Group Minds and Explanatory Simplicity

MARK SPREVAK AND DAVID STATHAM

Abstract

This paper explores the claim that explanation of a group's behaviour in term of individual mental states is, in principle, superior to explanation of that behaviour in terms of group mental states. We focus on the supposition that individual-level explanation is superior because it is simpler than group-level explanation. In this paper, we consider three different simplicity metrics. We argue that on none of those metrics does individual-level explanation achieve greater simplicity than a group-level alternative. We conclude that an argument against group minds should not lay weight on concerns of explanatory simplicity.

Introduction

In three papers, Robert Rupert presents arguments against group minds.[1] These criticisms motivate the conclusion that although the discovery of group minds is an open empirical possibility, there are strong reasons for thinking that no group minds exist. One of Rupert's principal arguments against group minds is an argument from *explanatory simplicity*. His claim is that, for any explanation of intelligent behaviour that appeals to group minds, there exists an alternative explanation couched solely in terms of the minds of individuals. According to Rupert, an *individual-level explanation*, which makes no reference to group minds will be simpler than its *group-level* alternative, which does make reference to group minds. Simpler explanations are, everything else being equal, better explanations. Therefore, individual-level explanations are, everything else being equal, preferable to their more complex group-level alternatives.

[1] R. D. Rupert, 'Minding One's Cognitive Systems: When Does a Group of Minds Constitute a Single Cognitive Unit', *Episteme* **1** (2005), 177–88; R. D. Rupert, 'Empirical Arguments for Group Minds: A Critical Appraisal', *Philosophy Compass* **6** (2011): 630–39; R. D. Rupert, 'Against Group Cognitive States' in *From Individual to Collective Intentionality: New Essays*, edited by S. Chant, F. Hindriks, and G. Preyer (Oxford University Press, 2014).

doi:10.1017/S1358246115000132 ©The Royal Institute of Philosophy and the contributors 2015

https://doi.org/10.1017/S1358246115000132 Published online by Cambridge University Press

In his invocation of the principle of explanatory simplicity, Rupert overlooks a great deal of this notion's complexity. We aim to explicate this notion and its various dimensions and, in doing so, to assess the relevance of each dimension to the issue of group-level explanation. We conclude that, whichever way explanatory simplicity is analysed, there is no warrant for thinking that individual-level explanations are simpler than group-level alternatives. Hence, this particular route to argue against the existence of group minds is blocked.

A Tale of Two Types of Agent

Rupert claims that no matter the intelligent behaviour a group displays, an explanatory story involving only cognitive properties of *individuals* is always available:

> In all of the cases that involve the relevant explananda, there seem to be available complete causal explanations couched in terms of the cognitive states of individuals (together with causal contributions of noncognitive, physical structures).[2]

Rupert claims that these individual-level explanations are preferable to group-level explanations. According to Rupert, this is because individual-level explanations are more simple. This consideration is the opening move in Rupert's inference to the best explanation against the existence of group minds.

Let us see an example of Rupert's argument in action.

An impressive display of intelligent behaviour by a group occurs when Microsoft develops a new operating system (OS). How does a group (Microsoft) produce this behaviour? Two types of story could be told. One is formulated exclusively at the level of individual agents: it refers to the cognitive properties of individual employees, to the interactions between employees, and to the organisational structure of the corporation that governs these interactions. The other introduces group agents as protagonists: it ascribes cognitive properties (such as beliefs, desires, and intentions), not only to individual employees, but also to groups of individuals, such as teams within Microsoft or the corporation as a whole.

What do these two stories look like?

An individual-level explanation of Microsoft's production of a new OS might begin by referring to the cognitive states of the individuals comprising the corporation's executive: their *beliefs* that a new OS is

[2] Rupert, 'Empirical Arguments for Group Minds', 635.

https://doi.org/10.1017/S1358246115000132 Published online by Cambridge University Press

required; their *desires* to produce this new OS within the next three years; and their *intentions* to undertake the necessary activity to ensure that these desires are fulfilled. The goal of producing a new OS can be broken down into several sub-goals: the production of its graphical content, its software compatibility, its compatibility with a range of hardware, and so on. Groups of programmers would be tasked with producing the code required to achieve each of these sub-goals and each programmer would adjust their cognitive states accordingly.

Each programmer would be ascribed a set of beliefs about the activities they would need to perform to achieve the sub-goal with which they have been tasked: the belief that producing a satisfactory user interface will require two years of work; the belief that a particular programming language would be most appropriate to this task; and the belief that weekly consultations with the corporation's executive will be necessary, and so on. Each programmer would have a set of relevant desires, intentions, and other cognitive states: she might, for example, have the desire to achieve her sub-goal within a particular time frame, to do so to the best of her ability, to exploit her newly gained competence with a particular programming language, and so on.

That an individual employee can be attributed cognitive states requires that the individual instantiate a *cognitive architecture*. This is to say that the individual is structured in such a way as to comprise a set of mechanisms whose tightly integrated activity contributes causally to that individual's intelligent behaviour. The *cognitive states* attributed to the individual are realised by the integrated activity of the mechanisms which comprise that individual's cognitive architecture. These states enter into causal series in which each state is caused by, and subsequently causes or modifies, other cognitive states. These causally connected series of cognitive states constitute *cognitive processes*. And it is in virtue of the realisation of these cognitive processes that the individual comes to behave intelligently and produce a wide range of cognitive phenomena (from choosing a suitable route to work in the morning to successfully coding a graphical user interface).

Thus, to treat an individual employee as genuinely *cognitive* is to attribute to that individual a related set of *cognitive properties*. The individual instantiates a cognitive architecture: a set of integrated cognitive mechanisms (such as working memory, a language processor, and visual object-recognition mechanisms). The integrated activity of these mechanisms realises cognitive states (beliefs, desires, and memories; representations of words, tasks, and problems; and so on). These states are connected in causal series to

https://doi.org/10.1017/S1358246115000132 Published online by Cambridge University Press

constitute cognitive processes (decision-making, problem-solving, and so on). And in virtue of the individual's instantiation of these various cognitive properties, that individual is able to produce a variety of cognitive phenomena.

Each individual employee is attributed in the individual-level explanation a set of related cognitive properties. Crucially, of course, it is only *individual* employees who are attributed cognitive properties, not any extra-individual entity or collection of entities.

An individual-level explanation of Microsoft's behaviour would of course need to appeal to more than this. It would need to refer to the various processes by which individual employees communicate the content of their cognitive states, and the results of their cognitive processes, to other employees. These communicative processes would include face-to-face linguistic interactions, the production of written memos, email, video-conferencing, record-keeping, and so on. An individual-level explanation would also need to refer to the non-cognitive physical structures used to facilitate individual cognitive processes and inter-individual communicative processes. A corporation like Microsoft makes use of all sorts of digital computers, hard drives and other storage devices, telecommunications devices, and old-fashioned pen and paper, amongst other things. Finally, of crucial importance would be a description of the corporation's organisational structure and the procedures and rules which determine how the activity of its numerous individual employees is coordinated and channelled towards the production of a new OS. This would include reference to sets of implicit and explicit rules, the hierarchical organisation of employees, the corporation's departmental structure, the procedural standards for communication between departments, and so on.

An individual-level causal explanation of Microsoft's production of a new OS would clearly be a complex and multifaceted explanation that would make reference to an enormous number of individual cognitive properties, interactive and communicative processes, rules and organisational structures, and various non-cognitive structures and processes.

What about the group-level story? A group-level explanation would ascribe cognitive properties, not only to individual employees, but also to groups of individuals. The relevant groups may be of various sizes and their membership could overlap. The groups may include small teams of programmers, larger engineering teams, entire software divisions, or the corporation as a whole. Just as individual employees are attributed a set of related cognitive properties, so a group would be treated as a cognitive system that instantiates a cognitive architecture and realises cognitive states that combine in causal series to constitute cognitive processes.

https://doi.org/10.1017/S1358246115000132 Published online by Cambridge University Press

For example, Microsoft's organisational structure, which plays the crucial role of integrating and coordinating the activity of individual employees in the individual-level explanation, may be equated with a cognitive architecture. This architecture is comprised of a large set of mechanisms whose activation may realise cognitive states. But whereas the cognitive mechanisms comprising an individual employee's cognitive architecture consist of various neural systems in the employee's brain, the cognitive mechanisms that comprise Microsoft's cognitive architecture consist of groups of individual employees and their supporting physical apparatus. The group-level cognitive states are realised by individual employees in interaction with each other and with their supporting physical apparatus. The group-level cognitive processes are the causal series in which each of these cognitive states is caused by, and causes or modifies, the others.

This presentation of the group-level explanation suggests that it is a rival to the individual-level explanation in the sense that it might eliminate the need for individual-level talk: individual-level cognitive properties, inter-individual interaction, and non-cognitive physical structures would be discarded in favour of group cognitive properties. But it need not be so. We will suppose in what follows that the group-level explanation retains reference to the various individual-level cognitive properties and non-cognitive physical structures described in the individual-level explanation.

It is conceivable that the similarity between a particular group and a cognitive agent is too superficial, or in other ways insufficient, to justify attribution of cognitive properties to that group. Both advocates and critics of group cognition agree that some level of functional similarity – the correct level of which is disputed – between a group and an individual cognitive agent is required in order for the attribution of mental properties to a group to be warranted at all. Nevertheless, in order to facilitate a discussion of explanatory simplicity, these considerations will be temporarily set aside. The claim at issue is not that Microsoft's production of a new OS genuinely constitutes a real-world example of group cognition. The claim is that, *if* a case such as the Microsoft example did warrant both an individual-level and a group-level explanation, and *if* two such complete causal explanations *could indeed be given*, the individual-level explanation would be superior in terms of explanatory simplicity.

Granted our task is to assess explanatory simplicity (and not some other dimension of adequacy of group-level explanation), it is important to tally up, as Rupert does, the nature and quantity of additional postulates made by group-level explanation. The additional postulates of group-level explanation consist of an additional set of

https://doi.org/10.1017/S1358246115000132 Published online by Cambridge University Press

cognitive properties: the instantiation, in virtue of the corporation's organisational structure, of a cognitive architecture; the realisation, in virtue of this architecture, of cognitive states over and above those realised by individual employees; and the realisation via the causal series of these cognitive states, of cognitive processes over and above those realised by individual employees.

Assuming, as per the above concession, that the group-level explanation retains reference to all individual-level cognitive properties and non-cognitive physical structures, the two explanations appear to sit roughly in a relation of containment. The group-level explanation posits all that the individual-level explanation posits, plus an additional set of properties, the group-level cognitive properties. Therefore it seems reasonable to conclude, as Rupert does, that individual-level explanation must be the simpler hypothesis. If one can explain the group's behaviour without introducing additional cognitive properties possessed by groups, why, everything else being equal, introduce those group-level cognitive properties?

The aim of this paper is to show that this conclusion is unwarranted. Individual-level explanation is not simpler than, and hence (all else being equal) preferable to, group-level explanation. Our strategy in this paper is to examine how simplicity should be measured. We consider three approaches: qualitative parsimony, qualitative parsimony, and theoretical elegance. We argue that on none of these measures does individual-level explanation count as more simple than group-level explanation. We conclude that concerns about simplicity cannot be straightforwardly used, as Rupert wants, to argue against the existence of group minds.

Qualitative Parsimony

One dimension of simplicity is parsimony. Parsimony concerns the number of entities posited by an explanation. How does one go about counting these entities? There seem to be two possibilities: count the *kinds* of entity postulated or count the *number of entities* postulated and subsumed under each kind. Counting the number of kinds gives a measure of *qualitative* parsimony; counting the number of entities subsumed under each kind gives a measure of *quantitative* parsimony.[3] Qualitative parsimony is the focus of this section; quantitative parsimony is the focus of the next.

[3] D. K. Lewis, *Counterfactuals* (Oxford: Blackwell, 1973).

https://doi.org/10.1017/S1358246115000132 Published online by Cambridge University Press

How do considerations of qualitative parsimony apply in the Microsoft case? On the assumption that both the individual-level and group-level explanations are otherwise adequate, the individual-level explanation might appear more qualitatively parsimonious for the simple fact that it includes only those kinds that are already in use throughout standard cognitive science. In other words, the individual-level explanation helps itself only to the cognitive kinds of architecture, state, and process that are already attributed to individual humans in cognitive scientific explanations. In contrast, the group-level explanation of the Microsoft's production of a new OS makes additional postulates: it attributes cognitive properties to Microsoft at the group-level.

It is of crucial importance, however, to the group-level explanation that the additional postulates involve attribution of *the same kind* of properties to Microsoft as are attributed to its individual employees. The architecture that the group-level explanation takes Microsoft to instantiate is, *ex hypothesi*, of the same cognitive kind as the architecture instantiated by individuals. And the states and processes that Microsoft is attributed in virtue of its instantiating this cognitive architecture are again, *ex hypothesi*, cognitive states and processes of the same kind as are attributed to its individual humans.

The cognitive states attributed to Microsoft are *additional instances* of individual-level cognitive states: beliefs, desires, intentions, and so on. The group-level states are individually distinct from any of the individual-level cognitive states attributed to particular employees. But the group-level states are individually distinct in the sense that they constitute new instances of the same cognitive kind as are attributed to individual employees. The claim – the innovation – that makes the group-level explanation of Microsoft's production of a new OS interesting is that Microsoft instantiates additional instances of the same kind of cognitive properties as are attributed to individual humans. The same is true of any group-level cognitive explanation. These explanations are interesting precisely because of the similarity they posit between individual-level properties and group-level properties. A group-level explanation that attributed utterly alien cognitive properties – so different as to fail to fall under any same-kind relation to those of individual humans – would not count as a group *mind* hypothesis at all.

Understood correctly, group-level explanations do not postulate additional cognitive kinds, only additional instances of existing kinds. Therefore, an appeal to qualitative parsimony as a strategy to argue against group-level explanations would miss the mark. There are no extraneous kinds to 'shave off' by dispensing with group-level explanation, only attribution of existing kinds to new particulars.

https://doi.org/10.1017/S1358246115000132 Published online by Cambridge University Press

Quantitative Parsimony

Might group-level explanations be less simple than individual-level alternatives because group-level explanations postulate extra instances of cognitive kinds? Does a group-level explanation of, say, Microsoft's behaviour introduce unnecessary instances of cognitive architecture, states and processes? In this section, we argue that concerns about quantitative parsimony do not tell against group-level explanation.[4]

Nolan provides an instructive case study on the value of quantitative parsimony.[5] Nolan's study concerns the postulation of the neutrino in the explanation of beta decay. In attempting to explain a puzzling drop in the energy of nuclei during beta decay, Wolfgang Pauli postulated the existence of a new kind of subatomic particle, the neutrino. When developing the neutrino-based explanation, there was no available means of determining how many of the new subatomic particles were released during beta decay. The total mass-energy to be accounted was known, but whether this mass-energy was best accounted by one, two, or a million neutrinos was an open question. Once the release of a single neutrino has been postulated, why not postulate the release of two neutrinos each with half the mass-energy of one neutrino? Or a million neutrinos with a millionth of the mass-energy? At the time, the one-neutrino and the million-neutrino explanations could both have accommodated the empirical data. Nolan argues that it was nevertheless rational, in light of consideration of quantitative parsimony, for Pauli to favour the one-neutrino explanation. Nolan defends this in two ways.

First, one may make a brute appeal to intuition. That the one-neutrino explanation is superior to the million-neutrino explanation (or any other such explanation) is an intuition that many otherwise rational enquirers share. It may seem reasonable to endorse this, at least pending countervailing evidence, even if it does not admit of explicit justification. Second, one may appeal to the successful track

[4] Appeal to quantitative parsimony is far from uncontroversial as a means of selecting between rival explanations. Lewis, *Counterfactuals*, says, 'I subscribe to the general view that qualitative parsimony is good in a philosophical or empirical hypothesis; but I recognise no presumption whatever in favour of quantitative parsimony' (87). Alex Oliver argues that, despite our undeniable bias towards explanations with quantitative parsimony, we should not trust this as a guide to the truth ('The Metaphysics of Properties', *Mind* **105** (1996), 1–80).

[5] D. Nolan, 'Quantitive Parsimony', *The British Journal for the Philosophy of Science* **48** (1997), 329–343.

https://doi.org/10.1017/S1358246115000132 Published online by Cambridge University Press

record of quantitatively parsimonious explanations. Quantitative parsimony has led to the selection of explanations that have gone on to prove their worth in other ways. A case in point is the one-neutrino explanation. Pauli's explanation was subsequently vindicated. Even if the reasons for the success of quantitative parsimony remain obscure, its success alone appears to lend some justification to the principle's continued application.

The cases in which quantitative parsimony play a part might seem trivial when the number of entities postulated (for example, atoms or subatomic particles) is compared with the potentially infinite number of other entities in the universe. This seems to pose a problem if one endorses an 'infinite ontology' view of the universe. One could argue that against this infinite background, a numerical increase in entities of one particular kind is trivial or negligible. Why should it matter that an explanation adds a few more entities to our scientific ontology when this ontology already includes the infinity of other kinds of entities which occupy an infinite universe?

Against this thought, Nolan contends that, irrespective of the potential infinity of entities subsumed under all extant kinds, it is nevertheless important to minimise the number of entities subsumed by any one particular kind. Thus, for example, even against the background of infinite mathematical entities to which the mathematical Platonist is committed, the addition of 'seventeen million little neutral particles being produced in every case of Beta-decay', as opposed to the addition of just one such particle, still makes a significant difference.[6]

We will follow Nolan in his understanding quantitative parsimony. Recall that the group-level explanation claims that Microsoft, as a group, instantiates a cognitive architecture with instances of cognitive states over and above those of its individual employees. Is an individual-level explanation more quantitatively parsimonious than this group-level account (assuming, as above, that both explanations can be given)?

At first glance, the answer appears to be a clear 'yes'. At the end of the previous section we said that the group-level explanation posits extra *instances* of cognitive architectures, states, and processes, although it does not posit new kinds. Granted this, the group-level explanation appears to do worse in terms of quantitative parsimony: it posits more instances of cognitive properties than individual-level explanation. On closer examination, however, the case against group-level explanation is not so clear.

[6] Nolan, 'Quantitative Parsimony', 340.

https://doi.org/10.1017/S1358246115000132 Published online by Cambridge University Press

An initial objection that might be levelled by an advocate of individual-level explanation is that the group-level explanation attributes to the Microsoft corporation a *cognitive architecture*, whilst at the same time appealing to the cognitive architectures of the individual employees. If this attribution involves the postulation of an additional entity, then it seems to fall foul of considerations of quantitative parsimony when compared to individual-level alternatives.

Understood correctly, however, the group-level explanation does not postulate the instantiation of a cognitive architecture as an *additional* entity over and above those postulated in the individual-level explanation. Rather, it *identifies* the organisational structure of the corporation, as described in the individual-level explanation, with a genuinely cognitive architecture. The individual-level explanation must already include this organisational structure in order to explain the interactions that take place between the corporation's individual employees. It is committed to there being an executive, external information stores, a variety of sub-components in the form of teams of individual employees, and various other non-cognitive physical structures. In order to make sense of the interactions between these features and to explain how they are coordinated in the production of a new OS, the individual-level explanation cannot merely postulate their existence. It must also give an account of the rules and procedures governing their interaction.

In attributing to the corporation a cognitive architecture, the group-level explanation does not postulate an extra structure over and above the structure already described at the individual-level. It claims that, properly understood, this structure *is* an instance of a cognitive architecture.

What about the additional cognitive states and processes? It seems that they must violate the principle of quantitative parsimony. Even if Microsoft as a corporation instantiates a cognitive architecture, an advocate of quantitative parsimony could say that the cognitive states attributed to the corporation in the group-level explanation constitute extraneous additions to number of instances of cognitive states postulated in the individual-level explanation. The individual-level explanation makes reference to the cognitive states of individual employees. The group-level explanation makes reference to these cognitive states plus an extra set of cognitive states at group level. The group-level explanation therefore seems inferior in terms of quantitative parsimony.

Let us pause and consider this more carefully. First, we need to ask how the instances of cognitive states postulated by each explanation are counted. Consider the cognitive state of *intending to develop a*

https://doi.org/10.1017/S1358246115000132 Published online by Cambridge University Press

new OS. Let us suppose that the individual-level explanation attributes this state to the corporation's CEO, amongst other individual employees. The group-level explanation retains these attributions and, additionally, attributes the state of *intending to develop a new OS* to the corporation as a whole.

Here is one method by which we could go about counting cognitive states. Beginning with the CEO, we say that his state of intending to develop a new OS is one cognitive state to which the individual-level explanation is committed. On top of this, let us suppose that he simultaneously *believes* that the development of a new OS is a financial necessity. This gives us two cognitive states. Simultaneously, he *desires* to make as much money as possible for the corporation during his time as CEO, giving us a total of three cognitive states. Assuming, for the sake of brevity, that the CEO's cognitive states are limited to only these three, we could go on to count the cognitive states of all the other individual employees in a similar manner. At the end of this process, we add up the cognitive states that have been attributed to each individual employee, and this gives us the total number of cognitive states to which the individual-level explanation is committed.

Suppose, for the sake of argument, that the group-level explanation is committed to those individual-level cognitive states too. But it goes further and attributes cognitive states at the group-level. We count these group-level cognitive states, add them to the total number of individual-level cognitive states, and this gives us the number of cognitive states to which the group-level explanation is committed. Patently, using this method of counting cognitive states, the group-level explanation is committed to a greater number of cognitive states than the individual-level explanation.

This method of counting cognitive states gives us the total number of *occurrent* cognitive states attributed by the explanations. The method might seem appealing. However, it relies on an implausibly atomistic view of the nature of cognitive states.

Cognitive states ought not to be treated as entities in the same way as material objects whose existence can be treated independently of the existence of all other such objects. If we are counting the number of chairs in a room, then counting only the 'occurrent' chairs – those that are present in the room at one particular point in time – is unproblematic. The existence of each of these chairs, and the location of each chair, stands largely independent of the existence or location of any other chair. The existence of any single chair, and the location of the chair, entails almost nothing about the existence or location of any other chair. Cognitive states are different, or so the

https://doi.org/10.1017/S1358246115000132 Published online by Cambridge University Press

objection goes. The fact that a cognitive architecture instantiates a particular cognitive state at a given moment in time *does* entail further facts about other cognitive states. The existence of a particular cognitive state cannot be treated in isolation from other cognitive states.

This approach to cognitive states falls under the broad heading of *dispositionalism*. The core of dispositionalist theories – the idea that is of the greatest significance for the discussion of group-level cognitive states – can be illustrated by some simple examples. We will give these to demonstrate the strength of the view as a broadly correct characterisation of cognitive states, and superior to the occurrent characterisation above.

Consider that *I believe that my house contains six rooms*. At a given point in time I am laying down plans for my latest DIY project and this has given me cause to think about the total number of rooms in my house. I say to myself 'there are six rooms in my house'. It seems entirely reasonable that at this given point in time, we should attribute to me the occurrent belief *that my house has six rooms*. It seems uncontroversial that I have at least this one belief at this point in time.

What about the belief *that my house contains more than three rooms*? I might never have cause during the course of my DIY project to entertain this as a proposition, and it might plausibly be the case that at no point in my life does the belief *that my house contains more than three rooms* ever become occurrent. Does this mean that I do not believe that my house contains more than three rooms? Maintaining that I do not seems to neglect a crucial feature of beliefs. I am *disposed* to assent to statements or answer questions in accordance with the non-occurrent belief.

It is the important sense in which I believe *that my house contains more than three rooms*, regardless of the actual occurrence of this belief, which is captured by the dispositionalist claim we are interested in here. The claim is that, in virtue of the structure possessed by my cognitive architecture, I am disposed to respond in a particular way to a whole set of possible stimuli and other dispositional mental states – to assent to a whole set of possible statements, to answer questions in a particular way, and so on.

Returning to Microsoft example, the CEO was attributed three cognitive states, one of which was the intention to develop a new OS. Construed in dispositional terms, however, his single state of intending entails many others. If the CEO intends to develop a new OS, he presumably also intends to develop a new OS which can be made available for the use of Microsoft's customers. He intends that this

https://doi.org/10.1017/S1358246115000132 Published online by Cambridge University Press

OS will run successfully on existing hardware, that it will enable the use of more than one piece of software, and so on. And these entailments are not limited to intentions. If the CEO intends to develop a new OS he presumably also *believes* that the development of a new OS is possible, and he believes that it is within the capability of the corporation to develop a new OS. Presumably he also has beliefs about what an OS is, what a corporation is, and his position within the corporation such that he can intend to develop a new OS in the first place, and so on.

To say that the CEO instantiates the cognitive state of intending to develop a new OS but none of the other inferentially-related cognitive states seems to neglect an essential feature of what it is to instantiate a cognitive state. We want to suggest, then, that cognitive states are not best characterised as solely occurrent states. Rather, in virtue of having a cognitive architecture and instantiating cognitive states, the CEO possesses a whole set of related dispositional mental states.

If we treat cognitive states in dispositional terms, how do we count them?

There are presumably limits in principle to the number of occurrent states that a single cognitive architecture can instantiate at any given point in time, even if those limits are difficult to discern. Limited neural resources and limited number of cognitive mechanisms that comprise the cognitive architecture of an individual of finite size mean that the number of occurrent cognitive states instantiated at any given moment is finite.

In contrast, the number of dispositions that an individual has, in virtue of their instantiating a cognitive architecture is, in principle, unlimited. If I believe *that my house contains six rooms*, and I also believe *that my house contains more than three rooms*, presumably I also believe *that my house contains fewer than seven rooms*. I am disposed to assent to the statement that my house contains fewer than seven rooms. Surely, then, I also believe *that my house contains fewer than eight rooms*, and fewer than nine rooms, and ten rooms, and so on *ad infinitum*. Construed dispositionally, there is no point in this sequence at which it makes sense to terminate. There is no principled limit on the number of rooms to which I would be disposed to respond *that my house contains fewer*.

In the Microsoft case, there is likewise no principled limit to the number of cognitive states we should attribute to its CEO once it has been accepted that he intends to develop a new OS. The simplest example of this fact could be given just by considering that, if the CEO intends to develop *one* new OS, he also intends to develop *less than two* new OSs, and *less than three* new OS, and so on. It is not

https://doi.org/10.1017/S1358246115000132 Published online by Cambridge University Press

just that the limits to the number of these cognitive states are difficult to discern. Rather, there are no such limits. There seems, then, to be the very real possibility that, construed in dispositional terms, the number of instances of cognitive states falling under each kind in individual-level explanation of Microsoft's behaviour is *unlimited*.

If individual-level explanation is committed to an unlimited number of instances of cognitive states (beliefs, intentions, desires, etc.), then the addition of further instances of these kinds attributed to groups would offer no overall increase in number. Even if the number of group-level cognitive instances were similarly unlimited, there would still be no overall increase. This is bad news for anyone wishing to use quantitative parsimony to argue in favour of individual-level explanation. Our first intuition that the number of individual cognitive states plus the number of group cognitive states must be greater than the number of individual cognitive states breaks down once we introduce an unlimited number of dispositional cognitive states.

Despite its initial intuitive plausibility, quantitative parsimony does not lend credence to individual-level explanation being more simple than group-level explanation. Insofar as cognitive property instances can be counted at all, adding further instances via group-level explanation does not increase our total commitments.

Theoretical elegance

The final measure of explanatory simplicity we consider is *theoretical elegance*. Theoretical elegance concerns the structure and length of the explanation, rather than the entities or kinds to which it is committed.[7]

Two ways of understanding theoretical elegance are considered below: explanations that minimise the number of primitive predicates employed, and explanations that minimise the length of the explanation. We argue that neither provides the grounds for thinking that individual-level explanations are more simple than group-level alternatives.

Explanations employ primitive and non-primitive predicates. Both kinds of predicate are used to express the ideas that a theory contains. The primitive predicates are those which are undefined: they receive

[7] A. Baker, 'Simplicity' in *The Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta, Fall 2013, http://plato.stanford.edu/archives/fall2013/entries/simplicity/.

https://doi.org/10.1017/S1358246115000132 Published online by Cambridge University Press

no definition or analysis within the explanation. The non-primitive predicates are defined: they receive a definition in terms of the predicates included elsewhere within the explanation.

In the Microsoft example, the primitive predicates might include *has the belief that a new OS is required*, *has the desire to produce the new OS within the next three years*, and so on. Additionally, many predicates related to the non-cognitive properties of physical structures and the predicates needed to explain the organisation of the corporation are likely to be primitive.

According to our assumption in Section 2, the group-level explanation will include all the primitive predicates of the individual-level explanation. The additional postulates in the group-level explanation are the group-level cognitive properties, and therefore the only additional predicates in the group-level explanation could be those predicates relating specifically to the instantiation of cognitive properties at the group-level: those needed to explain why and how a group-level architecture is a *group-level* architecture; why and how the additional states being postulated are *group-level* states, and so on.

These group-level predicates are the subject matter of group-level explanations and will therefore need to be explained and defined within the group-level explanation. These group-level predicates will *not* remain undefined because it is the concern of such an explanation to explain why and how cognitive properties are instantiated at the group level, and what it means to have the group-level properties in the sense intended by the explanation. Therefore, the (additional) group-level predicates will not be primitive.

This means that the group-level explanation will not include any more primitive predicates than the individual-level explanation. So, with regard to elegance *qua* minimum number of primitive predicates, group-level explanations are no worse off than individual-level explanations.

If the Microsoft group-level explanation replaces reference to individual-level cognitive properties and the various non-cognitive physical structures mentioned in the individual-level explanation with reference to only a single cognitive system and its properties, group-level explanation could be shorter. Even if group-level explanation retains reference to all the properties and structures of individual-level explanation, there will nevertheless be aspects of the group's activity whose lengthy individual-level explanations can be replaced with shorter group-level explanations.

For example, consider retrieval of information from a non-cognitive information store, such as a hard drive. For an individual-level explanation, this would need to be explained with description of an

https://doi.org/10.1017/S1358246115000132 Published online by Cambridge University Press

individual employee's processes of interaction with the hard drive and how information on the hard drive influences that individual's cognitive states and processes. This would involve, for example, facts about the physical encoding of the information on the hard drive, the content of the information, the employee's sensory interaction with the hard drive, their conceptualisation of the information, and their endorsement of the information. The explanation may plausibly draw on other elements too, such as translation between natural languages in order for the employee to capable of conceptualising the information in the first place. The explanation also needs to explain how, once the information has affected the cognitive states of an individual, this then affects the behaviour of the corporation as a whole.

In a group-level explanation, this set of processes by which the information contained on a hard drive comes to influence an individual's cognitive states and processes, and by which these states and processes are communicated to and influence other individual employees, could be given a much shorter explanation in terms of *the corporation*'s group-level cognitive states being influenced by states of *the corporation*'s long-term memory. Thus, even if group-level explanation elsewhere retains reference to the same individual-level cognitive properties and non-cognitive structures as individual-level explanation, there will be cases where group-level explanation can compress non-trivial sets of statements made by the individual-level explanation of certain processes.

This mirrors a discussion by Clark and Chalmers of the appeal of the extended mind hypothesis.[8] Clark and Chalmers argue that when explaining Otto's behaviour, certain stages in the explanation become unnecessary once one switches to an explanation couched in terms of a single extended cognitive system.[9] The description of Otto's transit to The Museum of Modern Art need not, they argue, make reference to Otto's belief that The Museum of Modern Art's location is recorded in his notebook, or to the fine-grained details of Otto's sensorimotor interaction with his notebook. In the extended-mind version of the explanation, this whole process is glossed simply as the retrieval of information from Otto's long-term extended memory in the service of fulfilling his desires. The non-extended alternative involving a detour through Otto's beliefs about his notebook takes 'one step too many' and is 'pointlessly complex'.

[8] A. Clark and D. J. Chalmers, 'The Extended Mind', *Analysis* **58** (1998), 7–19

[9] Clark and Chalmers, 'The Extended Mind', 13.

https://doi.org/10.1017/S1358246115000132 Published online by Cambridge University Press

Likewise, if a group-level explanation of Microsoft's production of a new OS is available, many stages of the individual-level explanation seem pointlessly complex. Ensembles of processes can be glossed as the execution of a single group-level cognitive process, such as retrieval of information from long-term memory. The set of statements describing this will be significantly shorter than the statements required to explain the same process in purely individual-level terms. As a result, even if the group-level explanation retains reference to individual-level cognitive properties and non-cognitive structures, there will be opportunities to substantially reduce the number of statements used in the explanation and there is a possibility that the group-level explanation will be shorter and more elegant.[10]

Conclusion

Individual-level explanations of intelligent behaviour are not always simpler than group-level explanations. This is not to say that particular individual-level explanations are not superior to group-level explanations. Nor is it to rule out an argument that individual-level explanation is better based on another, non-simplicity-based, metric. Nor is it to rule out a general non-explanatory-value-based argument against group minds. Our aim has only been to show that appeal to explanatory simplicity is not enough to show that group-level explanation is inferior to individual-level explanation.

University of Edinburgh
Mark Sprevak (mark.sprevak@ed.ac.uk)
David Statham (d.statham@ed.ac.uk)

[10] The authors are listed here in alphabetical order; both authors have contributed equally to this work. We would like to thank Anthony O'Hear for inviting us to present this paper at the Royal Institute of Philosophy, and the audience for their useful questions and suggestions. We would like to thank Robert Rupert for helpful comments on an earlier draft of this paper.

https://doi.org/10.1017/S1358246115000132 Published online by Cambridge University Press