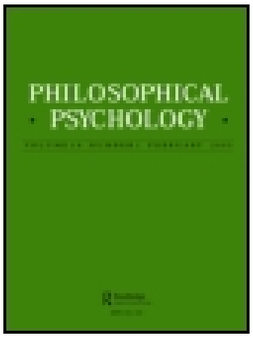

Philosophical Psychology

ISSN: 0951-5089 (Print) 1465-394X (Online) Journal homepage: http://www.tandfonline.com/loi/cphp20

The multiple, interacting levels of cognitive systems (MILCS) perspective on group cognition

Robert L. Goldstone & Georg Theiner

To cite this article: Robert L. Goldstone & Georg Theiner (2017): The multiple, interacting levels of cognitive systems (MILCS) perspective on group cognition, Philosophical Psychology, DOI: 10.1080/09515089.2017.1295635

To link to this article: http://dx.doi.org/10.1080/09515089.2017.1295635

Published online: 07 Apr 2017.

PHILOSOPHICAL PSYCHOLOGY, 2017
http://dx.doi.org/10.1080/09515089.2017.1295635

Routledge
Taylor & Francis Group

The multiple, interacting levels of cognitive systems (MILCS) perspective on group cognition

Robert L. Goldstone[a] and Georg Theiner[b]

[a]Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN, USA; [b]Department of Philosophy, Villanova University, Villanova, PA, USA

ABSTRACT
We lay out a multiple, interacting levels of cognitive systems (MILCS) framework to account for the cognitive capacities of individuals and the groups to which they belong. The goal of MILCS is to explain the kinds of cognitive processes typically studied by cognitive scientists, such as perception, attention, memory, categorization, decision-making, problem solving, judgment, and flexible behavior. Two such systems are considered in some detail—lateral inhibition within a network for selecting the most attractive option from a candidate set and a diffusion process for accumulating evidence to reach a rapid and accurate decision. These system descriptions are aptly applied at multiple levels, including within and across people. These systems provide accounts that unify cognitive processes across multiple levels, can be expressed in a common vocabulary provided by network science, are inductively powerful yet appropriately constrained, and are applicable to a large number of superficially diverse cognitive systems. Given group identification processes, cognitively resourceful people will frequently form groups that effectively employ cognitive systems at higher levels than the individual. The impressive cognitive capacities of individual people do not eliminate the need to talk about group cognition. Instead, smart people can provide the interacting parts for smart groups

ARTICLE HISTORY
Received 29 July 2016
Accepted 15 December 2016

KEYWORDS
Collective behavior; group cognition; measurement; networks; systems

1. Introduction

What kinds of things fall into the category of systems that are usefully described as employing cognitive processes? We propose that (most) humans fall into this category, but other systems do as well, including some collective swarms of insects as well as groups made up of multiple humans. Pushing this claim forward is revealing with regard to general principles of cognition as well as the specific nature of human cognition. Individual humans contain and integrate

CONTACT Robert L. Goldstone rgoldsto@indiana.edu

many cognitive mechanisms, resulting in systems that readily integrate with each other and their environment to produce higher level cognitive systems. Understanding group cognition, then, becomes a matter of determining how a cognitive system at a higher level can subsume cognitive systems at a lower level, and how the systems at multiple levels can strengthen rather than diminish one another.

2. Positioning a claim worth defending

In the current paper, we will be arguing for the explanatory fertility of a *multiple, interacting levels of cognitive systems* (MILCS) perspective. In adopting this perspective, we will deliberately steer clear of other, more assumptive, conceptually loaded, and controversial framings. We will not discuss "extended" minds because the notion of "extending" has the infelicitous implication that there is an original, core cognitive system that is merely augmented by a peripheral system. This implication is appropriate in the case of Clark and Chalmers' (1998) Otto, the man who relies on his notebook as an external memory aid that helps him find his way to the art exhibit he wishes to attend. For Otto and his notebook, talk about Otto's mind extending beyond Otto's skull is natural because Otto is undeniably a cognitive system in his own right, and a relatively small piece has been tacked onto this system. However, we will consider far more symmetric cases in which a cognitive system encompasses two or more people, each of whom can make a strong case for being a complete cognitive system on their own.

Our discomfort with the "extended" part of the "extended mind" thesis extends to the "mind" part as well. The term "mind" is evocative and spurs the imagination, but it is unnecessarily vague and difficult to operationalize. It brings along superfluous conceptual baggage for our present purposes. Accordingly, our arguments will concern "cognitive systems" instead of "minds." By "cognitive" we have in mind the kinds of systems that cognitive scientists spend their time researching, including perception, attention, memory, categorization, decision-making, problem-solving, and judgment. By directing attention away from minds writ large and toward the cognitive processes that make up thinking, we may be fairly accused of sidestepping the hard problem. Our only defense is that there is more to cognitive systems than meets the eye, and an investigation of their "endless forms most beautiful and most wonderful" (Darwin, 1859, p. 489–490) provides a helpful, grounded account of the functional components of things that we may eventually want to christen with the honorific "mind." We do not intend to deflate the rich notion of "mind" and reduce it to an interacting network of cognitive systems. For now, we leave it to others to determine if there is a useful notion of "mind" that goes beyond what would be covered by speaking of systems of cognitive systems. We suspect that much of the scientific work to be done in cognitive science's near future will consist in articulating the capacities of different arrangements of

different cognitive systems and documenting the different ways in which these systems are implemented. That is our main justification for reframing questions about extended minds into questions about what kinds of systems are capable of what kinds of cognitive functions.

We believe that progress in understanding how interacting groups of people can think and act as a coherent agent will depend on using many of the same systems principles that have been recruited for explaining how individuals think and act as agents. In an attempt to avoid possibly "triggering" language related to thinking and acting (e.g., "That's not real thinking. That's simply reacting."), we prefer to frame the accounts that unify individual and group behavior patterns in terms of the specific processes that are required at both levels. For example, both individuals and groups will often need to assure that their parts work in close coordination rather than at cross purposes, make decisions that integrate diverse information sources in an accurate and timely manner, and engage in useful behaviors by exploring, evaluating, and selecting appropriate actions. We focus our attention on processes like these rather than thinking and agency per se because their involvement in group and individual dynamics is relatively clear, and they provide the underpinnings for cognitive activities. In the final portion of our paper, we briefly speculate on how our systems-theoretic approach might be extended to encompass more traditional analyses of individuals and groups as "intentional systems."

In framing our interests in terms of *cognitive systems*, we emphasize the sets of interconnected parts that give rise to wholes that serve cognitive functions. To posit a system is to posit a division between elements inside and outside of the system. In the physics of an ideal gas, the relevant system consists of a containing vessel and point particles that interact via elastic collisions. All other elements are outside of that system, including sunlight, sub-atomic particles, and electrical charges.

The same physical stuff can be understood as manifesting different systems. For example, footbridges can be understood as resonating oscillators, as a network of connected edges spanning land nodes, or as a system of equilibrium tension incorporating a deck, towers, and cables. Ascribing systemic organization to physical stuff is to take a particular perspective toward it. These perspectives are better thought of as being more or less apt rather than as being true or false. What makes an ascription apt is that it provides a useful explanation of the behavior of stuff at some level and predictions for how the stuff will behave. We are particularly interested here in explanations that are useful for scientific purposes, and aptness is couched in the section "Inductive Power" in terms of the extent to which valid, task-relevant predictions are possible, given that a system has been ascribed. If a particular system perspective is not adequately considered, as was the case with the resonating oscillators perspective on the Tacoma Narrows Bridge, then its behavior (collapsing in 1940) may be disastrously surprising.

To say that a set of parts can be aptly understood as composing a particular system is to claim that the whole's behavior is explained or predicted by an understanding of that system's structure and dynamic processes. Whether parts are aptly considered to belong to a system is, in part, dependent on other notions related to distributed cognition. One such notion is the *coupling* between parts (Adams & Aizawa, 2008; Clark & Chalmers, 1998). Parts are structurally coupled if they are bound together into one unit, as when two mechanical shafts are coupled together by a sleeve or clamp; and they are dynamically coupled if they reciprocally interact, such that state changes in one part continually cause changes in the other part, and so forth. In Clark and Chalmers' original example, Otto is strongly coupled with his notebook to the extent that it is reliably available to Otto whenever he needs it, with its contents readily accessible and modifiable, and typically endorsed as trustworthy. Rupert (2010) presents a related notion of tightness in the co-contribution of mechanisms to the production of the cognition. Hutchins (1995) suggests individuating cognitive systems by the gradient of computational activity; parts belong to the same cognitive system to the extent that they are engaged in considerably more computations than subsets of the system. Sterelny (2010) suggests a graded approach to understanding whether a collection of parts forms a cognitive unit, including considerations of trust, interchangeability, and the collectivity of resources. In a similar vein, Heersmink (2015) analyzes "cognitive integration" as a multidimensional phenomenon in which the level of integration can vary across a number of more or less independent dimensions, including the kind and intensity of information flow between agent and scaffold, the accessibility of the scaffold, the durability of the coupling between agent and scaffold, the amount of trust a user puts into the information the scaffold provides, the degree of transparency-in-use, the ease with which the information can be interpreted, the amount of personalization, and the amount of cognitive transformation.

Another notion is the *bandwidth* with which subsystems are connected (Haugeland, 1998), expressible by a measure such as "bits per second." Subsystems become more prototypically modular as the bandwidths of the channels connecting the subsystems become increasingly narrow. As the channels' bandwidths approach the information transmission speed and capacity within a subsystem, the motivation to carve up the whole into discrete subsystems decreases.

Coupling and bandwidth do not directly address the question of whether parts are aptly construable as belonging to a particular kind of system. While parts which can be said to aptly compose a system are often strongly coupled and connected by high bandwidth channels, sometimes they are not. Sometimes a social network that is composed of less strongly coupled parts better serves a particular cognitive function (Weick, 1976). One well-documented such case is that of social networks and their search for good but hard-to-find solutions in large problem spaces. Both experiments with people (Derex & Boyd, 2016;

Mason, Jones, & Goldstone, 2008) and computational models (Goldstone, Wisdom, Roberts, & Frey, 2013; Lazer & Friedman, 2007) show that if the best solution in a problem space is difficult to find because it involves non-linear interactions of its elements, then it can be better to connect the social network more sparsely and more locally rather than densely and globally (see also Zollman, 2010). The problem with fully connected social networks is that decent, but not ideal, solutions may be stumbled upon by an agent early in the search. If everyone can see the solutions that everyone comes up with, then there will often be a premature convergence of agents on the good solution (Hutchins, 1995), thereby preventing the group from fully exploring the entire search space adequately for better solutions. For difficult searches, it is better to limit the connectivity between parts of the social network if the goal is for the network to find the overall best solution.

Another case in which a system better fulfills its information processing cognitive function by partially decoupling its parts is found in individual human brains (Sporns, Tononi, & Edelman, 2000; Tononi, Edelman, & Sporns, 1998). Brains need to quickly and reliably extract important features from sensory inputs, which biases perceptual systems toward developing relatively well-encapsulated modules. However, if the results of these specialized perceptual processes are not well-integrated with other modules, then a coherent interpretation will not arise that benefits from the entire brain's input. A neural network in which all brain regions are very tightly coupled risks having all brain regions redundantly "know" the same things, thus informationally overloading brain regions. At the same time, a neural network in which some brain regions are not connected by a path to other brain regions risks generating a cognitive system that develops incoherent and conflicting interpretations. The tradeoff between segregation and integration of brain regions can be formally captured by an information-theoretic measure of neural complexity that is maximized not by a fully connected neural network but rather by a network that optimizes total mutual information.

The juxtaposition of these two examples, purposely selected from the study of social and neural networks, provides an early glimpse into the possibility of a cognitive systems perspective that is equally relevant to both individual and group cognition. An issue that confronts all information processing systems is how the parts should be organized so as to make the flow of information useful. The above examples indicate that this is sometimes achieved by limiting the bandwidth of communication between the parts. It is exciting to contemplate that there might even be common principles underlying both neural and social networks that relate aspects of the connectivity of the networks, the nature of the task the network is trying to solve, and global measures of the network's performance (Dedeo, 2014). One example of an apparent generalization, shown in Figure 1, is that as the globally best solution to a problem becomes increasingly difficult to find, restricting a network to increasingly local connections

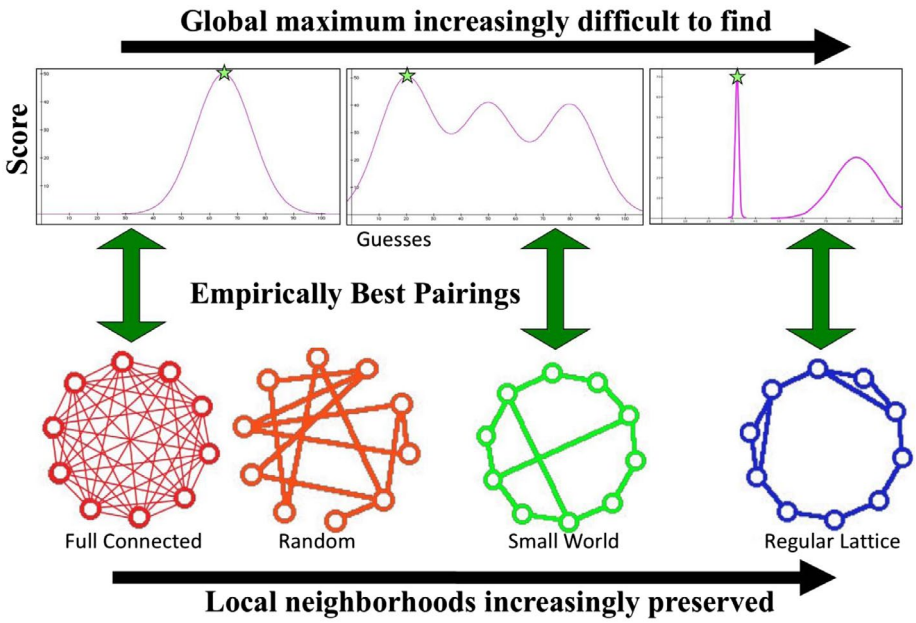

Figure 1. The upper three panels show three different functions used by Mason et al. (2008) for converting participants guesses on a 1–100 scale onto scores. In the left panel, the conversion is simple and the closer a guess is to particular number, the higher the score is. In the middle panel, the function has three peaks and so participants might get stuck on one of the two local maxima and miss the global maximum that is achieved with a guess of approximately 20. The right panel shows a function with an even harder global maximum to find because of the presence of a much broader local maximum. The lower portion of the figure shows examples of four different network structures for groups of 10 participants. Circles represent participants and lines indicate communication channels via which participants could see each others' guesses from the previous round and the scores earned by those guesses. The empirically best pairings of network structure and scoring function are shown by the bidirectional arrows. As the global maximum for a problem gets increasingly difficult to find, communication networks that increasingly limit long-range connections perform increasingly well. While the fully connected network serves to "get the word out" quickly when the global maximum is easy to find in a search space, it tends to cause participants to prematurely converge on decent but not optimal solutions when the global maximum is harder to find. For the "needle in the haystack" function on the top right, it is better to configure a social network so that different local regions of the network will explore different regions of the solution space (figure adapted from Goldstone et al., 2013).

allows different regions of the network to independently search in different areas of the problem space without over-hastily jumping onto a single bandwagon (Goldstone et al., 2013). This emerging area of network science is replete with examples of formalisms that have been applied to both individual neural networks (i.e., brains) and collective social networks (Barabasi & Albert, 1999; Boccaletti, Latora, Moreno, Chavez, & Hwang, 2006; Bullmore & Sporns, 2009; Girvan & Newman, 2002), often with the felicitous outcome of developing insight into both kinds of networks through their conceptual unification. From this perspective,

similar issues confront both kinds of networks: spread of information, modularization, robustness to attack, speed/accuracy tradeoffs in decision-making, and refinement so as to be responsive to signal but not noise. The same specific quantitative models have been applied to both kinds of networks with sufficient predictive success to allay concerns that the parallels are purely metaphorical or too vague. The successes of network science in unifying individual and group cognitive processes under common models provide an initial motivation for taking seriously the possible benefits of positing common principles that span from individual to collective cognitive systems.

Network science encourages us to think of the architecture of human cognition in terms of neural networks that participate in robust social networks, with both kinds of networks shaping and influencing one another (Goldstone & Gureckis, 2009). Our individual human thoughts both depend upon and determine the social structures that contain us as elements within those structures. The treatment of cognitive systems as arising at multiple levels is congruent with other frameworks in which multiple levels of analysis are necessary to fully understand phenomena such as language (Beckner et al., 2009), selves (Thagard, 2014), or evolutionary selection (Sloan Wilson & Wilson, 2007). Cognitive scientists often assume that individuals are the sole loci of organized cognition, but a network science perspective helps remind us that organized behavior can be described at multiple levels. In the next section, we will argue that some of this organized collective behavior can be effectively understood as implementing mechanisms that are importantly similar to those that generate cognitive functions in humans.

3. Cognitive systems in collectives

A cognitive systems perspective promises to unify collective and individual cognition in revealing ways. The kinds of cognitive systems that we have in mind are more fine-grained than a standard, coarse, flowchart of information processing that includes boxes for processes like short-term memory, long-term memory, perception, attention, and planning. Instead, our focus is on cognitive mechanisms like shifting attention by gating (Mayer et al., 2009), response discretization with winner-take-all lateral inhibition networks (Grossberg, 1976), developing differentiated and specialized units by having originally homogeneous units compete for the "right" to adapt to input patterns (Rumelhart & Zipser, 1985), interactive activation for combining top-down and bottom-up information (McClelland & Rumelhart, 1981), and decision-making via a diffusion process (Forstmann, Ratcliff, & Wagenmakers, 2016). Because these kinds of mechanisms may not be generally known outside of specialized cognitive science communities, we will describe two such cognitive mechanisms.

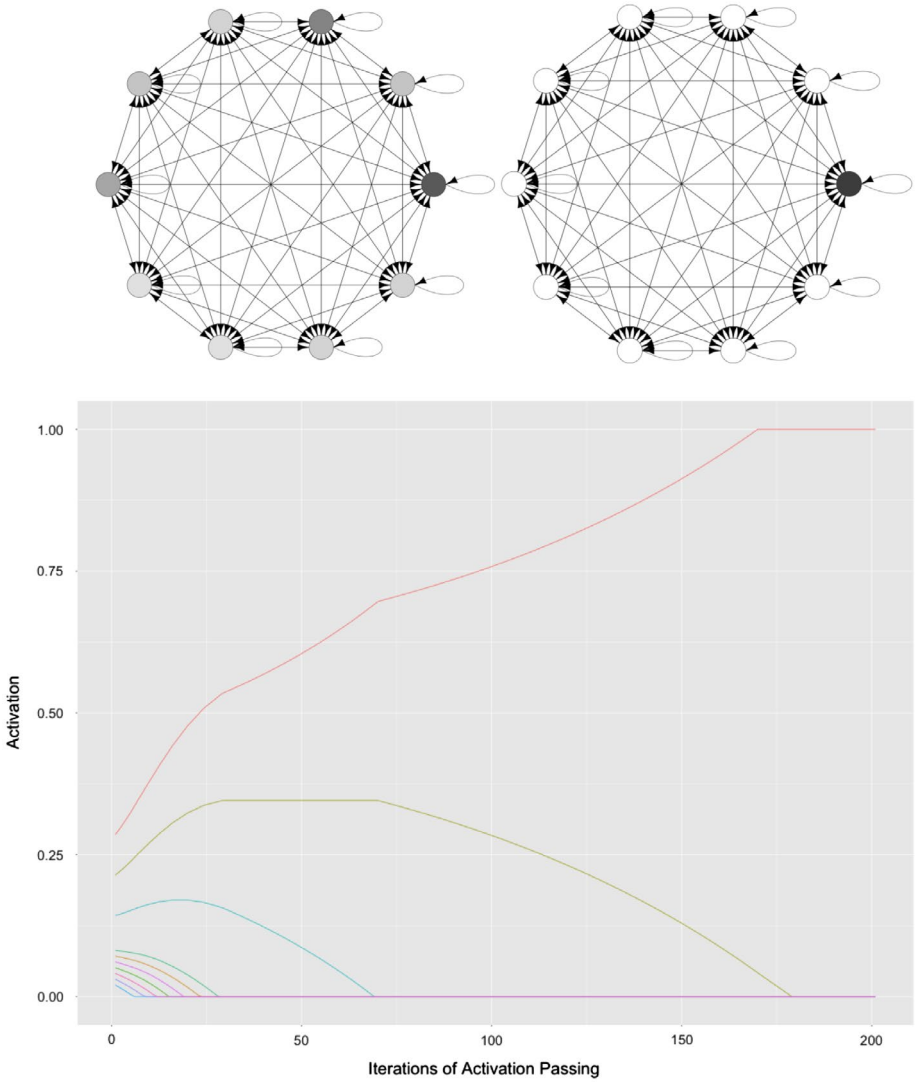

Figure 2. A lateral inhibition network for achieving winner-take-all decision-making. In the top left panel, an initial network contains 10 units each of which is connected to itself and the other units. The activation of each unit is shown by its darkness. After 200 iterations of activation passing among the units, one unit is maximally activated while the other nine units are completely inactive. The plot below shows the activation levels for each of the 10 units over the course of the iterations. By iteration 180, only a single unit, the unit with the largest initial activation, has a positive activation, and that activation level is the maximal level of 1.

3.1. Lateral inhibition networks

Figure 2 shows a simple *lateral inhibition* network consisting of ten units, whose activity levels are shown by the darkness of the unit. Initially, the units have various levels of activity, as shown on the left network. Each unit is connected to itself and all of the other units, and the units exchange activity along these connections

for 200 iterations. The network is called a "lateral inhibition" network because the units are understood as all being at the same, lateral level of a larger, possibly hierarchically structured network, and because all of the nodes are mutually inhibitory. As a result, by the end of activation passing, one unit is maximally active while all of the other units are completely inactive, as shown by the network on the right. The unit that was initially the most active becomes the maximally activated "winner," and the activity update can be formulated so that this is guaranteed no matter how many units are in the network, how small the initial advantage is for the most active unit, or how the other units' activities are distributed.

Notice that we have intentionally refrained from giving the units concrete designations such as neurons, words, concepts, candidates, or companies because the same lateral inhibition network has been used to describe the interactions among all of these units and others. The network's inhibitory dynamic is functionally useful whenever there is a need to decisively select one option out of many alternatives based on the options' initial evidentiary support. The network serves this function without any need for a "CEO unit" that makes the choice. The winner is selected in a decentralized manner. As an example, individual humans need to disambiguate the meanings of ambiguous words such as homonyms. There is evidence that when an ambiguous word like "bank" is first presented, both of its meanings (financial institution and side of a river) are first activated, but then the dominant meaning of the word becomes increasingly active while the non-dominant meaning simultaneously decreases its activity. This apparent lateral inhibition is particularly strong in the left compared to right hemisphere (Burgess & Simpson, 1988). As another example, there are models of the dynamics of consumer product popularity in which initial popularity differences become accentuated over time because consumers tend to want to purchase products that already have appreciable buy-in by others (Liebowitz, 1995).

3.2. Diffusion for decision-making

A universal problem for cognitive systems to solve is how to make decisions in a reasonable amount of time. Life is short, environments change, and opportunities do not always last. One rigorously formulated and empirically well-tested mechanism for decision-making is a diffusion process, by which information continuously accumulates and a relative rule is used to determine whether a threshold amount of evidence has been reached in order to respond (for a review see Forstmann et al., 2016). Figure 3 illustrates the diffusion process for a task requiring an observer to decide whether a subset of dots are moving to the left or right within a display of randomly moving dots. The diffusion model posits an evidence accumulator that begins at a *starting point* and continuously accumulates noisy evidence at a rate determined by *drift rate*. There are *upper and lower boundaries* that serve as thresholds for decision-making. When the accumulator crosses one of these boundaries, the decision-maker makes a "left" or "right" decision,

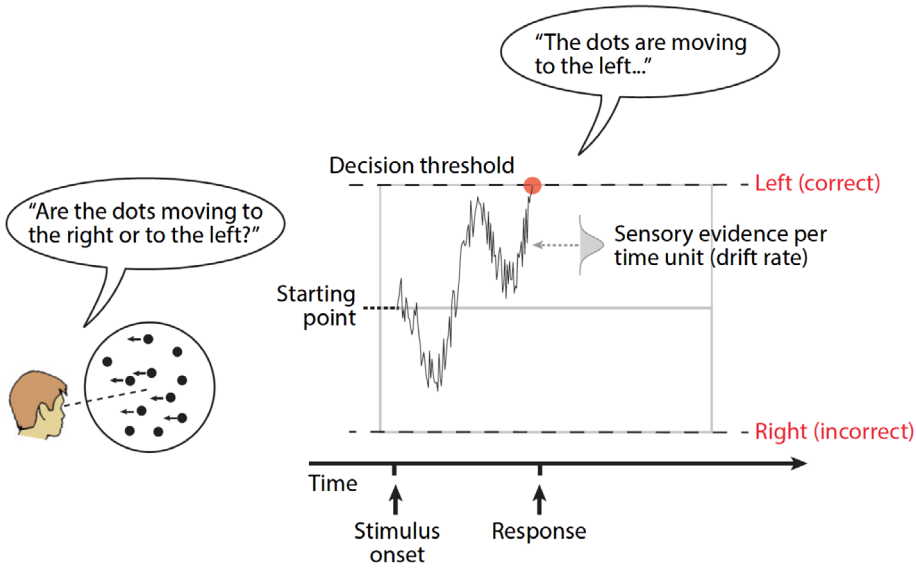

Figure 3. An illustration of a diffusion process for two-choice decision-making (from Mulder et al., 2012). The specific task asked of subjects in this case is to decide whether a subset of consistently moving dots in a field of randomly moving dots is moving to the left or right. The Starting Point determines the beginning location for the evidence accumulation process. At each moment, evidence is provided by senses that noisily changes the location of the evidence accumulator with a rate defined by Drift Rate. The subject responds "left" or "right" when the evidence accumulation cross the upper or lower boundaries, respectively. (Reprinted with permission from the authors).

respectively. The model makes predictions for response accuracy and the entire distribution of response times. Critical parameters in the model are influenced by individual differences and conditions of the task. The starting point reflects an a priori bias to respond to one choice over the other. The drift rate reflects the quality of information processing. The boundaries reflect the degree of caution in response; as the gap between the boundaries becomes larger, decision speed is sacrificed for accuracy.

The diffusion model is notable for its wide, level-spanning sphere of application. Originally formulated by Einstein (1905) to describe the random movements of particles as a result of molecular collisions, the model has been applied to many chemical, physical, neural, biological, and social phenomena in which the position of an entity reflects the cumulative effect of many small forces acting on it. Important for our purposes, the same diffusion model effectively captures the details of individual neurons' firing rates (Gold & Shadlen, 2007), individual human cognitive tasks such as recognition memory, categorization, and choice, and collective phenomena such as foraging in acellular slime molds (Latty & Beekman, 2011), and determining new nesting sites in ant colonies and bee hives (Marshall et al., 2011). In some ways, it is not surprising that the same model applies across these levels because all of these systems need to make accurate and

fast decisions. In this context, the diffusion model can be shown to be optimal in that it achieves the fastest mean decision time for a given accuracy (Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006). The optimality of the diffusion mechanism provides at least a partial response to Toon's (2014) dissatisfaction with distributed cognition (d-cog), namely, that "while d-cog might show how social interactions can implement a computation, it will not tell us why a particular representational scheme is used to carry out that computation" (p. 123). The optimality analysis suggests a good a reason for why many different systems that are in the business of making good decisions fast would benefit from adopting the diffusion mechanism. If such a mechanism can be evolved or learned, then the system will tend to stick around long enough to grow, reproduce, or extend its domain.

Generalizing from these two examples, the MILCS perspective holds that certain kinds of systems are able to achieve specific functionalities that are relevant for the cognitive processes involved in perception, attention, memory, problem-solving, and decision-making. These systems offer mechanisms for achieving cognitively important functions such as selecting an option that is most attractive among its competitors, achieving consensus in a network so that the network acts in a coherent manner, selectively attending to information globally assessed as relevant, developing specialized units from originally homogeneous networks through competition among units, and combining data from perception with top-down theories and models. These mechanisms help to explain a number of lawful regularities about how these systems behave under a variety of conditions. We call these systems "cognitive systems" because they instantiate cognitive mechanisms, and we call these mechanisms "cognitive mechanisms" because they are mechanisms in virtue of which these systems perform important cognitive functions associated with flexible, adaptive behavior. Thus, if we understand how these mechanisms work, we understand how cognition works. Human beings are undoubtedly fine specimens of cognitive systems in this sense. Thus, by extending models of human cognition to groups, we find it exciting to discover cases in which both instantiate the same types of cognitive mechanisms to solve similar problems.

3.3. What makes a system "cognitive?"

Our claim has been that system processes, such as a diffusion process for speeded decision-making or lateral inhibition for selecting from among competing alternatives, provide mechanistic underpinnings for cognition. We suspect that all naturally occurring cases of cognition are built out of processes like these, but these processes also occur in cases that would not normally be construed as cognitive. This raises the question: What distinguishes the cases that *are* versus *are not* aptly construable as cognition? While this question is natural, we prefer to approach it obliquely because we are doubtful of the existence of a simple adequate criterion for what counts as cognitive. "Cognition" and "thinking" are multifaceted terms

involving many partially overlapping notions. Empirically, different species have different portfolios of capabilities related to memory, space, attention, representation use, communication, and perception. Which species fall above or below the criterion for thinking will depend on the weightings given to different capabilities. These weighting are likely to be somewhat arbitrary, and very possibly anthropocentrically based on the degree to which we individual humans demonstrate the capability. If chimpanzees used the ability to accurately recall the locations of nine numerals when presented for 210 ms as the criterion for real thinking, then it would be they but not humans who could think (Inoue & Matsuzawa, 2007).

If cognition is indeed multifaceted, then a pluralistic conceptualization will likely offer the most promising approach for providing its characterization. In this regard, the behavioral flexibility that a system shows in achieving its goals provides a useful guide to designating a system as cognitive because flexibility is itself multiply determined. Systems with increasingly sophisticated memory, attention, planning, problem-solving, search, communication, and decision-making processes have increasing flexibility in the methods that they can deploy to achieve their goals despite environmental obstacles. So although we wish to avoid proposing a criterial "mark of the cognitive," we find it increasingly attractive, perhaps even eventually unavoidable, to consider a system to be capable of cognition to the extent that it can flexibly achieve its goals despite imposed challenges.

To resist the knee-jerk reaction that our systems approach to cognition is overly promiscuous, it will be instructive to compare it to a related proposal by Rupert (2009, 2010) which places a similar emphasis on the role of systemic integration.[1] For Rupert, the cognitive status of an entity X (e.g., neural activity in the hippocampus) derives from X's being part of an integrated cognitive system S, where S in turn is characterized as a collection of mechanisms or capacities which "contribute causally to the production of a wide range of cognitive phenomena, across a variety of conditions, working together in an overlapping way with a variety of other mechanisms of similar standing" (Rupert, 2009, p.41). We recognize it as a virtue of Rupert's proposal that it does not frontload philosophically charged notions such as consciousness, intentionality, or mental representation. Moreover, we take it as common ground here that the occurrence of flexible, adaptive, goal-directed behavior is among the cognitive explananda relative to which the question of cognitive systems individuation arises.[2] How, then, does our "cognitive systems" perspective fare when it is judged in terms of Rupert's proposal?

Consider first the required tightness in the co-contribution of elements that are jointly responsible for the production of the cognitive explananda, as measured by the frequency in which they are regularly co-employed. Citing this criterion, Rupert argues at length against (canonical examples of) the "extended mind" thesis, on the grounds that there is a steep gradient of computational activity which sets Otto's brain-bound cognition apart from the infrequent, specialized dealings with his notebook. In contrast, we claim that no such asymmetry between center

and periphery exists in the systems we have discussed, such as diffusion processes for decision-making and lateral inhibition networks for choice selection. In these systems, there is no single part which it is tempting to treat as a computationally privileged locus of control. For example, the choices that result from the activity passing in the lateral inhibition network arise from the densely coupled interactions among all parts. Furthermore, while abiding by our policy to avoid the term "mind" other than as a placeholder for systems of cognitive systems, we maintain that in many groups that we observe, mechanisms of attention, learning, decision-making, and the like causally interact in overlapping ways, under a variety of conditions, in the production of flexible, goal-directed behavior. In this regard, the fact that all the mechanisms we have discussed are entirely domain-general, and not restricted to specialized tasks, is bound to boost the frequency of their co-employment and thus their level of systemic integration sensu Rupert.

A second point of comparison concerns our characterization of the cognitive explananda. From Rupert's perspective, it might be objected that our description of cognitive mechanisms as "diffusion" or "lateral inhibition" processes is ontologically too permissive, because it cannot distinguish between *cognitive* and *non-cognitive* instances of the same (broadly information-theoretic) kind (Rupert, 2011). One way of getting at such a distinction is to focus on the nature of the respective outcomes of the mechanisms to which those processes are tied, assuming that a firm line between cognitive and non-cognitive outcomes can be drawn.[3] For example, a neural mechanism N, firing in isolation in a petri dish, does not by itself constitute a cognitive system. However, N might be part of a cognitive system if is functionally integrated in a larger neural network, lodged inside a human brain, where its activation contributes to a wide variety of cognitive outcomes. Indeed, we find that the individuation of mechanisms in general is context-dependent. Mechanisms are always mechanisms *for* doing certain things (Craver & Bechtel, 2006; Machamer, Darden, & Craver, 2000). Importantly, this means that our characterization of the phenomena which the operation of a mechanism is meant to explain depends in part on what kind of task we take the mechanism to perform (and vice versa). Less obviously, we often learn new things about how a mechanism works by examining the environment into which it is geared and the systemic properties on which its operation depends (Anderson, Richardson, & Chemero, 2012; Bechtel, 2009).

The same principle, we suggest, works also for groups. Here, too, a vital part of our motivation for considering the workings of (say) a lateral inhibition network as the mechanistic basis of *decision-making* processes stems from the fact that its dynamics contribute to the production of a wide variety of behaviors in which groups frequently engage, and that warrant being termed *decisions*. In particular, we maintain that a lateral inhibition network is deployed for a particular computational function—it selects the most attractive option from a set of options in a decentralized fashion. Furthermore, it implements the cognitively important function of warping sensory signals that are linearly related to physical

qualities in a non-linear manner, transforming analog inputs into quasi-discrete, quasi-symbolic encodings (Goldstone & Hendrickson, 2010). A basic feature of human symbolic thought is that people form equivalence classes. In the classical notion of an equivalence class, distinguishable stimuli come to be treated as the same thing once they have been placed in the same category. The lateral inhibition network provides a mechanism for the origin of these (near-) equivalence classes by assigning a single, maximal activation level to any graded activation that is higher than its competitors. Hence, the network can provide us with equivalence classes, the beginning of proto-symbolic thought (Harnad, 1987). So, although this network dynamic is not exclusively found in systems that we would want to call "cognitive," it does provide an example of the kind of mechanism that allows systems, including humans, to perform cognitive functions related to forming equivalence classes, action selection, and decision-making.

This last point confirms a general feature of mechanistic explanation, in particular as practiced in cognitive science. Earlier, we remarked on the epistemic gains derived from situating a cognitive mechanism in its wider context, but it seems to us that the reverse route to discovery is at least equally consequential. That is, in virtue of conceptualizing a plethora of physically diverse mechanisms in a common information-theoretic parlance, scientists are bound to arrive at new characterizations of the phenomena they were originally trying to study. Mathematical models in particular have the potential to unearth structural similarities that cut across intuitive category boundaries. Because of their formal nature, they do not have any in-built semantic restrictions regarding the domain to which they can be fruitfully applied. This opens the potential for significant conceptual drift when successful models of individual and collective behavior are either borrowed from, or extended beyond, the psychology of human beings. The history of cognitive science is already rife with examples of such perspectival shifts, or "reconstructing the phenomena" (Bechtel, 2008), and we predict that the shift in orientation that we advocate will likewise affect our understanding of what phenomena are legitimately viewed as "cognitive" explananda (see Figdor, in press).

3.4. The systems cure for cognitive bloat

Talk about extending minds beyond individual human brains and bodies leads to the criticism that the category of "cognition" becomes overextended to the point where it is explanatorily useless (Adams & Aizawa, 2008; Rupert, 2009, 2011; for responses see Clark, 2001, 2010). The worry here is that the principles underlying the ways in which people interact with their tools and other people do not form a natural kind for investigation. The way in which Otto consults his notebook to find his way to the art museum is different from how a person with an unimpaired brain would simply access her biological memory. Scientific regularities that are found regarding human memory, like primacy effects, recency effects, and

chunking would not necessarily show up in the Otto + notebook system (Adams & Aizawa, 2008).

However, network science and systems science provide exactly the kinds of unifying principles and mechanisms that yield compelling accounts for how and why superficially different systems adopt certain configurations and exhibit certain dynamics. These explanations have two desirable properties: they are inductively powerful and they are constrained. We describe these properties with respect to the two systems previously presented.

3.4.1. Inductive power

System dynamic models fall into a number of categories, such as: lateral inhibition network, diffusion, preferential attachment, competitive specialization, positive feedback, negative feedback, small-world network, scale-free network, back propagation, reinforcement learning, and multi-level deep learning. These categories are inductively powerful in the sense that if a natural phenomenon can be aptly cast as instantiating one of these systems, then a considerable amount of its behavior can be explained and predicted. These categories are, admittedly, different from standard taxonomic categories that would divide phenomena into social versus natural phenomena, then subdivide social phenomena into psychology, economics, and sociology, and then sub-subdivide psychology into cognitive, clinical, and social fields. We do not deny that these categories provide one way of structuring the world in inductively powerful ways, but they are not the *only* way, and a systems perspective offers another way.

Consider the system process of diffusion limited aggregation, in which many individual elements move randomly and if a moving element touches another element it becomes attached (Ball, 1999). The emergent result is a fractally connected branching aggregate. Four examples, selected from a variety of disciplinary fields, are shown in Figure 4. To a scientist armed with a model of diffusion-limited aggregation, there are striking similarities between the patterns found in London streets, electrical discharge in plexiglass, window frost, and lungs. All consist of fractal arrangements of branches on branches. Although highly perspective dependent, this system concept is inductively potent. Once frost is appreciated as an example of diffusion-limited aggregation, one can predict how it will change with time, how it will be affected by temperature, what shapes it can and cannot attain, and so on (Goldstone & Wilensky, 2008).

A similar inductive power is found for lateral inhibition and diffusion-based decision processes. For example, Figure 2 shows a stereotypical pattern of activity changes in lateral inhibition networks such that units that are initially nearly as active as the most active unit will *increase* their activity for a time before they *decrease* their activity. An example of this dynamic is that the non-dominant meaning of a homonym will initially increase its activity before it is eventually quelled by the more rapidly rising dominant meaning. Diffusion decision processes likewise form an inductively powerful category. For example, human

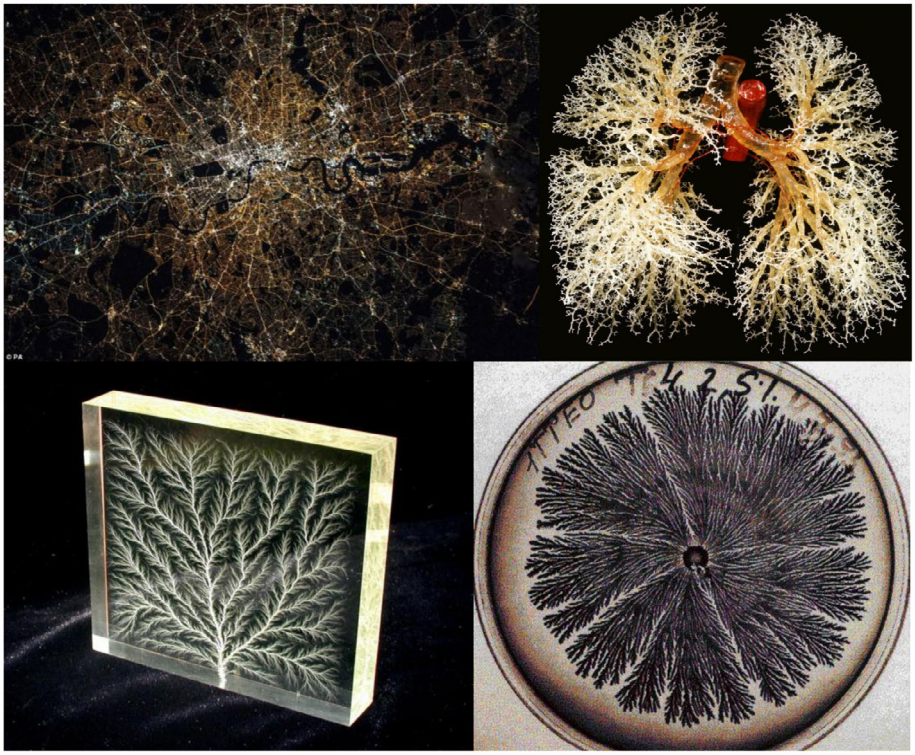

Figure 4. Four examples of Diffusion Limited Aggregation taken from sociology (upper left: a night satellite image of London), medicine (upper right: a human lung), physics (lower left: the pattern of high-voltage electrical discharge in plexiglass), and biology (lower right: the bacteria Bacillus subtilis growing on an agar dish). Photographs courtesy of Wikipedia.

decision processes that are governed by diffusion share a number of characteristics: speed increases lead to more errors (all else being equal), the mean response time is proportional to the standard deviation of the response times, manipulations that increase the speed of correct responses will also increase the speed of error responses, and response time distributions skew to the right and the extent of this skew increases with task difficulty. While not all of these properties have been empirically tested for colony insects, several have, and those have been found to be consistent with these model predictions (Marshall et al., 2011).

The inductive power of systems principles plays an important role for motivating the unification of accounts of individual and group cognition. Some analyses (e.g., Rupert, 2009) frame claims that individuals and groups both cognize as claims that the cognition in both is of the same "kind." Recruiting the notion of "kinds" when discussing cognition is dangerous because it conjures natural kinds such as water and squirrels. We do not want to commit to different examples of diffusion as being the same kind in the sense that they reflect the inherent structure of the natural world rather than the interests of human beings. Instead, by describing a common diffusion process for individual and group decision-making,

we are proposing that the account is inductively powerful for allowing scientists to make useful, common predictions about different levels of behavior. The notion of inductive power lets us maintain that different kinds of systems manifest important common behaviors without adding the overly restrictive stipulation that the different systems "really" be the same kind in a way that does not depend on human epistemic needs to predict and explain behavior.

3.4.2. Constraint

Scientists prefer their models to be constrained in terms of their account of data, but also want their models to be as general as possible in terms of domains of application. While the explanation "God made it that way" is commendable in terms of its potentially universal applicability, we generally prefer more defeasible explanations—those that could have been wrong but happen not to be wrong given our understanding of the best data currently available. Models are respected to the extent that they "stick out their necks" and make non-trivial predictions that nonetheless turn out to be true.

The diffusion model of decision-making generally has the property of being appropriately constrained. The model fails to account for artificially generated but plausible data patterns (Ratcliff, 2002). This is a *positive* characteristic because it indicates that the model is not so flexible that it would be able to provide an account for any kind of data that was obtained, as is the case for the "God made it that way" pseudo-explanation. Likewise, the standard diffusion limited aggregation model generally cannot provide an account of branches that cross directly through one another. Likewise, the lateral inhibition network is constrained to predict that the moment-to-moment change of a unit's activity is proportional to its current activity level relative to the other units. The standard lateral inhibition network does not have any memory for previous activity levels, and so would have limited utility in accounting for situations in which early, historic activity levels impact future activity levels.

System models generally *do* have some flexibility, provided by free parameters that govern the shape of predicted patterns of data. A free parameter is one that is estimated from data with the goal of maximizing the match (fit) between the model's predictions and the data. In the case of the diffusion model, three of the most important parameters for fitting empirically obtained data are starting point, drift rate, and boundary separation. Parameters such as these are looked upon with healthy skepticism, as vividly communicated by John von Neumann when he told Enrico Fermi, "With four parameters I can fit an elephant, and with five I can make him wiggle his trunk" (Dyson, 2004). However, practically speaking, this usually turns out not to be the case with cognitive models. Even with several parameters, most models are highly constrained in terms of the data that they can model. In addition, they are typically fit to complicated data-sets that often simultaneously vary two to four variables. Many cognitive modelers trying to get their PhDs can attest to the frustration of having a model incapable of fitting

apparently innocuous data, and often this is because critical aspects underlying the behavior have not been incorporated into the model.

The flexibility of models provided by their free parameters is often valuable because it integrates different situations by showing both their shared fundamental dynamic and the ways they differ on their specific parameters. For example, diffusion models can often model differences in task difficulty by varying just drift rate, differences in incentives (e.g., the penalty given for incorrect response), boundary separation, or priming from previous trials by varying starting point. These parameter variations allow cognitive scientists to pinpoint the specific ways in which otherwise similar mechanisms differ.

A particularly interesting application of this strategy to the investigation of differences between individual and group cognitive processes is the development of the Ornstein-Uhlenbeck extension (Bogacz et al., 2006; Busemeyer & Townsend, 1993) to the diffusion model. This extension adds one more parameter that allows the drift rate to depend on the current position of accumulated evidence, rather than leaving the drift rate independent of the position as in the standard model. For example, the parameter could be given a value that would make evidence accumulation favor movement toward a boundary as the evidence gets closer to the boundary. Exactly this biased accumulation of evidence has been found in colony insects such as honeybees (List, Elsholtz, & Seeley, 2008). Scout bees who find an attractive candidate nesting site recruit other scout bees to inspect the same site. As the colony gains evidence in favor of a nesting site, it becomes more likely to gain still more evidence in its favor. For many human judgments such as recognition, numerosity discrimination, and lexical decisions, this dependence of drift rate on current cumulative evidence is not found. However, this difference can be understood in terms of the needs of the spatially separated and independently acting bees to relatively quickly come to a consensual agreement about where the swarm should nest. The dependence of drift rate on current evidence allows for rapid convergence on a consensus decision in physically independent agents (Kao, Miller, Torney, Hartnett, & Couzin, 2014). Thus, the apparent disanalogy between individual and collective decision-making in this case need not reflect a fundamental difference in kind, but rather can be aptly viewed as supporting the postulation of a common diffusion process that has sufficient flexibility to accommodate independent or dependent accumulation of evidence. Whether dependence or independence is adopted by a particular system depends on a weighting of the relative risk of invalid conclusions due to confirmation bias against the risk of parts of the system fractionating and failing to converge on a single decision.

4. People as parts

Thus far, the cognitive systems we have considered have been purposely described in general terms, as containing interacting units, nodes, or agents, rather than humans per se. The question now arises: what difference does it make when people

with minds of their own function as the relevant units within such systems? [x] An obvious first difference is that people are cognitively far more capable than are the individual units that perform the rather limited operations described in the above two models. By itself, this point carries little or no weight, however, if we keep in mind that models in general selectively capture only a subset of the entities, properties, and relations in the empirical domain of their intended application. Considerations of parsimony suggest that we strive for models that accomplish a desired level of prediction or explanation with as few variables as possible. Thus, while the simple dynamic models that we have so far considered will likely need to become more complicated to fully accommodate the more complex decision-making, planning, and representation-building capabilities of people, this need not always be the case, as the success of some physics-based models of collective human behavior demonstrates (Ball, 2004). Even when these assumptions prove to be inadequate, the more austere models may still be apt and predictive, however. Whether models benefit from incorporating the added richness of people's individual cognitive machinery depends on the tradeoff between the decreased constraint and concision of the more nuanced models of humans and the models' hopefully greater fidelity to empirical results found in situations in which humans are in the loop.

The more substantial difference stems from the egocentric peculiarity that we feel as individual humans when we contemplate the larger, structured information processing systems that contain us as elements. A major source of resistance to thinking of a group of people as a cognitive system in its own right is very likely that the cognitive potential of the group is eclipsed by our focus on our own individual cognitive capabilities. The main worry, then, is that when we speak of *smart groups* of people, we are misrepresenting the human social domain by conflating different levels of analysis to which our models apply. Why do we need "smart groups" of people? Aren't smart people enough?

4.1. Multiple levels of cognitive systems

We contend that this question arises from assuming a "zero-sum" perspective in which there is a competition between accounts at different levels to explain the flexible, smart behavior of multi-level systems. To be sure, parsimony favors a zero-sum perspective towards competing accounts of a single phenomenon that are clearly pitched at the same level. For example, if our models could fully predict the scouting behavior of honeybees without taking accumulation bias into account, it would be gratuitous to posit a separate, independent search process that is sensitive to changes in drift rate. However, it is far less clear whether the same type of "exclusion" principle applies across a nested hierarchy of interacting levels of organization, as described by our MILCS framework. In this section, we elaborate our construal of groups as cognitive systems by reflecting on three potential

reasons for adopting a zero-sum perspective on group cognition enumerated by Huebner (2013) as part of his book-length treatment of collective mentality.

In the context of Huebner's book, the sobering effect of these three principles—which are perhaps more aptly described as injunctions—is to undercut deflationary appeals to "group minds" that are better, or at least equally well, explained by other means without having to treat groups as cognitive systems in their own right. Huebner introduces the three principles early to constrain what, from his perspective, can pass muster as a theoretically sound and empirically plausible defense of collective mentality. Having set himself this challenge, Huebner then develops a detailed, cognitive-scientifically informed account of what he calls "macrocognition"—system-level cognition that is implemented by an integrated network of specialized computational mechanisms—and argues that it avoids the injunctions which prevent (many) alternative conceptions of collective mentality from succeeding, such as establishing the irreducibility of collective mentality. As he frames his project, "my central hypothesis is that a plausible defense of collective mentality must take up the perspective of macrocognition, to demonstrate that groups are minded *in the same sense* that we are" (2013, p. 5).

In order to clarify what we hope to accomplish in this section with our discussion of exclusion principles, let us stress from the outset that our own approach differs considerably from Huebner's project. For starters, we do not wish to dispute Huebner's preferred computationalist manner of clearing the hurdles he set for himself, nor shall we question here his reliance on a distinctive model of computational architecture which he takes to underlie the mentality of individual human beings (and, per analogy for Huebner, the collective mentality of groups). In particular, we also do not aim to deny that the cognitive mechanisms we have canvassed involve some sort of mental representations or operations which can be captured in terms of their algorithmic structure. For example, the diffusion model can be seen as "accumulating evidence about the environment" and "messages about the popularity of consumer products" can be seen as being exchanged in lateral inhibition networks. For many cognitive models, such representational glosses are sensible because they allow us to connect specific aspects of our mathematical formalisms with particular features of the target systems we are trying to understand (Weisberg, 2013). However, precisely because these models also work in domains where a representational interpretation is dispensable, if not misguided, we do not inherently tie our analysis of cognitive systems to the aptness of our models to capture the creation and manipulation of mental representations (however defined). Instead, the structural and information-theoretic similarities among cognitive systems that inform our approach can be expressed formally as network structure-process pairs. As a result, the MILCS framework is bound up with, and actively encourages, a less anthropocentric perspective about the nature of psychological predicates. We are thus prepared to embrace the relative abundance of group cognition in the world, whereas the fruit of Huebner's admirable labor is to demonstrate that genuine macrocognition—of the sort that resembles

human minds—is in fact exceedingly rare. As opposed to Huebner, who compares the degrees of collective minds to the rungs of a ladder, reaching from the "minimal" mentality of bee hives to the "maximal" mentality of high-energy physics labs and their ilk, we envision the realm of cognitive systems to populate a more lattice-like, multidimensional region. With that caveat in mind, we proceed to a review of Huebner's three injunctions against positing group level cognition.

To begin with, we readily acknowledge that there are certain types of multi-level systems for which it is perfectly reasonable to adopt a zero-sum perspective on cognition. This means that to the extent that one level suffices to explain an episode of intelligent behavior, it makes the postulation of cognition at other levels unnecessary. For example, at least some of the stock examples of the "extended mind" literature, such as the case of Otto and his notebook, arguably fall into this category. Otto as a person is sufficiently smart that whatever augmented functionality is afforded him by his notebook, specifically how it transforms his ability to remember things, are largely dependent in a rather asymmetrical fashion on Otto's cognitive repertoire. Otto's "un-augmented" abilities are still doing most or all of the explanatory work here. Compared to the patterns of connectivity inside Otto's brain, the density of information flow from the notebook to Otto is minimal. Given the fixation of "extended mind" systems on a mid-level scale that typically involves solitary individuals plus external resources such as tools or artifacts, it makes sense to consider an organism's brain as the *center* (albeit not necessarily the exclusive locus) of extended cognitive activity (Clark, 2008, p. 95). Applying a zero-sum perspective, one would conclude that the Otto + notebook system is smart just because, and to the extent that, Otto is smart.

Our examples of systems like diffusion processes for decision-making and lateral inhibition networks for choice selection seem to have a quite different status than Otto + notebook. In our examples, there is no single part that it is tempting to treat as a cognitively privileged center. The choices that arise from the activity passing in the lateral inhibition network arise from the interactions of all parts. Furthermore, even if the parts (perhaps people) were capable of cognition themselves, the choice made by the lateral inhibition network need not match either the choice or decision process of any of the cognitive parts making up the network. In fact, when a decentralized network is tasked with solving a complex problem (like coloring any two-dimensional map using only four colors such that no two adjacent regions have the same color), adding global information to the cognitively enabled parts can sometimes hinder rather than help the network in achieving a solution (Kearns, Suri, & Montfort, 2006). Given that the decisions and solutions made by the higher level system do not always conform to those made by (any of) its cognitive parts, it is not satisfactory to simply explain the higher level decisions in terms of the cognitive capabilities of the parts.

But when we consider human collectives on much larger spatial and temporal scales, there are good reasons to doubt the generality of this zero-sum perspective. Again, there are situations for which we do not challenge its validity, and

we would be remiss not to mention them at the outset. Most evidently, this is the case whenever the intelligent behavior of a collective can be directly explained with reference to the psychological states of one or more smart individuals who are in charge of its operations, such as military commands issued by an army general or strategic decisions made by the CEO of a corporation. Here we derive no explanatory payoff from treating the army or company as a cognitive system in its own right. In this respect, we agree with the spirit of Huebner's first principle (or constraint) that we should not posit group level cognition "where collective behavior results from an organizational structure set up to achieve the goals or realize the intentions of a few powerful and/or intelligent people" (2013, p. 21). In practice, the formulation of this constraint may be too coarse to capture the flow of information in hybrid organizational structures that blend centralized forms of top-down management with distributed, bottom-up forms of mass collaboration such as "commons-based peer production" (Benkler, 2002), for instance wikis, or open-source projects. Crowdsourcing may be a hybrid model of this kind, whose distinctive feature is to leverage the collective intelligence of online communities for specific organizational goals set by management (Brabham, 2013). Thus, there is a sense in which the locus of creativity regarding the production of goods or ideas seems to reside somewhere "in between" organizational decision-makers and participant communities.

A zero-sum perspective loses its plausibility when we hope to understand the dynamics of large-scale social systems that contain multiple people and surrounding supports that facilitate their interactions (Stahl, 2006, 2009; Theiner, Allen, & Goldstone, 2010). To show this, our argument proceeds as follows. First, we articulate our "non-zero sum" position that part of what makes people smart is that they can form groups with the ability to behave in collectively intelligent ways, which in turn increases their potential to become even smarter as individuals. Then we consider Huebner's two other injunctions against positing group cognition and argue that they do not apply to the cognitive models on which our framework is based. As part of our discussion, we stress the non-anthropocentric character of our cognitive systems perspective.

4.2. What groups do people make?

People are strongly disposed toward creating groups that constitute their identity. One of the primary directions toward which people aim their cognitive efforts is creating robust communication channels, building lasting social organizations, and assuring the persistence of the groups to which they belong. Over time, groups of various sizes and purposes create rule and norms systems (e.g., laws) that self-organize the interactions in the group, monitors (e.g., police) to look for potential violations of the rules, arbiters (e.g., judges) to determine whether violations have occurred, and sanctions (e.g., prisons) to punish people found guilty of violating the rules (Ostrom, 2009). In this manner, people often build complex

and multi-faceted structures to help scaffold their interactions. Clark (2003) has described ways in which humans apply their powers of thought to create tools that improve those same powers of thought. This positive feedback loop is nowhere more apparent than in the *social* tools that we create to allow us to coordinate with, connect to, learn from, and create teams with others. In addition to social institutions like governments, courts, and laws, other examples of social tools include cell phones, scholarly journals, the World Wide Web, schools, and pubs. These tools connect us together, and once connected together, one of our immediate impulses is to look for ways to strengthen and augment those connections. For this reason, when it comes to individual people and the groups they form, there is often times not a zero-sum relation between the flexible problem-solving capacities of the two levels of systems.

Research from psychology provides strong support for a "smart people can make smart groups" argument. First, people have a fundamental drive to participate in close groups (Baumeister & Leary, 1995). When social bonds are broken, people look for substitutes. People in close relationships are healthier, happier, live longer, and have more resilient immune systems. Infants form close attachments even before they have any experience with their broader social world. People confuse themselves with others when they are connected to those others in close relationships. For example, when a person is asked to quickly decide whether a trait is true or not true of themselves, they are significantly slowed down if they differ in the trait from their partner (Aron, Aron, Tudor, & Nelson, 1991). This is consistent with people adopting close others into their own self-identity. Interestingly, people often even have difficulty knowing whether an event happened to themselves or a close other. For example, twins frequently have confusions, which they did not even know about before the experiment, regarding which twin had an experience (Sheen, Kemp, & Rubin, 2001). Typically, each twins think that *he or she* is the one who experienced the event, such as coming in 12th in an international cross country race or falling off a tractor and spraining a wrist, and the memory is usually from the first-person perspective. Even when the individual ownership of a memory is not disputed, people frequently possess vicarious memories in which they have recollections of salient life episodes that were told to them by another person, such as a friend or family member (Pillemer et al., 2015). These vicarious memories are often vivid, emotional, and experienced from a first-person perspective.

Other examples of people assimilating others into their identity come from groups that extend beyond intimate and close dyadic relations (Smith, Seger, & Mackie, 2007). When a person strongly identifies with a group, he or she experiences pride when that group does well and anger when the group is attacked, even when the person is not directly involved (Yzerbyt, Dumont, Gordijn, & Wigboldus, 2002). Individuals feel self-regarding emotions for events that happened before they were even born, for example when contemporary Caucasian-Americans feel

guilt about the catastrophic decimation of Native American populations or the enslavement of Africans (Doosje, Branscombe, Spears, & Manstead, 1998).

The above examples of individuals identifying with, and sometimes confusing themselves with, other people could be construed as cognitive and emotional phenomena at the individual level. The memory distortions, guilt, and pride are experienced, after all, by individual people. However, the point of the examples has thus far simply been to show that people are often highly motivated to participate in strong groups larger than themselves, and so, to the extent that they are smart, they will be able to form cohesive groups with strong structural integrity. Individual ants and bees also work toward the welfare of their colonies even to their own individual detriment, but because these insects do not have the cognitive wherewithal that humans do, there is reason to believe that humans will be able to find ways to create smart groups even in adverse conditions that would stymie social insects. After an apocalypse, there would be a push for the scattered remains of humanity to rapidly band together using whatever technologies remain for the purpose of reestablishing social connection, and if none remained, many would immediately begin to rebuild them. Of course, smart people can also act in self-serving ways that undermine cohesive groups, and social power structures can frustrate efforts of a group to self-organize in effective ways. Our point is only that most people have an inherent drive to connect to others, and individual cognitive flexibility often finds ways to support those connections in rich and resilient ways by creating communicative infrastructures, norms, and conventions.

Resuming our cognitive systems perspective, then, what are the likely characteristics of the group cognitive systems that are recreated in the post-apocalyptic world? Cognitively competent people would likely design the systems to be efficient in information processing, flexible, robust, and persistent, much like our current social structures. Many of the mechanisms reviewed above are instrumental in achieving these characteristics of groups. For example, vicarious and appropriated memories serve the purpose of letting the members of a group learn, vividly retain, and pass along information from episodes they were not the first to experience (Theiner, 2013). When a twin tells her twin sister about being in a car accident while texting and the second sister appropriates the memory as her own, she will be more influenced by the cautionary moral of the story, with the result that she will probably drive more safely and tell others about the dangers of texting. A similar group-level functional account can be given for the role of gossip in a community, which allows many people to gain access to valuable information, particularly about deviance from social norms, that would be impossible for them all to know directly (Levine & Smith, 2013). Likewise, people's strong emotions relating to group pride, guilt, and anger incite them to act in ways that the will reinforce, repair, and protect the group, respectively.

4.3. When can groups of people cognize?

We now return to the question of how our "inclusive" non-zero sum perspective fares with respect to the more stringent injunctions or "exclusion" principles against group cognition imposed by Huebner. Discussing termite nest construction as a canonical instance of self-organizing collective behavior, he suggests—as his second constraint—that we should not posit collective cognition "where collective behavior bubbles up from simple rules governing the behavior of individuals" (2013, p. 23). Since many of our models of collective cognitive systems fall into a similar category, it is important to clarify exactly how our position differs from that of Huebner. Upon closer inspection, we find that for Huebner the second constraint gets its bite not so much from a general metaphysical skepticism about emergence but from his commitment to a particular representational criterion of cognition. This can be gleaned from a comparison between two putative cases of group cognition which, for Huebner, come down on different sides of the fence erected by the second constraint.

The first type of case, which for Huebner does *not* qualify as an instance of macrocognition, is the aforementioned case of termite nest construction. Let us pause to review Huebner's reasoning for why that is the case (2013, p. 22–23). We certainly agree with Huebner's characterization that in order to build a wonderful termite mound, it is not necessary that any of the termites knows how to build one (see Dennett, 2009), and that the complex structure of the mound cannot in any straightforward manner be derived from the simple rules governing the behavior, including mental representations, of individual termites. We also share his view that the colony as a whole has no need to represent anything about the structure of the nest for the mound to come into existence. Based on these premises, Huebner then argues that if we can construct a complete "bubble-up" model of emergent collective behavior that refers only to *mechanisms* governing individual behavior, and *mechanisms* for aggregating these behaviors, we have effectively explained away group cognition. The same argument can be used, according to Huebner (2013, p. 23, note 4), to account for the trail formation capacities of foraging ants without ascribing cognition to the hive. Importantly, Huebner's conclusion follows only if we require—as Huebner does—that macrocognition must involve, or eventuate in, the creation of genuinely collective representations that modulate and coordinate the behavior of individual agents.

This crucial extra premise explains the contrast between termite mound construction (and relevantly similar cases, like ant foraging) and the goal-directed behavior of honeybee colonies, which for Huebner qualify as "minimally" minded macrocognitive systems (2013, p. 230–233). Huebner's argument here is that even though individual bees are capable of encoding information about things like the location of nectar, the quality of foraging sites, and the location and quality of nest sites, no individual bee could aggregate this information in the manner which allows the colony as a whole to make sophisticated judgments about

the comparative value of foraging sites and nest sites. Here, Huebner is equally impressed (as are we) with the computational equivalence between competitive "winner-take-all" mechanisms in the colony and similar mechanisms that occur in individual brains, on the grounds that they both generate genuinely cognitive representations of the environment. Provided that the criteria for positing such representations are met, Huebner argues that we should not adopt a zero-sum perspective.

Reflecting on the putative contrast between these two cases of emergent self-organizing behavior, the difference between our perspective and Huebner's view is not so much a dispute about inter-level exclusion principles as it is a disagreement over the central role which collective mental representations play in Huebner's account.

An example of groups of humans solving a problem that does not seem to require collective mental representation is collective path formation. As paths created by trudging students after a campus snowfall show, groups often create path systems that are mutually advantageous to the members of a group without being intended by any of the members. In laboratory experiments in which people are incentivized to take advantage of the trails left by their predecessors, we find that there are conditions in which groups come to close to optimally solving the problem of how to connect a set of destinations using the minimal total amount of trail (Goldstone, Jones, & Roberts, 2006). The produced path systems are solutions to a problem for the group, not the individuals. The individuals, at least those specified by our computational model (based on Helbing, Keltsch, & Molnár, 1997), are simply taking steps that compromise between going where they want to go and where others have gone before. The group's problem is different—creating an efficient path system for all members. Changes to the group process, for example increasing the built-in obsolence of paths, can increase the efficiency of the path system, and groups that stumble upon changes like this will create better path systems. Groups that can, via random or systematic variation, search through a parameter space involving group/environment factors like path half-life, step impact, and path diffusion can develop increasingly good path systems without engaging in explicit representation building, or having its members even realize that the group is trying to solve a path length minimization problem.

Finally, as a segue to Huebner's third and final constraint, we would at least like to speculate how the MILCS framework might be reconciled with an analysis of groups as intentional systems (Clark, 1994; Dennett, 1987; Tollefsen, 2015). Taking up the "intentional stance" is a shorthand for our widespread interpretive practice of ascribing what philosophers call "propositional attitudes" (such as beliefs, desires, and intentions) to people in order to predict and explain their behavior, but also to make sense of one another as agents having a rational point of view. There can be little doubt that we regularly and successfully extend this practice to things other than individual human beings, such as computers, non-human animals, and specifically groups. In common parlance as well as social scientific

research, we often talk about what a company intends to do, what the market desires, or what a court of law judges illegal. Within philosophy, Gilbert (1989, 2013) has long argued for the collective intentionality of groups conceived as plural subjects. Drawing on formal results in the theory of judgment aggregation, List and Pettit (2011) offer a far-ranging account of the metaphysics, organizational design, and normative status of groups as corporate agents. Rooted in a cognitive systems perspective, Dedeo (2014) presents a game-theoretic analysis of the dynamic interplay between cooperative and conflicting attitudes among Wikipedia editors and shows that the simplest explanation of large-scale trends can be obtained by attributing aggregate expectations ("beliefs") and preferences ("desires") to the group as a whole.

Huebner embraces the usefulness of the intentional stance as a starting point, as an indispensable tool to get a grip on the kinds of cognitive functions that creatures with minds perform, but denies that it bears any cognitive-scientific fruit unless wedded to a detailed theory of the mechanisms by which these functions are implemented (2013, chapter 2). But for Huebner, not just any mechanism will do. The collective intentionality of groups has to be produced by a massively distributed, but sufficiently integrated computational architecture that eventuates in the construction and manipulation of complex, collective mental representations (chapter 4). Thus, in his discussion of the stock market as exhibiting the "wisdom of crowds" (Surowiecki, 2004), Huebner concedes that the decentralized "buy-and-sell" interactions among traders in a stock market implement a "winner-take-all" computation analogous to what is found in competitive neural networks (2013, p. 69), or the collective selection of nest sites by honeybees (p. 68, note 11). However, due to a lack of genuinely collective mental representations, he denies that there is any substance to the claim that the stock market is an intentional (or cognitive) system in its own right, for example, judging the value of certain companies or desiring certain outcomes over others. In this context, he then adds his third constraint that we should not posit group cognition if

> the capacities of the components belong to the same intentional kind as the capacity that is being ascribed to the collectivity and where the collective computations are no more sophisticated than the computations that are carried out by the individuals who compose the collectivity. (p. 72)

Since our approach allows, but does not require, that group cognition involve the creation and manipulation of collective mental representations, Huebner's analysis seems to puts us into the following double-bind. On the one hand, we could completely disavow any attachment to the intentional stance as a proto-scientific theory or model of collective behavior, despite its considerable predictive success; or, we could drop the rather sensible requirement that the intentional stance must be amenable to a deeper, mechanistic understanding of cognitive systems. Neither alternative is prima facie attractive to us. Fortunately, there is a way to go between the horns of this dilemma. By adopting a measurement-theoretic understanding of propositional attitude attributions (Matthews, 2010), it is possible to salvage a

broadly functionalist construal of intentional systems theory without incurring an explicit commitment to a certain kind of computationalist cognitive architecture (à la Huebner) to sustain it. Moreover, the modestly realist understanding of propositional attitude ascriptions that it underwrites, and that can in principle be applied to individuals as well as groups, also reveals why we reject Huebner's third constraint as too demanding. Here, we can offer only the vaguest of sketches about how such a reconciliation might proceed.

The basic intuition behind a measurement-theoretic approach is to compare the practice of attributing propositions to the subjects of intentional mental states to our practice of assigning numerical values to objects having a certain weight. For example, it encourages us to think of the role played by the sentence "Otto believes that the interest rates will fall" by analogy with that of "Otto has a mass of 80 kilo." On this view, to ascribe beliefs and desires to a person can be thought of as a procedure for locating her in a linguistically defined measurement space that indexes semantically and pragmatically evaluable information about her causally efficacious aptitudes for thought and behavior. Concerning different types of attitudes, one could say—to a very rough approximation—that "beliefs" are environment-shaped states, and "desires" are environment-shaping states; and for more fine-grained explanations to be possible, different kinds of beliefs and desires must be individuated in terms of the specific states of the environment which they shape or are shaped by, which is achieved by ascribing propositional contents to the attitudes in question (Matthews, 1994, p. 139). Importantly, the success of this practice does not require that the subject of an attitude literally stand in a psychological relation to a mental representation that directly or indirectly corresponds to the structure of its propositional content. In fact, it does not require that there are any mental representations at all! Instead, it fundamentally hinges only on the satisfaction of two basic measurement-theoretic properties. First, the selected measurement system must preserve the properties and relations of the empirical system that is to be measured (the "representation problem"). In the case of weight, this would involve arithmetical relations (defined over the real numbers) such as "x has the same mass as y" or "x has twice the mass of y," and so on. In the case of beliefs and desires, the analogs are inferential relations (defined over semantic contents) that represent a subject as being in a psychological state that is apt to have certain characteristic effects. This includes purely logical inferences (such as "If a person *S* believes that *P*, and also believes that $P \rightarrow Q$, then *S* also believes that *Q*"), but also modes of practical reasoning (such as "If A desires that *P* and believes that doing *A* is an effective way to achieve *P*, *S* will likely try to do *A*"). Since the first principle can in principle be accomplished by any number of such measurement systems, the second condition is to specify, for a given scale, the set of admissible transformations to other scales which preserve the empirical properties and relations that it represents (the "uniqueness problem"). For example, empirical relations that are measured on an interval scale (e.g., temperature) are preserved under linear transformations of the scale. Likewise, in the case of

belief, contextually interpreted belief reports couched in different languages ("*S* believes that Inga is a septuagenarian," "*S* glaubt dass Inga in ihren Siebzigern ist") are mapped onto the same point in propositional attitude space if and only if they count as admissible transformations of one another (for a detailed articulation of this view, see Matthews, 2010).

If a measurement-theoretic approach to propositional attitude ascriptions can be defended (which we have not done here), it has several consequences that we welcome from our position. First, it serves to decouple our understanding of intentional systems theory from the more stringent demands on computational architecture imposed by Huebner. Thus, if we can "fully and voluminously" predict the collective behavior of groups by adopting the intentional stance, we do not have to be "dyed-in the-wool instrumentalists" (Huebner, 2013, p. 37) to justify its empirical adequacy. Instead, we can assert in a moderately realist tone of voice that the intentional stance works *because* it tracks psychologically relevant internal structures of cognitive systems. The dog is still wagging his tail. Again, it is important to point out that a measurement-theoretic approach to propositional attitude ascriptions is consistent with, but does not require, a computational theory of mind (Matthews, 2010). All that is required to preserve the psychological properties and relations of the cognitive system to be measured is that the intentional stance be a *homomorphic* image of its empirical domain. Suppose we discovered, for example, a cognitive system which implements a classical computational architecture that operates on syntactically structured and semantically evaluable mental representations, akin to a "language of thought" (Fodor, 1975, 1987). With respect to that system, for any propositional attitude ascription, it is true that the psychological subject stands in a corresponding computational relation to an internal sentence tokened in the language of thought. Such a neat coincidence would be theoretically convenient because we could directly read many critical properties of the measurement system—for example, constituent structure or inferential relations among propositions—back into the system's inner psychology. For a computational system of this sort, which is "semantically transparent" (Clark, 1989, p. 18), the empirical psychological domain and the linguistic measuring system would be virtually isomorphic. Since every isomorphism automatically preserves a homomorphic image, the language of thought clearly counts as an ("industrial strength") solution of the measurement problem. Of course, Huebner (2013, chapter 4) rejects—wisely, in our view—the requirement of semantic transparency in favor of a "kludgy" cognitive architecture that is implemented by massively distributed, highly integrated, specialized processing systems, and is thus "computationally opaque" from the standpoint of propositional attitude ascriptions. Still, such an architecture is not inconsistent with a measurement-theoretic approach to propositional attitudes, provided that a homomorphic image of the target domain is preserved. But contrary to Huebner, and in line with our more promiscuous approach, we do not want to prejudge for

which kinds of systems the intentional stance is appropriate. Hence, we prefer to remain non-committal about the exact mechanisms by which it is implemented.

Second, the measurement-theoretic approach puts the strategy of taking the intentional stance towards individual human beings on a par with its extension to groups of people, without necessarily implying that the kinds of intentional mental states that we attribute to groups are equally fine-grained. It is surely unlikely that we can measure the psychological life of groups using the same rich, nuanced, and intricately interwoven network of intentional predicates by which people make sense of one another. This does not mean that the predicates we use cannot be represented within the same measurement space, but the remaining challenge here is to define structural correspondences between individual and group attitudes of similar types. This last point also makes it clear why we beg to differ with Huebner's third constraint. Even if it is true, as Huebner alludes, that *qua* intentional systems, groups are less complex than individuals, the measurement-theoretic approach we prefer encourages a non-zero sum perspective on intentional systems theory. A helpful analogy to keep in mind here is the transition from classical to statistical mechanics. Statistical mechanics can explain the macroscopic properties of a gas (i.e., temperature, pressure, and distribution of kinetic energies) without being able to, nor having to, predict where any particular molecule will be in the future. Similarly, if the intentional stance can account for psychological "bulk" properties of physical systems that we could not deduce from the behavior of the system's components, its adoption is justified on any given level where this is the case, even if those levels are nested within each other.

5. Conclusions

The MILCS perspective on understanding the cognitive capacities of groups of people focuses on understanding cognition in terms of systems dynamics. Certain kinds of systems are able to achieve specific functionalities that are relevant for cognitive processes involved in perception, attention, memory, problem-solving, and decision-making. These systems offer mechanisms for achieving cognitively important functions such as selecting an option that is most attractive among its competitors, achieving consensus in a network so that the network acts in a coherent manner, selectively attending to information globally assessed as relevant, developing specialized units from originally homogeneous networks through competition among the units, and interactive networks for combining data from perception with top-down theories and models. We call these systems "cognitive systems" not because they appear *only* in things that we intuitively think of as thinking. Instead, we call these systems "cognitive systems" because they instantiate cognitive mechanisms in virtue of which a variety of systems perform important cognitive functions associated with flexible, adaptive, and intelligent behavior, and what we have argued is that those are also the systems which allow those thinking things to think.

We suspect that some of our colleagues will find this treatment of group cognition to be a deflationary disappointment. Nowhere have we talked about mental representations or languages of thought, let alone group minds or global brains. However, one hopefully offsetting advantage of the MILCS framing is that these kinds of systems are performing the kinds of cognitive functions that cognitive scientists are actively trying to understand. For example, scientists are trying to figure out whether the rate of evidence accumulation is independent of the current evidence or not, whether and how the output from a perceptual processing module is affected by information available outside of the module, and how groups can be structured to find excellent solutions most effectively. The answers to questions like these depend on developing and empirically validating models of system dynamics in networks of interacting units. A second advantage is that these systems can be comfortably implemented using parts and interactions among parts that are well understood. We do not know how to go about building minds or perhaps even to recognize when we have built one, but we can make effective progress in cobbling together networks of neurons, words, families, scientists, and employees, and shaping how they influence each other to achieve cognitively useful functions. A third advantage is that these cognitive systems can be described in a unified language of networks, nodes, edges, inhibition, excitation, and message passing. When expressed in this language, we see that the systems principles apply across many scales and domains and yet are inductively powerful. Despite dissimilarities in scale (e.g., neurons versus employees) and domain, different networks that are aptly describable by the same system dynamic share theoretically powerful commonalities that can be exploited to make accurate predictions about behavior at multiple levels. While this account may miss out on some of the excitement that extended and group minds inspire, there is nonetheless a certain cohesive grandeur to understanding society as consisting of infrastructural networks that connect social networks of neural networks.

Notes

1. Following Rupert's own treatment, we leave it open here whether his account is intended as providing not only necessary but also sufficient conditions of cognition (Rupert, 2010). As noted by Rupert, Hutchins (1995, p. 157) similarly suggests individuating cognitive systems by the gradient of computational activity; parts belong to the same cognitive system to the extent that they are jointly engaged in considerably more computations than elements at the periphery. For a detailed discussion of Rupert's "systemic" mark of the mental, see Theiner (2011), in particular chapter 6.2.
2. For an alternative systems-based account, see Weiskopf (2010).
3. We thank Robert Rupert for drawing our attention to this important point.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Adams, F., & Aizawa, K. (2008). *The bounds of cognition*. Malden, MA: Blackwell.

Anderson, M. L., Richardson, M. J., & Chemero, A. (2012). Eroding the boundaries of cognition: Implications of embodiment. *Topics in Cognitive Science, 4*, 717–730.

Aron, A., Aron, E. N., Tudor, M., & Nelson, G. (1991). Close relationships as including other in the self. *Journal of Personality and Social Psychology, 60*, 241–253.

Ball, P. (1999). *The self-made tapestry*. Oxford: Oxford University Press.

Ball, P. (2004). *Critical mass: How one thing leads to another*. New York, NY: Farrar, Straus & Giroux.

Barabasi, A. L., & Albert, R. (1999). Emergence of scaling in random networks. *Science, 286*, 509–512.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117*, 497–529.

Bechtel, W. (2008). *Mental mechanisms: Philosophical perspectives on cognitive neuroscience*. London: Routledge.

Bechtel, W. (2009). Looking down, around, and up: Mechanistic explanation in psychology. *Philosophical Psychology, 22*, 543–564.

Beckner, C., Blythe, R., Bybee, J., Christiansen, M. H., Croft, W., Ellis, N. C., … Schoenemann, T. (2009). Language is a complex adaptive system: Position paper. *Language Learning, 59*, 1–26.

Benkler, Y. (2002). Coase's penguin, or, Linux and the nature of the firm. *Yale Law Journal, 112*, 369–446.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., & Hwang, D. U. (2006). Complex networks: Structure and dynamics. *Physics Reports, 424*, 175–308.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced choice tasks. *Psychological Review, 113*, 700–765.

Brabham, D. C. (2013). *Crowdsourcing*. Cambridge, MA: MIT Press.

Bullmore, E., & Sporns, O. (2009). Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience, 10*, 186–198.

Burgess, C., & Simpson, G. B. (1988). Cerebral hemispheric mechanisms in the retrieval of ambiguous word meanings. *Brain and Language, 33*, 86–103.

Clark, A. (1989). *Microcognition: Philosophy, cognitive science, and parallel distributed processing*. Cambridge, MA: MIT Press.

Clark, A. (1994). Beliefs and desires incorporated. *Journal of Philosophy, 91*, 404–425.

Clark, A. (2001). Reasons, robots and the extended mind. *Mind and Language, 16*, 121–145.

Clark, A. (2003). *Natural born cyborgs*. Oxford: Oxford University Press.

Clark, A. (2008). *Supersizing the mind: Embodiment, action, and cognitive extension*. Oxford University Press.

Clark, A. (2010). Much ado about cognition. *Mind, 119*, 1047–1066.

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis, 58*, 7–19.

Craver, C. F., & Bechtel, W. (2006). Mechanism. In S. Sarkar & J. Pfeifer (Eds.), *Philosophy of science: An encyclopedia* (pp. 469–478). New York, NY: Routledge.

Darwin, C. (1859). *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life, nature* (Full image view 1st ed.). London: John Murray, 5 (121): 502.

Dedeo, S. (2014). Group minds and the case of Wikipedia. *Human Computation, 1*, 5–29.
Dennett, D. (2009). Darwin's "strange inversion of reasoning". *PNAS, 106*(Suppl. 1), 10061–10065.
Dennett, D. C. (1987). *The intentional stance*. Cambridge, MA: MIT Press.
Derex, M., & Boyd, R. (2016). Partial connectivity increases cultural accumulation within groups. *PNAS, 113*, 2982–2987.
Doosje, B., Branscombe, N. R., Spears, R., & Manstead, A. S. R. (1998). Guilty by association: When one's group has a negative history. *Journal of Personality and Social Psychology, 75*, 872–886.
Dyson, F. (2004). A meeting with Enrico Fermi. *Nature, 427*, 297.
Einstein, A. (1905). Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen [On the movement of small particles suspended in a stationary liquid demanded by the molecular-kinetic theory of heat]. *Annalen der Physik, 322*, 549–560.
Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in uncertain environments. *Psychological Review, 100*, 432–459.
Figdor, C. (in press). *Pieces of mind*. Oxford University Press.
Fodor, J. (1975). *The language of thought*. Cambridge, MA: Harvard University Press.
Fodor, J. (1987). *Psychosemantics: The problem of meaning in the philosophy of mind*. Cambridge, MA: MIT Press.
Forstmann, B. U., Ratcliff, R., & Wagenmakers, E. J. (2016). Sequential sampling models in cognitive neuroscience: Advantages, applications, and extensions. *Annual Review of Psychology, 67*, 641–666.
Gilbert, M. (1989). *On social facts*. London: Routledge.
Gilbert, M. (2013). *Joint commitment: How we make the social world*. New York, NY: Oxford University Press.
Girvan, M., & Newman, M. E. J. (2002). Community structure in neural and social networks. *PNAS, 99*, 7821–7826.
Gold, I. J., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience, 30*, 535–574.
Goldstone, R. L., & Gureckis, T. M. (2009). Collective behavior. *Topics in Cognitive Science, 1*, 412–438.
Goldstone, R. L., & Hendrickson, A. T. (2010). Categorical perception. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*, 69–78.
Goldstone, R. L., Jones, A., & Roberts, M. E. (2006). Group path formation. *IEEE Transactions on System, Man, and Cybernetics, Part A, 36*, 611–620.
Goldstone, R. L., Wisdom, T. N., Roberts, M. E., & Frey, S. (2013). Learning along with others. *Psychology of Learning and Motivation, 58*, 1–45.
Goldstone, R. L., & Wilensky, U. (2008). Promoting transfer by grounding complex systems principles. *The Journal of the Learning Sciences, 17*, 465–516.
Grossberg, S. (1976). Adaptive pattern classification and universal recoding I: Parallel development and coding of neural feature detectors. *Biological Cybernetics, 23*, 121–134.
Harnad, S. (1987). *Categorical perception*. Cambridge: Cambridge University Press.
Haugeland, J. (1998). *Having thought: Essays in the metaphysics of mind*. Cambridge, MA: Harvard University Press.
Heersmink, R. (2015). Dimensions of integration in embedded and extended cognitive systems. *Phenomenology and the Cognitive Sciences, 14*, 577–598.
Helbing, D., Keltsch, J., & Molnár, P. (1997). Modelling the evolution of human trail systems. *Nature, 388*, 47–50.
Huebner, B. (2013). *Macrocognition*. New York, NY: Oxford University Press.

Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: MIT Press.
Inoue, S., & Matsuzawa, T. (2007). Working memory of numerals in chimpanzees. *Current Biology, 17*, R1004–R1005.
Kao, A., Miller, N., Torney, C., Hartnett, A., & Couzin, I. D. (2014). Collective learning and optimal consensus in animal groups. *PLoS Computational Biology, 10*, e1003762.
Kearns, M., Suri, S., & Montfort, N. (2006). An experimental study of the coloring problem on human subject networks. *Science, 313*, 824–827.
Latty, T., & Beekman, M. (2011). Speed-accuracy trade-offs during foraging decisions in the acellular slime mould *Physarum polycephalum*. *Philosophical Transactions of the Royal Society B, 278*, 539–545.
Lazer, D., & Friedman, A. (2007). The network structure of exploration and exploitation. *Administrative Science Quarterly, 52*, 667–694.
Levine, J. M., & Smith, E. R. (2013). Group cognition: Collective information search and distribution. In D. E. Carlston (Ed.), *Oxford handbook of social cognition* (pp. 616–633). New York, NY: Oxford University Press.
Liebowitz, S. J. (1995). Path dependence, lock-in, and history. *Journal of Law Economics & Organization, 11*, 205–226.
List, C., Elsholtz, C., & Seeley, T. D. (2008). Independence and interdependence in collective decision making: An agent-based model of nest-site choice by honeybee swarms. *Philosophical Transactions of the Royal Society B, 364*, 755–762.
List, C., & Pettit, P. (2011). *Group agency: The possibility, design, and status of corporate agents*. Oxford: Oxford University Press.
Machamer, P. K., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science, 67*(1), 1–25.
Marshall, J. A. R., Bogacz, R., Dornhaus, A., Plaque, R., Kovacs, T., & Franks, N. R. (2011). On optimal decision-making in brains and social insect colonies. *Journal of the Royal Society Interface, 6*, 1065–1074.
Mason, W. A., Jones, A., & Goldstone, R. L. (2008). Propagation of innovations in networked groups. *Journal of Experimental Psychology: General, 137*, 422–433.
Matthews, R. J. (1994). The measure of mind. *Mind, 103*, 131–146.
Matthews, R. J. (2010). *The measure of mind: Propositional attitudes and their attribution*. Oxford: Oxford University Press.
Mayer, A. R., Hanlon, F. M., Franco, A. R., Teshiba, T. M., Thoma, R. J., Clark, V. P., & Canive, J. M. (2009). The neural networks underlying auditory sensory gating. *NeuroImage, 44*, 182–189.
McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part I. An account of basic findings. *Psychological Review, 88*, 375–407.
Mulder, M. J., Wagenmakers, E. J., Ratcliff, R., Boekel, W., & Forstmann, B. U. (2012). Bias in the brain: A diffusionmodel analysis of prior probability and potential payoff. *The Journal of Neuroscience, 32*, 2335–2343.
Ostrom, E. (2009). A general framework for analyzing sustainability of social-ecological systems. *Science, 325*, 419–422.
Pillemer, D. B., Steiner, K. L., Kuwabara, K. J., Kierkegaard Thomsen, D., & Svor, C. (2015). Vicarious memories. *Consciousness and Cognition, 36*, 233–245.
Ratcliff, R. (2002). A diffusion model account of response time and accuracy in a brightness discrimination task: Fitting real data and failing to fit fake but plausible data. *Psychonomic Bulletin and Review, 9*, 278–291.
Rumelhart, D. E., & Zipser, D. (1985). Feature discovery by competitive learning. *Cognitive Science, 9*, 75–112.
Rupert, R. (2009). *Cognitive systems and the extended mind*. Oxford: Oxford University Press.

Rupert, R. (2010). Extended cognition and the priority of cognitive systems. *Cognitive Systems Research, 11*, 343–356.

Rupert, R. D. (2011). Empirical arguments for group minds: A critical appraisal. *Philosophy Compass, 6*, 630–639.

Sheen, M., Kemp, S., & Rubin, D. (2001). Twins dispute memory ownership: A new false memory phenomenon. *Memory & Cognition, 29*, 779–788.

Sloan Wilson, D., & Wilson, E. O. (2007). Rethinking the theoretical foundation of sociobiology. *The Quarterly Review of Biology, 82*, 327–348.

Smith, E. R., Seger, C. R., & Mackie, D. M. (2007). Can emotions be truly group-level? Evidence regarding four conceptual criteria. *Journal of Personality and Social Psychology, 93*, 431–446.

Sporns, O., Tononi, G., & Edelman, G. M. (2000). Connectivity and complexity: The relationship between neuroanatomy and brain dynamics. *Neural Networks, 13*, 909–922.

Stahl, G. (2006). *Group cognition: Computer support for building collaborative knowledge*. Cambridge, MA: MIT Press.

Stahl, G. (2009). *Studying virtual math teams*. New York, NY: Springer Verlag.

Sterelny, K. (2010). Minds: Extended or scaffolded? *Phenomenology and the Cognitive Sciences, 9*, 465–481.

Surowiecki, J. (2004). *The wisdom of crowds*. New York, NY: Doubleday.

Thagard, P. (2014). The self as a system of multilevel interacting mechanisms. *Philosophical Psychology, 27*, 145–163.

Theiner, G. (2011). *Res cogitans extensa*. Frankfurt: Peter Lang.

Theiner, G. (2013). Transactive memory systems: A mechanistic analysis of emergent group memory. *Review of Philosophy and Psychology, 4*, 65–89.

Theiner, G., Allen, C., & Goldstone, R. L. (2010). Recognizing group cognition. *Cognitive Systems Research, 11*, 378–395.

Tollefsen, D. (2015). *Groups as agents*. Boston, MA: Polity Press.

Tononi, G., Edelman, G. M., & Sporns, O. (1998). Complexity and coherency: Integrating information in the brain. *Trends in Cognitive Sciences, 12*, 474–484.

Toon, A. (2014). Friends at last? Distributed cognition and the cognitive/social divide. *Philosophical Psychology, 27*(1), 112–125.

Weick, K. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly, 21*, 1–9.

Weiskopf, D. A. (2010). The Goldilocks problem and extended cognition. *Cognitive Systems Research, 11*, 313–323.

Weisberg, M. (2013). *Simulation and similarity: Using models to understand the world*. New York, NY: Oxford University Press.

Yzerbyt, V. Y., Dumont, M., Gordijn, E., & Wigboldus, D. (2002). Intergroup emotions and self-categorization: The impact of perspective taking on reactions to victims of harmful behavior. In D. M. Mackie & E. R. Smith (Eds.), *From prejudice to intergroup emotions* (pp. 67–88). New York, NY: Psychology Press.

Zollman, K. J. (2010). The epistemic benefit of transient diversity. *Erkenntnis, 72*, 17–35.