The Journal of Mind and Behavior
Autumn 2009, Volume 30, Number 4
Pages 249–266
ISSN 0271–0137

Distributed Mental Models: Mental Models in Distributed Cognitive Systems

Adrian P. Banks and Lynne J. Millward

University of Surrey

The function of groups as information processors is increasingly being recognised in a number of theories of group cognition. A theme of many of these is an emphasis on sharing cognition. This paper extends current conceptualisations of groups by critiquing the focus on shared cognition and emphasising the distribution of cognition in groups. In particular, it develops an account of the distribution of one cognitive construct, mental models. Mental models have been chosen as a focus because they are used in a number of theories of high level cognition from different areas of research such as cognitive science and human factors and so the implication of this development is wide reaching. This paper reviews the unconnected literatures on distributed cognition and mental models and integrates them in order to extend the theory of mental models to distributed cognitive systems such as groups. The distributed cognition literature is reviewed and the importance of considering the group as single cognitive system is adopted. A range of mental model theories are reviewed leading to the conclusion that they all have, in some form, the central feature of a mapping onto the cognitive system. Combining these two ideas, it is proposed that the model can be a mapping onto the whole group, if the information is distributed appropriately and the connections between parts of the model maintained through communication. This cognitive construct is referred to as a distributed mental model. Implications and applications of this theory are discussed.

Keywords: distributed cognition, mental models, shared cognition

Recently, several different approaches have emerged which treat the group as a unit of analysis for information processing in place of the traditional focus on the individual (e.g., Hinsz, Tindale, and Vollrath, 1997; Hutchins, 1995a; Larson and Christensen, 1993). Although there are differences between these perspectives, they are similar in that they consider the transmission and manipulation of information between group members to be part of a cognitive process

Requests for reprints should be sent to Adrian P. Banks, Department of Psychology, University of Surrey, Guildford, Surrey, GU2 7XH, United Kingdom. Email: a.banks@surrey.ac.uk

akin to any similar process within an individual.[1] Accepting that the group is an information processing unit permits the events that occur between group members to be studied as part of the cognitive processing of a task rather than solely as a social phenomenon. Much high level information processing, especially in relation to real-world tasks, relies on cognitive representations in the form of mental models. In group contexts, these cognitive representations have been studied as "shared mental models." However, we argue that there are many situations where mental models are more efficiently distributed amongst group members than shared by them. In practice, complex tasks more often than not require the interdependent application of different (e.g., specialist) knowledge and skills. Drawing on the distributed cognition approach (Hutchins, 1995a) in combination with the concept of mental model, we propose a "distributed mental model" theory.

Group Level Cognition

The notion that a group engages in mental activity that can be analysed in its own right, in addition to the analysis of individuals, has a long history. The concept of a group mind was prominent in the later nineteenth and early twentieth centuries in the writings of McDougall, Wundt, Durkheim, and Le Bon, but was used to describe how otherwise rational individuals could become "disinhibited" and "irrational" within the collective. Accordingly, groups were not considered viable cognitive entities and so historically the individual has been the dominant unit of analysis in cognitive science.

Since the late 1980s, however, the idea that groups can be studied as cognitive systems has begun to take hold in the effort to better understand group and team performance. Larson and Christensen (1993) for instance, introduced the term social cognition to refer to:

> those social processes (e.g., introducing information into a group discussion) that relate to the acquisition, storage, transmission, manipulation and use of information for the purpose of creating a group-level intellective product. (p. 6)

Here the idea that group interactions can be cognitive processes is emphasised, contrasting with the more usual meaning of the term social cognition which pertains to cognition about social events. Later, Hinsz et al. (1997) introduced a more detailed specification of information processing in groups. This approach to conceptualising group level cognition is to identify the components of a generic information processing model, divorced from who or what processes that information. Generic processes include such things as attention, encoding,

---

[1]Cognitive processing is defined here as information processing and so these terms will be used interchangeably throughout this paper.

storage, and retrieval processes. Any activity that occurs within or amongst group members that fulfils one of these functions (e.g., through social processes) is an instance of group information processing.

One example of how this might work is the concept of transactive memory (Wegner, 1987). This group level model of memory suggests that the group is not simply comprised of individuals who each try and remember as much information as they can, the sum of which is the group memory. Rather, groups use strategies to maximise the memory capacity of the group. Each member of a group has specialist categories of information which she remembers and these categories are divided between group members through discussion. The group then shares knowledge about who knows what. As a result, information is efficiently stored because the person most likely to remember a fact does so, and the redundancy in the time and effort involved in remembering is reduced. As a result, the information recalled by the group is often larger than that of an individual. The key feature of this theory is that memory performance in groups is not explained in terms of social influences such as social facilitation (Zajonc, 1965) or social loafing (Latane, Williams, and Harkins, 1979). Instead, differences between individual and group performance are explained in terms of the efficient processing — storage, transmission etc. — of information by the group.

Research on information sampling also describes how the group decision is influenced by sharing information with the group (e.g., Stasser and Stewart, 1992; Stasser and Titus, 1985). Groups are more likely to discuss information that is shared by all group members. When information is not shared, it is less likely to be used in the decision-making process and this omission can bias the outcome. This example demonstrates that there are phenomena which only become apparent when information processing is considered at the group level, highlighting the importance of studying this level of analysis.

Several authors have sought to define what group cognition is. Typically the fact that cognition is shared has been identified as the defining feature. For example, Larson and Christensen state that "an individual could hardly be expected to introduce into a group discussion a piece of information that she cannot recall. And yet, simply recalling that information is not enough. It must also be shared with others" (p. 7). That is, they argue that only by sharing can cognition become social. Hinsz et al. (1997) also claim that "At the group level, information processing involves the degree to which information, ideas or cognitive processes are shared, and are being shared, among the group members" (p. 43). Tindale and Kameda (2000) maintain that "the concepts "shared" and/or "sharing" are what make group information processing possible, and distinguish it from individual-level processing" (p. 124). However, whilst sharing information in groups may be important for some tasks, the idea that it is the *sine qua non* of group level cognition in groups is too restrictive. Many important information processing tasks in groups may depend, on the contrary, on what is "unshared" cognition.

There are in fact two meanings of shared cognition to consider. Information processing can be shared in the sense of transmitting information between group members (i.e., shared through communication) and it can also be shared in the sense that all group members hold the same information (i.e., shared knowledge). However, there are several counterexamples to both of these as defining features of group cognition.

Examining transactive memory provides a counterexample to the first meaning. Whilst this does involve communicating information about the memory task, it also depends on the fact that information is not always shared (i.e., not discussed nor held by everyone in the group). In complex tasks requiring specialist input, much of the information is only ever held by one person and, unless relevant, will not be communicated. It is nonetheless information that is in the system somewhere, if required. The efficiency of transactive memory is due to each individual only searching their own expert memory for the required information. Thus the efficient memory search occurs at the individual level. The subsequent sharing of the solution is important but is only part of the process. To fully understand transactive memory it is necessary to describe the memory search at the individual level as well as information sharing at the group level. Hence sharing information is only one part of the information processing in groups.

A counterexample to the second meaning can be taken from many real world tasks. In these complex situations it is unlikely that group members will have identical (shared) information. In laboratory studies of group decision making, information consists of a small number of discrete facts which all members can readily appreciate. However, in a more complex task this may not be the case. For example, an engineer and a designer may be discussing a car and one member of this group may mention a concept such as the shape of the wing mirror. For the designer this concept has meaning associated with aesthetics whereas for the engineer there may be meaning associated with aerodynamics or field of view. The basic concept is shared but the full implications of the meaning of the concept are quite different. It is likely that the notion of shared information as representations held in common by group members is not typical in everyday situations, in which case it is problematic to rely on shared information processing as a defining feature of group level cognition.

These two counterexamples indicate that neither sense of sharing cognition fully captures what is meant by group cognition. We therefore argue that attempting to define what is unique about group cognition, especially by relying on shared cognition as the defining feature, is not possible. Instead, we argue groups should be studied simply as information processing systems that are different to individuals but are not special in their own right. A framework that adopts this stance and could readily be applied to information processing in groups is distributed cognition.

Distributed Cognition

The individual is the conventional unit of analysis for cognitive science. Distributed cognition can simply be described as an approach which applies classical cognitive science principles, namely that cognition is information processing, to a unit of analysis other than the individual. Commonly the unit of analysis is larger than an individual; it might include the environment, artefacts in the world, social interactions, and culture. The implication of this approach is that studying the individual alone does not present a full picture of the phenomenon; some effects might be omitted or misattributed to individuals when they are in fact properties of a larger system. It has been argued that this approach could demand a reappraisal of established theories if indeed they are based on a misattribution of cognitive processes (Hutchins, 1995a). There are two related theoretical implications which follow from expanding the unit of analysis for describing and explaining a cognitive process. Firstly, the boundary becomes larger and so other people and artefacts become part of the process and engage in some computation themselves. This may reduce the individual task load, but it adds the problem of communicating and coordinating with other agents in the system. In addition, the range of mechanisms which are thought to be involved in the cognitive process may expand to include other artefacts. Secondly, the nature of the computation itself may change. For example, solving a problem using tools may transform it such that different operations are required compared to solving the problem without tools. The theory outlined here mostly involves the former so only this literature will be reviewed. (See Hutchins [1995a] for more detailed comment on the latter as distributed cognition and also Clancey, [1997], Greeno and Moore [1993], and Suchman [1987] for the similar concept of "situated cognition.")

In his portrayal of distributed cognition, Hutchins (1995a, 1995b) focuses mainly on the representations themselves. He suggests that information can be represented in different media, but it is clear that he considers all of these media equal in status, including the brain. The object of study is the propagation of representational states through different representational media and the process by which these media coordinate to allow this propagation to occur. As the representational states move between different media, different agents perform the computation, and this can influence what computation takes place. In principle, the representations could travel through any number of representational media, so there is no obvious boundary to this unit of analysis. The boundary can only be defined in terms of the representational media brought into coordination with each other thereby forming a system separate from those that are not.

In an ethnographic study to illustrate this process occurring naturally, Hutchins (1995a) describes the bridge of a U.S. Navy ship as it approaches harbour. Navigating the ship through the correct channel requires expert navigation by a team and is a complex cognitive task. Within the distributed cognition framework this is treated as a single task. The procedure, much simplified, involves taking bearings on each side of the boat and communicating these to the navigator who plots their location on the chart. The information being represented is the location of the boat; this state must be propagated from the people who take the bearings to the chart. Initially the boat is represented as a physical configuration of the alidade (compass) which is converted to a number (the bearing). During this propagation process the representation alters in order to facilitate the task. The analogue representation of the alidade is transformed into a digital representation which is easier for verbal communication and calculation. This is then communicated to the navigator who converts the number into a physical configuration of the hoey (protractor) and draws a line on the chart. This process is repeated for two other bearings so that three lines are drawn on the chart. Where they cross represents the location of the ship. Several representational media are used here and the propagation of the representational state between them is essential for computing the location of the boat. It is preferable to have bearings from both sides of the boat and it is essential to have these recorded at as nearly the same time as possible because the boat is moving. Thus at least two people are required, one for each side of the boat. These two bearings, or representations of the position of the boat, must be combined so they are propagated from the alidade to the navigator who combines them. So, in studying the navigation task the importance of the representations and where they are is essential to understanding how the whole system operates, along with the computations that are performed using the representations at each stage of the process.

Rather than focusing on the group level per se, Hutchins focuses on the information processing that takes place, the representations that are used in the process, how they are transformed, and what representational media they occur in. Thus, information processing may occur partly within an individual as she thinks through a problem, partly collectively in a group discussion and partly with external artefacts such as a piece of technology. Therefore, although this approach can be used to conceptualise information processing in groups, it does not define group level cognition. Indeed, it strongly questions even individual level cognition as a useful level of analysis.

The distributed cognition approach provides a useful framework for studying information processing, but it does not offer any precise predictions about how information processing takes place. The question of what representational media are used, and how, remains open depending on the task that is being studied. We argue that high level cognition frequently relies on cognitive representations in the form of mental models, especially in real-world tasks.

The Concept of Mental Models

The concept of a mental model as a cognitive representation used in thinking has a long history. Craik (1943) introduced the term and it has subsequently been used as a fundamental concept in many theories of reasoning, such as those of Johnson–Laird (1983). The concept has also flourished in the fields of applied cognitive psychology and human factors. It is employed to explain a wide range of human behaviour, for example users' interaction with technology e.g., Gentner and Stevens (1983), Norman (1988), naturalistic decision making e.g., Klein (1993) and many others.

One difficulty in using the term mental model is that there are wide differences in conceptualisation (e.g., Wilson and Rutherford, 1989). Moray (1999) argues that these are due to the range of contexts in which the construct is applied, not the construct itself. The context-specificity of most mental model research is illustrated by studies such as Kieras and Bovair (1984) in which the role of mental models in operating a simple machine is investigated. In this they define a mental model as "how a device works in terms of its internal structures and processes" (p. 255). This definition is highly specific to its field and so does not allow the same concept to be used in any other way apart from the operation of a device. In order to apply mental models to groups a more generally applicable theory is required. This section will review three influential theories of mental models in order to establish such a general theory. The theories considered here are developed by Craik (1943), Johnson–Laird (1983), and Moray (1999).

Craik (1943) is widely attributed as the first psychologist to describe the use of models in reasoning. These he proposed were used in order to predict events, which he considered to be a fundamental property of thought. The example he gives is of building a bridge. This is not done haphazardly, choosing materials at random. Rather, he claims, a model can be constructed and tested mentally in order to ensure a sufficient safety factor is built in, instead of actually building a bridge and waiting to see if it collapses. Thus Craik is referring to a purely internal, qualitative process of modelling. He also offers a definition of models used in this process. Interestingly, this is not a definition specific to mental models, it applies to all models. Hence Craik defines a model as:

> any physical or chemical system which has a similar relation-structure to that of the process it imitates. By "relation-structure" I do not mean some obscure non-physical entity which attends the model, but the fact that it is a physical working model which works in the same way as the process it parallels, in the aspects under consideration at any moment. (p. 51)

It seems from this definition that the model is isomorphic (i.e., structurally identical) to the process it is modelling, at least in the specific aspects under consideration. However, later on, Craik suggests that as a model is an analogy it will eventually break down at some point, revealing properties not found in the process it is imitating. In other words, the model cannot be identical to reality. Craik's conceptualisation can perhaps be thought of as a rough-and-ready kind of model, constructed with pragmatic aims. The intention to build a detailed model is there, but the implication is that it is unnecessary to bother with any aspects not of direct relevance. Realistically the model must be expected to fall down at some level of analysis; the aim is to build a model which is accurate for the purposes it is to be used for.

In contrast, Johnson–Laird (1983) developed a more detailed and well-specified account of the construction and manipulation of mental models in reasoning. There are three reasons for the greater detail of his account. Firstly, the principles of cognitive science applied in the development of the theory involved the construction of several computational models, which lead to an emphasis on effective procedures for computing the models. A benefit of this is that the processes hypothesised must be fully specified in order for the computational model to be implemented. This means it is difficult to create anything but an explicit and well-specified theory.

Secondly, Johnson–Laird lays out several principles governing the construction of mental models. These combine to generate mental models that are parsimonious representations. The principle of constructivism states that "a mental model is constructed from tokens arranged in a particular order to represent a state of affairs" (1983, p. 398). The principle of structural identity states that "the structures of mental models are identical to the structures of the states of affairs, whether perceived or conceived, that the models represent" (p. 419). All of the structural relations in the model play a symbolic role. The principle of economy states that "a description of a single state of affairs is represented by a single mental model even if the description is incomplete or indeterminate" (p. 408). Therefore the arrangement of the tokens in the model is isomorphic to one particular state of affairs that it is modelling. Unlike Craik's theory there is no proviso of the model breaking down at a certain level of analysis or of only some aspects of the state of affairs being modelled. This means that a model is an economical and accurate representation of the object modelled.

A third reason for the precision of Johnson–Laird's theory is that he chose to study well defined tasks and therefore was able to collect a range of experimental data to establish the cognitive processes used to reason about these tasks in detail. For example, an important area of his research investigated syllogistic reasoning (e.g., Johnson–Laird and Bara, 1984). Syllogisms are created by combining a limited set of premises, for example "All artists are beekeepers. Some beekeepers are chemists. What follows?" By arranging all syllogistic premises in every possible combination, sixty-four different problems can be created and the potential number of mental models required to represent each of these can be calculated. For the syllogisms with valid conclusions a maximum of three models are required. This is a lot for the limited capacity of working memory to manipulate. However, it is much more reasonable to consider the exact models held by participants when reasoning syllogistically than in other tasks for which mental models are used.

For example, nuclear power plants have in the order of forty-five degrees of freedom in a complex, dynamic system. To analyse this problem to the same degree of thoroughness would be a difficult if not impossible task (Moray, 1997). In short, Johnson–Laird provides a well-specified account of the use of mental models in reasoning. He is able to do this because of the emphasis placed on computational modelling of the reasoning process, the general principles established for model construction and the well-specified problems on which he has focused his research. Therefore Johnson–Laird's position is not necessarily incompatible with the other theories of mental models although it may appear different as a result of these different emphases in its development.

Meanwhile, Moray (1999) approaches mental models from a human factors position, although his aim is explicitly to unify the field. He claims that mental models are best defined as representations that are "homomorphic mappings" of real systems. This term is taken from Ashby's (1956) work on models. Essentially, a model is a homomorphic mapping if it preserves the elements in a system and the relations between them to a certain level of abstraction. Elements are not deleted in the mapping, they are simply represented more abstractly. For example, a driver may know how a car engine is related to other parts of the car, but may not know all of the elements within the engine. These have been abstracted to a single element "engine" that has the features of an engine and is causally linked to the wheels etc. allowing understanding of how a car functions, to a certain level of abstraction. This abstracted representation is a homomorphic mapping from the physical car.[2]

Despite both Moray's and Johnson–Laird's claims for universal mental model theories, some striking differences between them can be identified. Most obviously Johnson–Laird's theory requires models to be isomorphic to the situations that they represent whereas Moray's theory requires them to be homomorphic. Indeed, much of each theory's explanation of characteristic human behaviour follows from its chosen form of representation. In the case of Johnson–Laird, reasoning occurs through the construction of one or more mental models, each depicting a different situation. Where multiple mental models are required, humans' limited working memory capacity is taxed and according to Johnson–Laird this is a source of errors. But if these models were homomorphs, as suggested by Moray, then there would only be a single mental model constructed covering

[2]"Homomorphism" should not be confused with "homeomorphism" which is the term used to describe icons by Johnson–Laird (2006, p. 435). Homeomorphisms are topological isomorphisms.

several situations (the many-to-one mapping) and some of the cognitive load would be reduced. Hence working memory capacity would not be exceeded and this limitation would not be a cause of error. Instead, errors would be caused because the mental model is only an imperfect representation. This is the explanation offered by Moray for errors in performance. Therefore the different forms of mental models explain behaviours in different ways.

Much of these differences, it can be argued, may be largely a function of the task to which the mental model is applied. Moray's example of a mental model "*par excellence*" (Moray, 1997, p. 279) is the operator of a complex system such as a process plant. This means the model will be repeatedly used over a long period of time and may well be stored in long term memory. There are many opportunities to discover if different models are effective through interactions with the system and it may well be that some homomorphisms will lead to appropriate behaviour, at least in many situations. They will therefore be reinforced. In contrast, deductive reasoning tasks in the laboratory will often only be addressed once. This does not allow much opportunity for the model to be stored in long term memory, nor does it allow repeated feedback to find models that will lead to the correct answers without using isomorphic models. It is also less likely that effective homomorphs exist for these problems. The form of the mental model develops in response to the nature of the problem it is being used for and therefore apparent differences arise between the forms of mental models. But this does not mean that they are fundamentally different.

Thus, despite the apparent differences, there is a central idea which captures the key features of mental models in all of these areas. The concept common to all of the theories is that the mental model is a mapping of a system in the world onto the cognitive system. That is, a mental model is a representation of the elements of a system and the relationships between them in which each of these components are analogous to those in the system being modelled. Therefore the structure of the mental model directly corresponds to the structure of the real system unlike other representational forms where this can be lost such as language or predicate calculus. The detail of this mapping may vary according to the domain studied. Mental models of very complex systems are likely to contain more abstracted elements than mental models of simpler systems and are therefore homomorphic mappings. Mental models of smaller, well defined domains may be isomorphic. But there will always be a correspondence between parts of the real system and the model. Thus the main function of a mental model is as an analogous representation of reality.

Different tasks may result in the mental model being used in different ways, such as in reasoning or decision making. Deductive reasoning may rely on the construction of multiple mental models whereas predictions about causal systems may rely on transitions from one state of the model to another. As a result, differences may arise such as the source of errors in thinking. However, how the mental model is used in a cognitive process is distinct from the nature of the representation itself. Here we have focused on the defining feature of the representation and this remains the same in all of these applications. It only remains to consider how mental models, as defined here, might operate in distributed cognitive systems.

Distributed Mental Models: A General Theory

The central idea of mental models can now be combined with the framework of distributed cognition to propose the concept of a "distributed mental model." This concept provides a basis for understanding high level cognition in distributed systems, specifically cognition which relies on mental models including reasoning, naturalistic decision making, the cognitive basis of teamwork, etc. Many of these applications will be in groups, but distributed cognition is not limited to groups and incorporates any cognitive artefacts, the environment, culture, social factors and other elements which might be adopted into the unit of analysis. In short, the concept of distributed mental models provides a basis for understanding much high level cognition in all of these diverse systems too.

As was noted by Craik (1943) and Moray (1999) in particular, a model is not exclusively a mental construct. Models take all manner of forms, for example, mathematical or computer based. Therefore, there is no requirement for a model to be instantiated solely in the brain as models are not conceived fundamentally as cognitive representations. They have simply been used that way. We can therefore look at the model as the unit of analysis in its own right and assess the information that it processes rather than limiting the research to only those models that are instantiated in a single individual.

The central feature of a mental model is that it is a representation of the elements of a system and the relationships between them that is analogous to those in the system being modelled, although parts of the system may be represented at different levels of abstraction. That is, there is a mapping from each part of the system (the source of the model) onto the model, although some details may be lost as the model is more abstract. The distributed cognition approach applied here proposes that the medium in which this representation exists need not be a single individual. Instead, the model could be a mapping onto a group of people, or a group plus cognitive artefacts, and each person could hold part of the model. If the connections between the different parts of the model are maintained where necessary, e.g., by communication between the group members, then this model can be treated as a single unit of analysis rather than taking the individual who holds only part of the model as the unit of analysis. A model mapped onto a group in this way will be referred to here as a "distributed mental model."

For example, Banks, Macklin, and Millward (2002) studied causal reasoning about ecosystems in groups. The ecosystem used in this study was comprised of a number of plant and insect species and the predator–prey relations between them. Figure 1 illustrates a typical ecosystem used in the study. A line connecting two species indicates a predator–prey relationship.

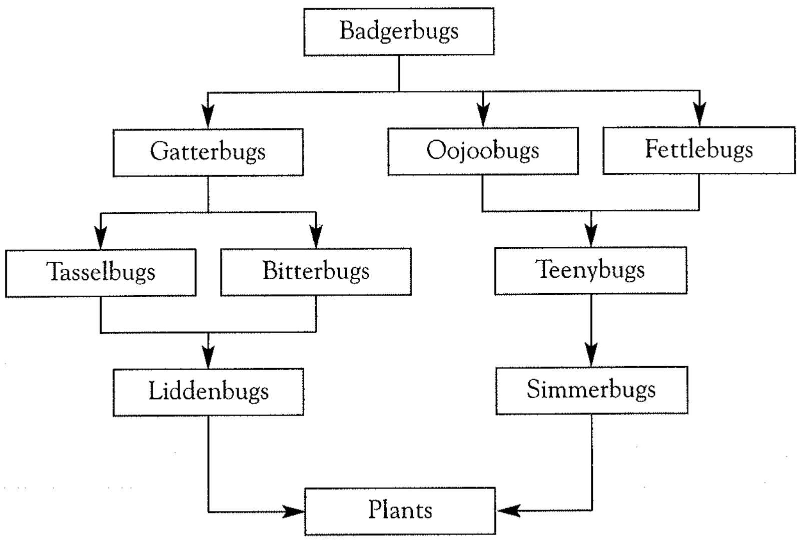

Figure 1: Typical ecosystem studied by Banks, Macklin, and Millward (2002).

In the experiment, groups were given diagrams illustrating an ecosystem and were asked to generate predictions about how the population levels of species within the ecosystem would change over time in response to changes in population numbers of certain species within it. However, the information about the ecosystem (i.e., the model of the ecosystem) was distributed amongst the group. This was achieved by giving one group member a full account of half of the ecosystem and only an abstract level of information about the second half, whereas the other group member was given detailed information about the second half of the ecosystem and only abstract information about the first half. Examples of these diagrams are presented in Figures 2 and 3. One group member was given Figure 2 and the other Figure 3.

The entire model of the ecosystem is known by the group, but it is distributed between the group members. In order to form predictions about the ecosystem as a whole, participants made predictions individually about parts of the ecosystem that they alone knew about and these fed into predictions made collectively about the parts of the ecosystem that they shared with their group

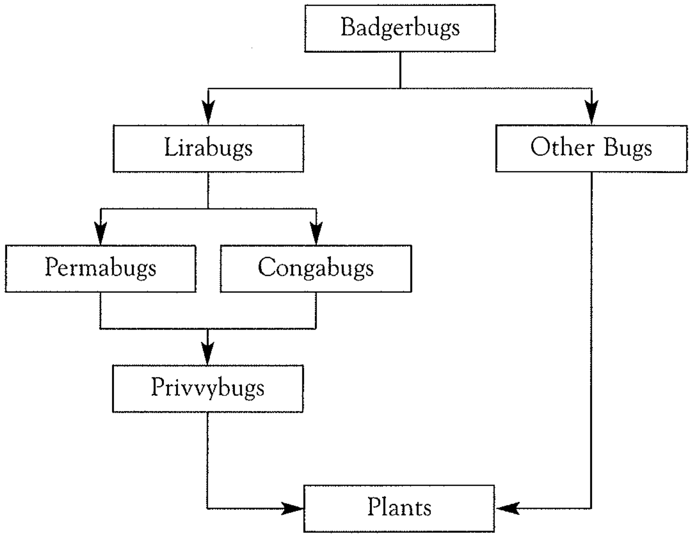

Figure 2: Partial representation of an ecosystem used by the first group member.

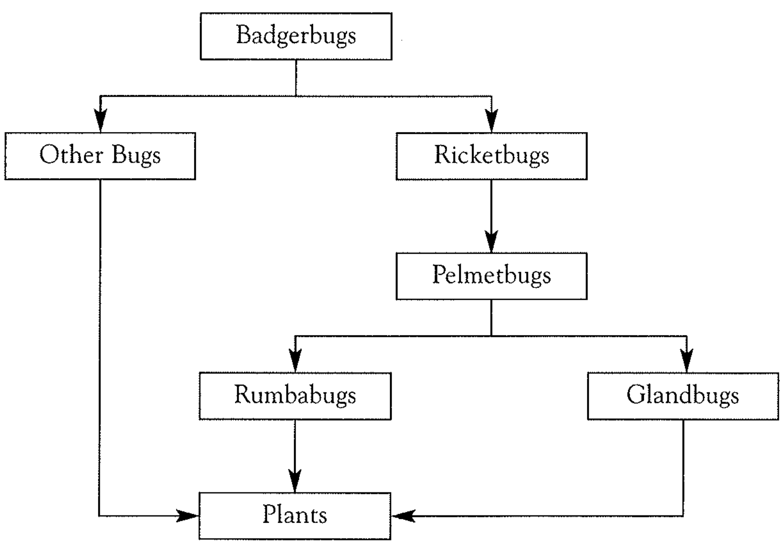

Figure 3: Partial representation of an ecosystem used by the second group member.

member. Hence the whole model was used to generate predictions, but the cognitive process was partially completed by individuals and partly collaboratively, according to the mapping of the model onto the group.

The ecosystem example demonstrates one way in which information can be distributed in a group. However there are many ways in which the model could have been distributed, that is, the ecosystem could have been divided in different ways, and different distributions will have different implications for the cognitive process.

One important feature of the distribution is the division of the model between people. For example, Banks and Millward (2000) describe how models can be decomposed into relatively independent components, referred to as "*modules*" by Chandrasekaran (1981). The modules are subsystems which have a relatively high number of intra-connections, but a relatively low number of interconnections. This means that group members who hold parts of the model which are whole modules will be able to complete some reasoning independently as modules allow relatively autonomous cognitive processing, by definition. As a result different group members can be engaged in reasoning in parallel, simultaneously forming inferences about different parts of the model. These can then be combined at a later point. In contrast, if the module was divided such that group members held only parts of modules then individuals could not engage in much independent reasoning and would be required to discuss much more.

For example, in Banks et al. (2002) the species Gatterbugs, Tasselbugs, Liddenbugs, and Bitterbugs shown in Figure 1 form a module. There are complex relations between them but comparatively simple inputs and outputs. The ecosystem in Figure 1 is arranged into two distinct modules and this is what enabled the group members to reason independently about most of the species; only combining to discuss the shared species, the plants and Badgerbugs. A benefit of this is that less communication is required. As communication is time consuming and cognitively demanding the whole process can be quicker. A disadvantage is that there are reduced opportunities to check what other group members are thinking and detect potential errors. Conversely, if many elements in the model or relationships between them are split between group members, these will all need to be discussed in order to ensure that the predictions can be made. As a result the process will be slower and potentially prone to error because of the cognitive demands of this discussion. However, this discussion could prove a valuable check of group members' thinking, depending on the context and model.

A second important element influencing the distributed cognitive process is the timing of the process. Given a modular distribution of the mental model, group members will work in parallel and explicitly discuss only parts of the model. It is important for effective thinking that when they do discuss the model they do so in a coordinated fashion. As the cognitive process unfolds over time and possibly at different rates amongst group members it is important that when they do communicate they refer to the same thing and have not lost coordination by working at different speeds. For example, in Banks et al. predictions were made by considering the impact of population change in one species on another over a certain time period. For accurate discussion group members must be talking about the same time periods in order for their parallel predictions to be compatible.

A third important feature which can influence a distributed mental model is the filtering of the information processed. For example, the hidden profiles identified by Stasser occurred because information relevant to the problem was not perceived as relevant whereas in fact it was. Individuals filtered the information presented to the group on the basis of their understanding of the problem and so inadvertently biased the outcome of the decision. This could also occur if a model is distributed amongst a group. The problem would be reduced if model is distributed between group members so that each holds whole modules as these allow relatively autonomous information processing. This means that there is less information to be communicated and so less chance that some of it will be inappropriately filtered out. But if the distribution is complex and much communication between group members is required, the chance that some of this information is filtered out becomes greater.

A final factor influencing how the distributed mental model may be used is the cognitive architecture onto which the model is mapped. A formal definition of a cognitive architecture suggests that it "includes the basic operations provided by the biological substrate, say, for storing and retrieving symbols, comparing them, treating them differently" (Pylyshyn, 1984, p. 30). In other words, a cognitive process takes place in a physical system, normally considered to be a person although here we have extended that unit to wider systems. This physical system provides some constraints and limits on what is possible. There are limits of speed, cognitive capacity, etc. which constrain the information processing in various ways. These limits will affect the distribution of the mental model.

In his approach to distributed cognition, Hutchins does not discuss the influence of the limits of these basic operations on the cognitive process. In fact, a result of his move towards any unit of analysis that is processing information is that he does not consider the differences between cognitive agents in a distributed cognitive system. They are all considered equally. In doing this Hutchins is adopting a classic information processing approach to cognition in which the process is outlined without consideration for the architectural constraints that might exist, such as biological plausibility. However, whilst it is useful to treat all information processing agents as part of the cognitive process, they all have different information processing capabilities and this will impact on performance. A group member who is able to write notes (and so distributes memory between herself and the paper) will be able to use this information to support recall, for example, and so perform differently to a group member without notes. Therefore it is not just the way the model is distributed that matters, it is also important to consider the different cognitive architectures onto which it is mapped. For example, an efficient distribution of the ecosystem in Figure 1 was to divide the model into two modules which were split between the two group members and so allowed the processing of these to occur in parallel. This was in part effective because the cognitive workload involved in forming predictions for each module was compatible with the capacity of the group members and the natural division of the model into two halves fitted the number of people engaged in this task. So here was a good mapping between the distribution of the model and the cognitive architecture (i.e., the group). If the same model had been mapped onto a group of three people there would not have been as efficient a distribution as the model cannot as easily be divided into three. Thus, the cognitive architecture of a group of three is not well aligned with the layout of this model. When considering the impact of the distribution of the mental model onto a distributed cognitive system such as a group it is not just the division of information that affects the information processing but rather the compatibility between the distribution and the cognitive architecture onto which it is mapped.

The concept of distributed mental models also provides a rich source of further research questions. For example, we have argued here that distributing a mental model can lead to efficient information processing by a group. But there are also likely to be times when sharing a mental model can be beneficial (e.g., Mathieu, Heffner, Goodwin, Salas, and Cannon–Bowers, 2000). In which situations are distributed or shared mental models likely to arise? We predict that more complex or specialised tasks lead to distributed mental models whereas time pressured tasks or times when there is little opportunity to communicate lead to distributed mental models. The type of knowledge within the mental model may also make a difference. We predict that mental models of tasks, which may be quite specialised, are more likely to be distributed than mental models of teams, which improve team functioning because they are shared. We discuss several features of distributed mental models and these raise empirical questions about their effect. For example, is the distribution of a mental model into relatively discrete modules always more efficient? How is the timing of coordination between group members managed? When do filtering effects influence the cognitive processing? What limits are placed on a distributed mental model by the underlying architecture, for example the cognitive processing capacity of the individuals in the group and what effect does this have on performance? Finally, this paper has considered the group purely as an information processing system whereas it is also a social group. How do social factors influence distributed cognition?

We believe that work to answer these questions will aid the understanding of cognition in groups, in particular high level cognition of groups and teams engaged in complex, real world tasks. The distributed mental model concept can help inform this work by offering a specific account of the cognitive representations used by groups in distributed situations.

Conclusion

This paper argues that distributed mental models are an important construct which adds to the explanation of much high level cognition in groups and teams. Many groups, especially in real world contexts, do not share knowledge. They may have distinct roles or expertise which means they cannot fully share an understanding of the task. Indeed, often the purpose of a group is to bring together different perspectives. As a result, previous approaches to understanding cognition in groups that focus exclusively on sharing cognition are inappropriate and a greater emphasis on distributing cognition allows a more accurate description of what groups actually do. However, it is necessary to go beyond the general claim that cognitive processes are distributed. A theory of what is distributed enables specific accounts of cognition in groups to be developed. This paper has focused on mental models as these are a theoretical construct which are a widely used to explain high level cognition such as reasoning and decision making. The central idea of a mental model as a mapping from the system in the world to the cognitive system is readily applied to distributed cognitive systems such as groups. This notion of a distributed mental model greatly adds to conceptualisations of groups as information processors through emphasising the importance of how a model is distributed, e.g., into relatively independent modules or the compatibility between the distribution of the model and the properties of the cognitive architecture of the distributed cognitive system. The approach can help develop accounts of the ways in which models are distributed between group members and therefore help explain how groups process information.

References

Ashby, W.R. (1956). *An introduction to cybernetics*. London: Chapman and Hall.

Banks, A.P., Macklin, C.M., and Millward, L.J. (2002, September). Distribution of causal reasoning in groups. *British Psychological Society Cognitive Section Annual Conference, Canterbury, United Kingdom*.

Banks, A.P., and Millward, L.J. (2000). Running shared mental model as a distributed cognitive process. *British Journal of Psychology, 91*, 513–531.

Chandrasekaran, B. (1981). Natural and social system metaphors for distributed problem solving: Introduction to the issue. *IEEE Transactions on Systems, Man, and Cybernetics, SMC–11*, 1–5.

Clancey, W.J. (1997). *Situated cognition: On human knowledge and computer representations*. Cambridge: Cambridge University Press.

Craik, K.J.W. (1943). *The nature of explanation*. Cambridge: Cambridge University Press.
Gentner, D., and Stevens, A.L. (1983). *Mental models*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
Greeno, J.G., and Moore, J.L. (1993). Situativity and symbols: Response to Vera and Simon. *Cognitive Science, 17*, 49–59.
Hinsz, V.B., Tindale, R.S., and Vollrath, D.A. (1997). The emerging conceptualization of groups as information processors. *Psychological Bulletin, 121*, 43–64.
Hutchins, E. (1995a). *Cognition in the wild*. Cambridge, Massachusetts: MIT Press.
Hutchins, E. (1995b). How a cockpit remembers its speeds. *Cognitive Science, 19*, 265–288.
Johnson–Laird, P.N. (1983). *Mental models: Towards a cognitive science of language, inference and consciousness*. Cambridge: Cambridge University Press.
Johnson–Laird, P.N. (2006). *How we reason*. Oxford: Oxford University Press.
Johnson–Laird, P. N., and Bara, B. (1984). Syllogistic inference. *Cognition, 16*, 1–61.
Kieras, D.E., and Bovair, S. (1984). The role of a mental model in learning to operate a device. *Cognitive Science, 8*, 225–273.
Klein, G.A. (1993). *Sources of power: How people make decisions*. Cambridge, Massachusetts: MIT Press.
Larson, J.R., and Christensen, C. (1993). Groups as problem-solving units: Toward a new meaning of social cognition. *British Journal of Social Psychology, 32*, 5–30.
Latane, B., Williams, K., and Harkins, S. (1979). Many hands make light work: The causes and consequences of social loafing. *Journal of Personality and Social Psychology, 37*, 822–832.
Laughlin, P.R., and Ellis, A.L. (1986). Demonstrability and social combination processes on mathematical intellective tasks. *Journal of Experimental Social Psychology, 22*, 177–189.
Mathieu, J.E., Heffner, T.S., Goodwin, G.F., Salas, E., and Cannon–Bowers, J.A. (2000). The influence of shared mental models on team process and performance. *Journal of Applied of Psychology, 85*, 273–283.
Moray, N. (1997). Models of models of . . . mental models. In T.B. Sheridan and T. Van Luntern (Eds.), *Perspectives on the human controller* (pp. 271–285). Mahwah, New Jersey: Lawrence Erlbaum Associates.
Moray, N. (1999). Mental models in theory and practice. In. D. Gopher and A. Koriat (Eds.), *Attention and performance XVII: Cognitive regulation of performance: Interaction of theory and application* (pp. 223–258). Cambridge, Massachusetts: MIT Press.
Norman, D.A. (1988). *The design of everyday things*. New York: Doubleday.
Pylyshyn, Z.W. (1984). *Computation and cognition: Towards a foundation for cognitive science*. Cambridge, Massachusetts: MIT Press.
Stasser, G., and Stewart, D. (1992). Discovery of hidden profiles by decision-making groups: Solving a problem versus making a judgement. *Journal of Personality and Social Psychology, 63*, 426–434.
Stasser, G., and Titus, W. (1985). Pooling of unshared information in group decision making: Biased information sampling during discussion. *Journal of Personality and Social Psychology, 48*, 1467–1478.
Suchman, L.A. (1987). *Plans and situated actions: The problem of human-machine communication*. Cambridge: Cambridge University Press.
Tindale, R.S., and Kameda, T. (2000). 'Social sharedness' as a unifying theme for information processing in groups. *Group Processes and Intergroup Relations, 3*, 123–140.
Wegner, D.M. (1987). Transactive memory: A contemporary analysis of the group mind. In B. Mullen and G.R. Goethals (Eds.), *Theories of group behavior* (pp. 185–208). New York: Springer Verlag.
Wilson, J.R., and Rutherford, A. (1989). Mental models: Theory and application in human factors. *Human Factors, 31*, 617–634.
Zajonc, R.B. (1965, July). Social facilitation. *Science, 149*, 269–274.