, and a Confusion about Embodied, Embedded, Extended, and Enacted Cognition”, *Neuroethics*, 3, pp. 61–72.

Watanabe, T. (2010), “Metascientific foundations for pluralism in psychology”, *New Ideas in Psychology*, 28, pp. 253–262.

Weaver, W. (1949), “The mathematics of communication”, *Scientific American*, 181, pp. 11–15.

Wechsler, D. (1991), *Echelle Clinique de Mémoire de Wechsler (forme révisée)*. Paris: Adaptation et traduction du Centre de Psychologie Appliquée.

Wegner, D. M. (1986), “Transactive Memory: A Contemporary Analysis of the Group Mind” in B. Mullen and G. R. Goethals (eds.), *Theories of group behaviour*. New York: Springer-Verlag, pp. 185–208.

Wegner, D. M., Erber, R., Raymond, P. (1991), “Transactive memory in close relationships”, *Journal of Personality and Social Psychology*, 61, pp. 923–929.

Wegner, D. M., and Guiliano, T. (1982), “The forms of social awareness”, in W. J. Ickes and E. S. Knowles (eds.), *Personality, roles and social behaviour*. New York: Springer-Verlag, pp. 165–198.

Wegner, D. M., Guilano, T., Hertel, P. T. (1985), “Cognitive Interdependence in Close Relationships”, in W. Ickes (ed.), *Compatible and Incompatible Relationships*. New York: Springer-Verlag, pp. 253–276.

Weisberg, M., Okasha, S., Mäki, U. (eds.) (2011), “Modelling in Biology and Economics”, *Biology and Philosophy special issue*, 26 (5), pp. 613–315.

Welze, H. and Markowitsch, H. (2005), “Towards a bio-psycho-social model of autobiographical memory”, *Memory*, 13 (1), pp. 63–78.

Wertsch, J. V. and Roediger III, H. L. (2008), “Collective memory: Conceptual foundations and theoretical approaches”, *Memory*, 16 (3), pp. 318–326.

Wessel, I. and Moulds, M. L. (2008), “Collective memory: A perspective from (experimental) clinical psychology”, *Memory*, 16 (3), pp. 288–304.

Wheeler, M. (2005), *Reconstructing the Cognitive World: The next step*. Cambridge MA: MIT Press.

Williamson, T. (2007), *The Philosophy of Philosophy*. Oxford, Blackwell.

Wilson, R. A. (2005a), “Collective memory, group minds, and the extended mind thesis”, *Cognitive Processing*, 6 (4), pp. 227–236.

---------- (2005b), *Genes and the agents of life*. Cambridge: Cambridge University Press.

---------- (2010), “Meaning Making and the Mind of the Externalist”, in R. Menary (ed.), *The Extended Mind*, Cambridge Massachusetts: The MIT Press, pp. 167–188.

Wimsatt, W. C. (1994), “The Ontology of Complex Systems: Levels of Organisation, Perspectives and Causal Thickets”, in M. Matthen and R. Ware (eds.), *Canadian Journal of Philosophy*, supplement, 20. University of Calgary Press, pp. 207–274.

Wittenbaum, G. M. (2003), “Putting Communication into the Study of Group Memory”, *Human Communication Research*, 29 (4), pp. 616–623.

Wittgenstein, L. (1921), *Tractatus Logico-Philosophicus*. London: Routledge.

---------- (1953), *Philosophical Investigations*. Oxford: Blackwell.

Wolpaw, J. R. (2007), “Brain-computer interfaces as new brain output pathways”, *The Journal of Physiology*, 5793, pp. 613–619.

Wundt, W. (1916), *Elements of folk psychology* (Trans.). New York: Macmillan. (Original work published 1910).

Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., Wager. T. D. (2011), "Large-scale automated synthesis of human functional neuroimaging data", *Nature Methods,* 8 (8), pp. 665–70.

Yoo, C., Ayello, E. A., Robins, B., Salamanca, V., R., Bloom, M. J., Linton, P., Brem, H., O'Neill, D. K. (2012), "Perioperative use of bispectral (BIS) monitor for a pressure ulcer patient with locked-in syndrome (LIS)", *International Wound Journal*, doi: 10.1111/iwj.12001.

Zagorski, N. (2005), "Profile of Elizabeth Loftus", *Proceedings of the National Academy of Sciences of the United States of America*, 102 (39), pp. 13721–13723.