Logique & Analyse **228** (2014), 683–699

HOW SHOULD WE STUDY CONCEPTS IN THE COGNITIVE SCIENCES?

THE EXAMPLE OF MEMORY

RUTH HIBBERT

ABSTRACT

There is considerable variation in the concept of memory employed in different branches of the cognitive and social sciences. This paper is about how a philosopher of science can make sense of this divergence. First I consider the reasons for focussing on *concepts* specifically. Then I pose a question about the *classifying practices of scientists*, and consider various methods for investigating the answer. I defend a historically situated case study method as the best option, and suggest some appropriate case studies for the example of the concept of memory.

Introduction

Memory is a subject of importance right across the cognitive and social sciences, from experimental psychology and neuroscience to sociology, anthropology, philosophy and A.I. But these disciplines and their approaches to the subject are so diverse that it barely seems they are talking about the same thing. They have different *concepts* of memory. This paper is about how a philosopher of science can make sense of this divergence.

First I will defend the focus on scientists' *concepts* specifically, and pose a question about the *classifying practices of scientists*. I will go on to discuss two possible methods for trying to answer such questions: Experimental philosophy and a case study method.

I will consider how appropriate each method is for studying MEMORY[1] specifically, and argue that the case study method is superior here. I will then go on to outline how one should work with the case study method, providing a guide to choosing appropriate case studies, and advocating an historical version of the case study method. This historical approach allows further normative claims to be made than the case study method alone.

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

doi: 10.2143/LEA.228.0.3078179.

1. What is the Question and why Study Concepts to Find the Answer?

What question exactly can the study of concepts help us to answer? The kind of question I want to propose a method for addressing here is: How do users of a scientific concept classify phenomena according to that concept, i.e. how do they employ the concept?[2] This is a pragmatic question about the *role* that the concept plays in discourse and practice (in scientific theory and experiment). Why ask this question about concepts?

I take the project of science to be about conceptualising the world; carving it up into entities, states, processes etc. in order to explain, predict and manipulate it. It is scientists' concepts that allow them to do this; concepts are the tools of science. This idea is far from new (e.g. work descended from the later Wittgenstein's metaphor of language as tool use, Wittgenstein, 1953).

A similar pragmatic approach to concepts is taken by Sally Haslanger with respect to KNOWLEDGE, and by Justin Fisher mainly concerning philosophical concepts (Haslanger, 1999; Fisher, 2006). 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, in line with the tool-use metaphor.

Griffiths and Stotz, whose work I will discuss later in this paper, apply the idea of concepts as tools specifically to scientific concepts. They say

> For the scientific practitioner, concepts are tools which classify experience in ways that meet their specific needs and which are reshaped in the light of new empirical findings. This attitude is sometimes made explicit, but is implicit whenever scientists describe a statement as a 'definition' and yet regard it as hostage to future empirical findings, as they commonly do. (Griffiths and Stotz, 2008: 508).

Again the focus is on the role that a concept plays — what the tool is used for. As this quote highlights, scientific concepts change over time. This is what we should expect; as discoveries are made, how science views the world changes in response, and the best tools for describing and manipulating it change. In addition to this change over time, scientific concepts[3] often display heterogeneity across the scientific community employing the

[2] One could also ask about what physically constitutes concepts (whether they supervene on individual brain states, the behaviour of social groups etc.), or about how concepts are encoded at the psychological level (are they prototypes, exemplars etc.). Neither of these are the kind of study of concepts I am interested in here.

[3] It may be that all concepts are like this, but scientific concepts are my particular concern here.

concept. The variation is particularly marked between members of different subdisciplines. For example, Stotz, Griffiths and Knight (2004, hereafter SGK), using the experimental method for investigating concepts discussed below, identify differences in the gene concept used between molecular, developmental and evolutionary biologists.

Scientists themselves acknowledge the variation in their concepts, and the particular problems that can arise when scientists from different subdisciplines want to collaborate and communicate. Interdisciplinary work can therefore be hampered by conceptual variation and this is of particular concern in cognitive science since this is an interdisciplinary enterprise.

Much of the literature on memory refers to this problem of variation in MEMORY between different subdisciplines (e.g. Sutton (2004): 188, Wilson (2005), 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.' (*ibid*.: 211). [4]

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.

The relevance of the problem for interdisciplinary collaboration is highlighted in various places in addition to Sutton's paper quoted above (e.g. Hirst and Manier (2008): 197, Wertsch and Roediger (2008): 324, Barnier et al (2008): 34-35, Welze and Markowitsch (2005): 74, Figdor (2013), Roediger et al (2007)). Dudai, Roediger and Tulving argue 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

[4] For an excellent source of references for the memory concept in all its diversity, see John Sutton's bibliography on Memory (http://www.johnsutton.net/Memory.html, accessed 23/04/13).

venture' (Roediger et al, 2007: 1). They argue that what is required is 'a direct confrontation of the issue at what we regard as the most fundamental level of knowledge and analysis — the conceptual level' (*ibid*).

In summary, a study of scientists' concepts and the role they play is important because there is variation in these concepts, and this variation causes problems for science. More specifically, different scientists, particularly those from different subdisciplines, can have different concepts of the phenomena they are researching, and this can cause problems for collaborative work. This problem is acknowledged in the literature.

Before moving on, I want to clarify a couple of points with respect to what I have said so far. The first is that when I talk about "different concepts" here, I do not mean to ally myself to any particular theory of concept meaning. There are large issues in this vicinity concerning concept individuation and how those with different concepts can communicate.[5] Here I will not get into these complexities in the philosophy of language. On my view, difference is a matter of degree, so partial communication is possible when concepts are similar. This is the usual situation when scientists talk past one another; it is not the case that they are not communicating at all (no Kuhnian incommensurability here), merely that important misunderstandings may sometimes occur.

A second clarificatory point is to note that all this talk of concepts is of course not to say that there is not an objective world that science studies, or that we should not be interested in that world as philosophers. However, the problems of variation and communication I am interested in here are already at the level of concepts. Concepts are the means by which scientists theorise about and manipulate the world, and concepts mediate their communication with one another. A study of concepts can therefore contribute to the study of the world itself, and help to facilitate communication about it. Understanding the relationship between scientists' differing concepts can also help us to understand the relationships between the theories of different disciplines. Ultimately this can contribute to many of the major debates in the philosophy of science, such as those over reductionism, emergence, levels of explanation, etc. These debates are particularly relevant to the philosophy of the cognitive sciences (and the social and biological sciences) as "higher level"[6] sciences.

The study of scientists' concepts is important and worthwhile, and there is some recognition of the need for such a study in the literature. The modest goal of this paper is to discover the best method for carrying out this

[5] Thanks to an anonymous reviewer for raising the issue of concept individuation.

[6] Or special sciences, or more causally complex sciences etc. I don't mean to commit myself here to any particular claim about what makes these sciences differ from the physical sciences, or whether the difference is one of degree or kind.

study for MEMORY. The goals of work carried out using this method are more ambitious and include facilitating interdisciplinary work and communication within science, as well as contributing to longstanding debates in the philosophy of science.

The best way to investigate scientists' concept use is not immediately obvious. Which method is best may depend to some extent on the concept being investigated, the degree and kind of variation it exhibits etc. I will therefore focus on a specific example (MEMORY) in the remainder of the paper. I will now consider two alternative methods and how suitable they are for investigating the role of MEMORY as a tool in use by scientists in theory and practice.

2. 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 SGK (2004) to study the variation in GENE, a similar problem to our issue with MEMORY.

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 this situation. On this approach, scientists from a range of backgrounds are presented with a range of examples in a questionnaire, and must decide which of them conform to the concept and which do not, or which of several examples best conforms. The experimental method allows a large number of scientists' concepts to be probed and correlated with data about the subdisciplines in which they were trained and in which they work. Other information such as age and gender can also be collected to build up a full picture of the variation.

As well as describing the variation in the concept, SGK's work explains this variation using the idea of the *epistemic niche* inhabited by the concept. The epistemic niche consists of the needs the group of scientists have in their investigation (Griffiths and Stotz, 2008: 508). As these needs change over time, the concept changes to adapt to them. This can result in 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.' (*ibid.*: 508). They christen the study of this diversification *conceptual ecology*.

This vocabulary gives us another way to describe the pragmatic approach being taken here. In agreement with Haslanger, I said that the role the concept plays is of primary importance, and we should look for a concept that

could fill that role. In SGK's terms which I will use here, the epistemic niche is the role the concept plays (what scientists need it to do), and this niche shapes the concept.

It may appear that conceptual ecology is a purely descriptive tool, but this is not the case; there is also a normative dimension. Normative force is provided by the surrounding network 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 'It 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).[7,8]

This kind of experimental conceptual ecology would allow us to analyse the extent of the variation in MEMORY and explain it by using the notion of the epistemic niche. It could also suggest improvements to the concept or resolve confusions by analysing fit between concept and niche. However, 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.[9] The situation of answering a questionnaire is very different from the situation of designing and carrying out an experiment. I would argue that the situations are different enough to constitute partially different epistemic niches. We should therefore expect that different concepts may be most appropriate in each case. This is the very phenomenon we set out to investigate in the first place.

[7] 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.

[8] Haslanger and Fisher also talk about the normativity of the pragmatic approach (Haslanger, 1999: 466; Fisher, 2006: ch4).

[9] This is a general problem for 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.

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 epistemic niche is rich and multidimensional and cannot be fully captured in a questionnaire, however well designed. 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 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 answered; on this occasion they needed the concept to help them communicate a certain idea about their field of research to the questionnaire-setter.

Raising worries about ecological validity with respect to experimental philosophy is not new. Knobe and Samuels mention, with respect to their experimental studies on innateness, the possibility 'that the conditions that obtain in real scientific research are dissimilar to those found in our case-based condition'. They reference 'a strain of thought within the study of science that emphasizes…the ways in which the behavior of scientists is molded by characteristic features of their external situations (e.g., Kitcher, 1995; Mercier & Sperber, 2011).' (Knobe and Samuels, 2013: 84). I am pointing to a specific way in which the behaviour of scientists may be affected, namely the worry that a scientist's concept may not be stable across different contexts, and therefore may differ between the questionnaire and the research contexts.

SGK do acknowledge 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 apply their concept of the gene to deal with examples, rather than articulating a definition explicitly. This seems to show that SGK accept some kind of divergence of concepts depending on the context of concept application (response to a direct question about the concept, versus a scenario that demands use of the concept). It seems reasonable to suppose that different epistemic niches are responsible; the scientist has different aims and needs in each case.

Similarly I am arguing that the very fact of answering a questionnaire may constitute a different epistemic niche from that in which scientific work usually takes place. This may therefore be enough to cause the concept to vary. It is taking seriously the idea of the epistemic niche that brings

me to this conclusion, and the notion of the epistemic niche is the main focus of this kind of experimental approach.

This is not to say that the experimental approach is useless. 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 the results of the experimental approach may not be an accurate reflection of the variation in actual scientific practice. The next method aims to get around this problem.

3. 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 angle.

The method involves looking closely at particular pieces of current or recent research, treating them as case studies. 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; *Memory*, 2008, 16 (3), Special issue: "From individual to collective memory: Theoretical and empirical perspectives"). 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' (Collins and Evans, 2002) terms, she needs to develop at least interactional expertise in the science.

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 subdisciplines to see whether they have common concepts, and how they do or could communicate. 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 no single individual could develop enough expertise in all of the sciences of memory to carry out a full survey.

Case study work can be done within the framework of conceptual ecology, and SGK do mention the possibility of carrying out conceptual ecology by looking at published work from different scientific fields (SGK: 648). The emphasis on the epistemic niche can therefore be retained from the previous method, giving us the possibility of doing explanatory and normative work as well as description. This is the approach I advocate because it retains these advantages, while improving on the experimental method 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.

However, there are some apparent downsides of case-study-based methods. 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.' (*ibid.*: 186).

I believe this problem can be partially got around by careful choice of case studies. Case studies should be chosen to have the greatest predicted variation in the target concept along a particular dimension of interest. They can therefore be chosen to test a prediction about variation between subgroups, and investigation could falsify this prediction if the concepts are found to be similar after all. We can therefore examine whether different subgroups endorse different norms etc. as Machery and Cohen wish, but without statistically significant sample sizes.

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' (*ibid.*). This is a fair criticism and I think Machery and Cohen are correct 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.

However, this issue may be more serious when combined with the previous problem about detecting variation between subgroups. It may be that our investigation appears to confirm the predicted variation, but only because we have selected an article with an idiosyncratic concept compared to the subgroup to which the scientists belong. We then cannot correctly infer that there is variation between the subgroups from our contrasting case studies. The first part of the solution to this is to look at multiple case studies. It may not be possible to look at a statistically significant sample, but looking at a handful of cases will at least reduce the chances of having selected one or two idiosyncratic papers. However, it may be that some kind of implicit bias makes the experimenter choose a whole handful of idiosyncratic cases. How should this be overcome?

It seems that this is a good place for experimental and case study methods to work together. The case study approach can give detailed and in-depth analysis of contrasting concepts, and the experimental approach can be used to find out whether these concepts generalise beyond the cases examined. Although neither method is fool proof, if there is agreement between them about the concept possessed by members of a certain sub-group, the chances of either being in error are reduced.

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 (*ibid.*). 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. This will also help with the previous problem, since a broad range of case studies are unlikely to be idiosyncratic in similar ways. 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.

Carefully selecting case studies can therefore go some way toward alleviating the problems with the case study approach. In addition, the detail and depth provided by case study work makes the method amenable to the historical approach discussed in the next section, whilst the experimental philosophy method is not so obviously amenable. This is important because, as we will see, the historical approach allows further normative claims to be made than either the experimental philosophy or case study methods alone.

In summary, the case study method is a better approach than its critics suggest, and at least as good an approach as the experimental one. In many cases, both approaches can fruitfully be used together. After this rather abstract section, I will now give some more specific guidance on how to choose case studies for MEMORY and how they should be analysed using a historical version of the case study method.

4. How to Go about the Case Study Method

First, how should we select case studies? I said that case studies should be selected that are predicted to vary along the dimension(s) of interest. For MEMORY, some important dimensions of variation have already been identified in the literature. Some central examples are:

- Individual to collective memory: From memory as a property of individuals to memory as a property of groups. (e.g. Barnier et al., 2008: 36).
- Veridical to non-veridical memory: Must something be an accurate reflection of an event in order to count as a memory? And relatedly, must the rememberer believe that it is an accurate reflection? Research on reconstruction of memories is important here (e.g. that described in Campbell, 2006).
- Internal to external vehicles: From memory as something that can only be stored in brains, to something that is functionally defined and can be stored in any appropriate medium. Extended cognition research is of particular importance here (e.g. Clark and Chalmers 1998, Clark 2008).

If we are interested in the first dimension, appropriate case studies could be an experimental psychology study on individual recall in a laboratory, and a sociological or anthropological study of the memory of a tribe or nation.

If we are interested in the second dimension, case studies could be experimental psychological work on word recall, and discursive psychological work on the construction of a memory to meet certain personal or social needs.

To investigate the third dimension, case studies could include a neuroscientific study of the brain areas involved in memory, and a work in the embedded or distributed cognition tradition where people use external props and aids to remember.[10]

The case studies should be analysed carefully to determine the features of MEMORY in play in each case. These features should then be looked at alongside the epistemic niche in which the researchers were employing the concept in each case. In this way, we can hopefully get some idea of how the niche applies pressure to scientists' concepts, how the concepts adapt and whether they are well-adapted.

However, I think there is still room for improvement on this method. While it allows us to analyse fit with the epistemic niche, it says nothing about whether the niche is appropriate in the first place. This is a problem for both case-study-based and experimental conceptual ecology. It may be

[10] I have given two suggestions for each dimension here, but there is no reason to limit exploration of each dimension to only two cases, and the more that can be explored without sacrificing depth, the better.

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. Although what constitutes "good science" is notoriously hard to quantify, realist philosophy of science wants to say something about theory choice. As mentioned above, explanation and prediction are important goals of science, so getting the world right is an important pressure on the concept. However, there are always multiple competing explanations of the phenomena and we want a way to choose between them. There are also more pragmatic pressures that can shape scientific theory and practice, such as need for certain practical applications, availability of funding etc. These can also exert selective pressure on scientific concepts. Which factors should be allowed to exert such pressure, i.e. which would constitute parts of a "good" or appropriate niche? Answering this question allows us to make stronger normative claims than merely analysing fit between concept and niche.

Haslanger also suggests a normative approach to concepts. Like me, she suggests that it is our purposes in using the concept (the niche in the terminology I am using here) that are important, and that we need to ask whether those purposes are legitimate (Haslanger, 1999: 468). I aim to provide a more explicit account of how we should do this.

An example of current interest involving MEMORY is the use of external memory aids like mobile devices. For those working within the extended cognition tradition (e.g. Clark, 2008; Rowlands, 1999, 2010), the epistemic niche applies pressure to the concept to include extended examples within the memory concept. For their internalist opponents (e.g. Adams and Aizawa, 2001, 2008; Rupert, 2004), the niche applies pressure to the concept to exclude these same examples. This results in those working on the two sides of the debate having different concepts.

As I said above, this difference is a matter of degree, so it is not the case that the two sides are entirely talking past one another. There is enough overlap between their concepts to constitute a substantive debate. This debate is not a purely terminological dispute because how memory is conceived of will affect the theories and practices of the science in which the concept is embedded. This will have significant impact on how science progresses.

Both sides of the debate are producing explanations of the phenomena, but the kinds of explanations differ. It might be that both explanations are equally good in terms of fit with their respective niches.[11] In that case, the debate is over which niche is more appropriate for the progress of science.

When we have a complex of interacting epistemic and pragmatic factors, different groups of researchers may respond to different pressures, as in the

[11] I do not have space to defend this claim here. It is certainly plausible, but if it is not the case, other examples could be found which do meet this criterion.

example of the debate over extended cognition. The extended and non-extended concepts of cognitive states and processes (e.g. MEMORY) inhabit different but overlapping niches made up of these different factors. Neither the case study method nor experimental philosophy gives us a good way to decide which niche is better. We can isolate epistemic and pragmatic factors which shape the concept, but which of those factors are most important?

I claim that an historical version of the case study method can help. Griffiths and Stotz briefly mention taking an historical approach to concepts. They say that '[i]n the study of conceptual evolution, the history of genetics provides a 'conceptual phylogeny' of the gene.' (Griffiths and Stotz, 2008: 516) and note that there are a variety of ways of studying concepts, '[o]ne is via the history of science, an invaluable approach when a concept is changing over time.' (*ibid*.: 508). I am arguing that a study of the development of a concept can do more than help us track that change over time; it can also help us to make sense of current variation. The primary focus should be to trace the history of the epistemic niche because in this way we can make normative claims about the niche itself, not just the concept's fit with it. In other words, we should study the evolution of the environment in which current concepts are employed.

When good case studies have been identified and the epistemic niches of contrasting concepts investigated, the history of those niches should be traced. By thus situating the epistemic niche in its historical context, we can assess its appropriateness. The factors that make up the niche can be assessed in terms of how they came to be seen as important by the scientists in question. To come back to the example of MEMORY and extended vs. non-extended approaches, the relevant question would be: how did giving explanations in terms of extended cognition or purely internal cognition come to be seen as important? This is a large project that I will make no attempt to carry out here. I will just give an example of the sort of thing the method might achieve. [12]

Assume that science is currently internalist, and extended cognition is suggested as a new alternative. Consider the case where both extended and internalist cognition theories can provide equally good explanations and predictions of the phenomena. But extended cognition is found to have been developed largely in order to make money for the manufacturers of memory aids. The development of internalist cognition theory on the other hand was based only on its explanatory and predictive successes.[13] We do not consider the financial success of manufacturers to be

[12] I don't believe the following story for a moment; it is purely for illustrative purposes.

[13] Or on its explanatory and predictive successes plus some legitimate pragmatic factor such as helping amnesia patients. I do not mean to imply that only epistemic factors can be legitimate.

a legitimate factor in shaping scientific theory, so we have no reason to switch to talking in terms of extended cognition. Although this factor is part of the epistemic niche for the memory concept, it is not a legitimate part of the niche.

Why certain factors are viewed as legitimate and others are not is a very difficult question with a long history in the history and philosophy of science. Legitimate factors are often taken to be those that are "internal" to science, while "external" factors are illegitimate (see Shapin, 1992 for an overview and a relatively sympathetic perspective on the debate). However, this is not as simple as dividing "scientific" factors (such as empirical adequacy, explanatory power, elegance and simplicity) from social and political factors. Many factors that are social or political are perfectly legitimate pressures on the development of science, for example new possibilities arising due to technological advances, or the need for certain practical (e.g. medical) applications. Additionally, what it means to be explanatorily powerful or to be a simple or elegant theory is partly determined by social and political factors. But despite the difficulty in identifying them, there are nonetheless some factors which we view as legitimate pressures on the development of science and others which we do not.

The reasons why we do not consider certain pressures as legitimate factors in the niche is a product of the development of the methods and practices of science itself. A historical approach to a case study can trace this development of science over time, i.e. the development of *what counts as a good scientific theory* in the relevant subdiscipline. On my own view, the outcome of this development is contingent. For a similar view, see Shapere (1984, 1986, 1987). However, you need not accept this contingency to see the worth of the approach I am advocating here. It might be that there are certain factors that are legitimate *simpliciter*, and over time science is getting better at recognising these factors. A historical study of the niche can still be useful because it reveals the discovery of these factors.

5. Conclusion

I have argued that a historical approach to carefully selected case studies is the best way to study MEMORY. Case studies should be chosen to maximise expected variation in the concept along the dimension of interest (e.g. from internal to external vehicles if you are interested in extended cognition). The epistemic niche (the needs of the scientists using the concept) can be identified and used to explain the variation, and normative claims can be made based on the fit between concept and niche. The development of the niche over time can then be traced, to see how different factors came to be seen as most important to different subdisciplines. This allows

a further normative claim to be made about the appropriateness of the niche itself. The concept can then be assessed in its historical context according to whether the niche is made up of legitimate factors that should be allowed to apply pressure to our scientific concepts or not. This may allow disputes over the concept to be resolved. For example, if there is a dispute over whether memory should be seen as extended or purely internal, it may be that one of these concepts of memory has developed as a result of a niche containing illegitimate factors and the other has not.

The methods advocated here can therefore help to alleviate conceptual disputes in the sciences. This can help to facilitate collaborative work. There is a particular need for this between scientific subdisciplines because different subdisciplines often have different concepts and can end up talking past one another.

In addition, work carried out using these methods can contribute to questions in the philosophy of science, for example those over reductionism or levels of explanation, because it can help to clarify the relationship between different concepts at different levels, e.g. social and individual memory. I have not said much about these wider issues here. How much light can be shed on them will largely depend on the results of the study of MEMORY using these methods. My aim here has just been to identify the best method to make sense of the variation in the concept of memory in science.

References

Adams, F. and Aizawa, K. (2001), "The Bounds of Cognition". *Philosophical Psychology*, **14**, pp. 43-64.

Adams, F. and Aizawa, K. (2008), *The Bounds of Cognition*, MA, Oxford, Victoria: Blackwell.

Barnier, A. J., Sutton, J., Harris, C. B. and Wilson, R. A. (2008), "A conceptual and empirical framework for the social distribution of cognition: The case of memory". *Cognitive Systems Research*, vol. 9, issues 1-2, pp. 33-51.

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.

Campbell, S. (2006), "Our Faithfulness to the Past: Reconstructing Memory Value". *Philosophical Psychology*, vol. 19, no. 3, pp. 361-380.

Clark, A. and Chalmers, D. (1998), "The Extended Mind". *Analysis*, **58**, pp.7-19.

Clark, A. (2008), *Supersizing The Mind: Embodiment, Action, and Cognitive Extension*. New York: Oxford University Press.

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. (2002), "Interlevel Experiments and Multilevel Mechanisms in the Neuroscience of Memory". *Philosophy of Science*, **69**, pp. S83-S97.

FIGDOR, C. (2013), "What is the "Cognitive" in Cognitive Neuroscience?". *Neuroethics*, vol. 6, issue 1, pp. 105-114.

FISHER, J. (2006), *Pragmatic Conceptual Analysis*. PhD thesis, University of Arizona, USA. Available from ProQuest Dissertations & Theses (PQDT), 2007.

GRIFFITHS, P. E. (2002), "Lost: One Gene Concept. Reward to Finder". *Biology and Philosophy*, **17**, pp. 271-283.

GRIFFITHS, P. E. and STOTZ, K. (2008), "Experimental Philosophy of Science". *Philosophy Compass*, **3**, pp. 507-521.

HASLANGER, S. (1999), "What Knowledge Is and What It Ought to Be: Feminist Values and Normative Epistemology". *Noûs*, Vol. 33, Supplement: Philosophical Perspectives, 13, Epistemology, pp. 459-480.

HIRST, W. and MANIER, D. (2008), "Towards a psychology of collective memory". *Memory*, **16** (3), pp. 183-200.

HUTCHINS, E. (1995), *Cognition in the Wild*. Cambridge, MA: MIT Press.

KITCHER, P. (1995), *The advancement of science: Science without legend, objectivity without illusions*. 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.

MACHERY, E. and COHEN, K. (2012), "An Evidence-Based Study of the Evolutionary Behavioral Sciences", *Brit. J. Phil. Sci*. **63**, pp. 177-226.

MERCIER, H. and SPERBER, D. (2011), "Why do humans reason? Arguments for an argumentative theory". *Behavioral and Brain Sciences*, **34**, pp. 57-74.

ROEDIGER III, H. L., DUDAI, Y. and FITZPATRICK, S. M. (eds.) (2007), *Science of Memory: Concepts*, USA: Oxford University Press.

ROWLANDS, M. (1999), *The body in mind: understanding cognitive processes*. Cambridge: Cambridge University Press.

ROWLANDS, M. (2010), *The new science of the mind: from extended mind to embodied phenomenology*, Cambridge: MIT Press.

RUPERT, R. (2004), "Challenges to the Hypothesis of Extended Cognition". *The Journal of Philosophy*, vol. 101, no.8, pp. 389-428.

SHAPERE, D. (1984), "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.

SHAPERE, D. (1986), "External and Internal Factors in the Development of Science". *Science and Technology Studies*, vol. 4, no. 1, pp. 1-9.

SHAPERE, D. (1987), "Method in the Philosophy of Science and Epistemology: How to Inquire about Inquiry and Knowledge", in N. J. Nersessian (ed.) 1987. *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.

STOTZ, K., GRIFFITHS, P. E. and KNIGHT, R. (2004), "How biologists conceptualize genes: an empirical study". *Studies in History and Philosophy of Biological and Biomedical Sciences*, **35**, pp. 647-673.

SUTTON, J. (2004), "Representation, reduction and interdisciplinarity in the sciences of memory" in H. Clapin, P. Staines and P. Slezak (eds.) 2004. *Representation in mind: New approaches to mental representation*, Oxford: Elsevier.

SUTTON, J. (2007), "Integrating the Philosophy and Psychology of Memory: Two Case Studies" in M. Marraffa, M. De Caro and F. Ferretti (eds.), 2007. *Cartographies of the Mind: Philosophy and Psychology in Intersection, Volume 4 of Studies in brain and mind*, Dordrecht: Springer, pp. 82-92.

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.

WILSON, R. A. (2005), "Collective memory, group minds, and the extended mind thesis". *Cognitive Processing*, vol. 6, no. 4, pp. 227-236.

WITTGENSTEIN, L. (1953), *Philosophical Investigations*. G.E.M. Anscombe and R. Rhees (eds.) G.E.M. Anscombe (trans.), Oxford: Blackwell.

Ruth HIBBERT
reh30@kent.ac.uk