

The Group Mind In Commonsense Psychology

Bryce Huebner, Georgetown University

In the late 19th century, powerful labor unions, anarchist collectives, and radical workers' parties were struggling against the unfair demands of powerful corporate agents. *Psychological* forces were commonly posited to explain the behavior of these groups, and it was often suggested that the stability of social norms revealed 'collective representations', or even the workings of a 'group mind' (Durkheim 1982, 1995; Freud 1990; Jung 2001; Le Bon 2002; McDougall 2005; MacKay 1980; Spencer 1897). But while the group mind hypothesis quickly fell into disrepute (Allport 1924), discussions of collective intelligence, maddening crowds, and group minds have persisted in the popular imagination (Gladwell 2000; Surowiecki 2004). Once again, we live in an age of popular uprisings, Occupy movements, and failing political parties; appeals to corporate personhood are deeply entrenched in the American legal system; and it is hard to ignore the consolidation of state and corporate power that lurks behind many forms of collective action. None of this provides evidence that there are group minds, and there is reason to think that most appeals to collective mentality are explanatorily superfluous (Rupert 2005; Huebner 2013). But it remains unclear how discussions of collective intelligence and corporate power are interpreted by non-experts.

In this chapter, I examine recent work in psychology and experimental philosophy that has targeted the commonsense understanding of group minds. I begin, in Section 1, by setting up the conceptual and empirical terrain on which claims about the group mind in commonsense psychology have been constructed. In Section 2, I then consider experiments based on explicit judgments, and experiments examining the tendency to see groups as unified entities. And in Section 3, I turn to an analysis of the cross-cultural data, which suggest a greater willingness to ascribe collective mentality in East Asian cultures. I conclude my initial discussion of the empirical data in Section 4, where I turn to research programs in social psychology that have attempted to distinguish groups that are more and less entity-like. Finally, in Section 5, I argue that these different strands of data together support the claim that commonsense psychology is committed to a form of methodological holism, which allows for properties of individuals that critically depend on group membership; and I argue that these data provide weak support at best for the hypothesis that groups are commonly seen as possessing minds of their own. But, I conclude by suggesting that this does not cut against philosophical and scientific attempts to defend the existence of collective intentions and group minds, it merely suggests that such projects must move beyond commonsense strategies for predicting and explaining *individual* behavior.

1. Do we think about group minds?

Let's begin a case where people accept discussions of group minds without a second thought. The Borg of *Star Trek*, the bugs of *Starship troopers*, the Formics of *Ender's game*, and the Precogs of "Minority report" all think and act as *groups*. These are fictional entities, and they are embedded in narratives that help us to map commonsense assumptions onto new possibilities (Fauconnier and Turner 1998, 2003). And it is unclear how people interpret claims about 'group egos' that cannot die until every body they inhabit are destroyed (Heinlein 1941). Do they read them as a literal claims about group minds, or do they treat them as literary flourish? A person who has never heard of tachyons may accept that *something* makes cloaking possible, but she is unlikely to think too much about what fills the 'tachyon' role. By contrast, an expert who

already has beliefs about tachyons will be more likely to see claims about cloaking devices as literary flourish based on false or ill-formed views about what tachyons are. The reason why claims about group *minds* are interesting is that most people are in a position much closer to our imagined expert. By the time we can interpret science fiction scenarios, we already have some beliefs about other minds, and some beliefs about the mental states necessary for intentional action. So we must ask: do we intuitively assume that only individuals have minds, or do our commonsense assumptions have a wider berth?

In one of the most familiar experiments in cognitive psychology, Fritz Heider and Marry-Ann Simmel (1944) asked participants to describe what happened in a short animation where geometric shapes moved systematically around a screen. Most explanations made reference to the purposeful actions of intentional agents. While it would be surprising if these people really believed that the circles and triangles were genuine agents, these data suggest that they found it easy to adopt a perspective from which the purposeful behavior could be described in intentional terms. Far more strikingly, people in such tasks find it difficult to override their initial sense that apparently purposive behavior call for intentional explanations—and they continue to use intentional language even when they are explicitly told not to do so (Dittrich and Lea 1994). So, it is worth noting that people also describe animations where groups of objects move together in a coherent and organized way by reference the intentional states *of groups* (Bloom and Veres 1999, B4). This doesn't establish much, but it may suggest that people find it easy to *see agency* in apparently purposeful group behavior (or it may only suggest that people are prolific users of non-literal language, a possibility to which I return below).

Beyond this tendency to use intentional explanations to describe observed behavior, we find that discussions of collective action are pervasive in the news: “Israel keeps a quiet eye on turmoil in Egypt” (Kershner 2013); “Iceland aims to seize opportunities in oil exploration” (Gardner 2013); and “Google, a pioneer of self-driving cars, is quietly lobbying for legislation that would make Nevada the first state where they could be legally operated on public roads” (Markoff 2011). On a deflationary reading of such headlines, we only speak as-if there are group agents to highlight the absence of individual agency or to highlight the fact that we don't really know who the agents are. But people are inclined to ascribe mental states wherever they find purposive behavior, and they opt for intentional explanations of even potentially purposeful behavior (Rosset 2008). So perhaps people reflexively posit group minds when they encounter purposeful behavior in the absence of obvious individual agency.

Indeed, recent neuroimaging data suggest that people rely on the same neural systems to interpret individual and group behavior, and these systems are equally active when people draw inferences about individual and group mental states (Contreras et al. 2013). In both cases, people seem to rely on mentalizing systems. But the data are inconclusive, for while “the same brain regions represent the mental states of groups and individuals, most of them carry information about whether the target of the mental state inference is a group or an individual” (Contreras et al. 2013, 1415). So these data leave open the possibility that we see groups as agents (though as different kinds of agents than individuals), as well as the possibility that we see individuals as agents and rapidly infer their mental states from facts about group membership. They do show that mentalizing is occurring, but they do not show where it is occurring. Consequently, investigating any more inflationary hypothesis will require looking to other sources of data.

2. Explicit judgments about group minds

There is broad consensus that we appeal to mental states to explain intentional action, and that we appeal to causal histories to explain non-intentional behavior (Malle 1999; Malle and Knobe 1997). Our explanations of intentional actions typically make reference to an agent's reasons for acting, and they require an agent who is aware of what they are doing, capable of acting

rationally, and capable of exerting intentional control over their actions. By contrast, explanations of unintentional behavior are typically framed in quasi-mechanical terms, and they require nothing more than causal forces that can bring about an outcome (Malle 2010). So if people posit group minds, they should sometimes explain group behavior by appealing to a *group's* reasons for acting; but if people are individualists by default, their explanations of group behavior should always make reference to the causal history of a group or the mental states of group members. Unsurprisingly, people are more likely to appeal to causal histories to explain group behavior. However, they do sometimes appeal to a *group's* reasons for acting, and they are more likely to do so when the group is thought to be *deliberatively unified*. (O'Laughlin and Malle 2002).

Of course, this is precisely what would be predicted by philosophers working on collective intentionality; but as the arguments in this literature show, the fact that a group's actions are rationally intelligible cannot establish that the group is an agent with a mind of its own (Gilbert, 1987, 1989, 1996; List & Pettit 2011; Tollefsen, 2002; Velleman, 1997). At best, the fact that people sometimes appeal to a group's reasons for acting suggests that people think well-organized groups can form and act on shared intentions, in the weak sense that they all have similar intentions *qua group members*.¹ It takes a lot of philosophical work to get from claims about shared intentions in this weak sense, to arguments for the existence of group mentality. In part, this is because it's not obvious what could make these sorts of shared intentions count as the *mental* states of a group. Indeed, it has seemed clear to many philosophers of mind that groups lack the kind of subjective experiences or conscious agency that is necessary for genuine mentality (Block 1978). But then again, non-experts seem to be far more willing to acknowledge the possibility that there are agents who lack subjective experience (Gray, Gray, and Wegner 2007)

Against this backdrop, Joshua Knobe and Jesse Prinz (2008) asked people to evaluate sentences that ascribed mental states to the ACME Corporation; their participants thought that sentences ascribing beliefs and desires to groups sounded 'natural', but that sentences ascribing feelings and other phenomenal states to groups sounded 'weird'. But the interpretation of these data is complicated by the fact that carefully paired sentences do not reveal significant differences in the acceptability of sentences like 'Microsoft is feeling depressed' and 'Microsoft is depressed' (Sytsma and Machery 2009). Furthermore, the inclusion of contextual information seems to attenuate the impact of the 'feels' locution on the acceptability of sentences ascribing mental states to groups (e.g., "Microsoft feels sad when it loses customers"; Arico 2007). These data seem to suggest that people accept some explicit ascriptions of group mental states. But the fact that these sentences all sound acceptable does not, on its own, discriminate between interpretations on which these sentences are seen as literally true, and interpretations of these sentences on which they are seen as metaphorical or literary flourish.

To examine this question, Adam Arico and his colleagues (2010) asked participants to categorize sentences as making figurative or literal claims. They included several sentences about *individual* mental states (e.g., "Some millionaires want tax cuts"), several sentences about *group* mental states (e.g., "Some corporations want tax cuts"), and several sentences that served as a manipulation check (e.g., "Einstein was an egghead"; "Carpenters build houses"). Much to their surprise, they found that people tended to interpret sentences about group mentality literally—and the effect was not moderated by excluding people who were bad at recognizing figurative language from the analysis. But while this is a striking result, it cannot settle the issue either. After all, participants may interpret these sentences literally, but treat them as claims about the individuals who compose the groups under consideration. Plural noun-phrases can be read collectively, as claims about groups, or distributively, as claims about the individuals that compose groups. For example, a collective reading of "Iceland aims to seize opportunities in oil exploration" has the goals of *Iceland* as its subject matter; but a distributive reading treats it as a claim about the goals of Icelandic citizens, or members of the Icelandic

government—and this reading is made true by the fact that these individuals are making decisions that lead to oil exploration by Icelandic citizens, or Icelandic-owned business.

To examine whether individuals interpret sentences referencing the mental states of groups collectively or distributively, Mark Phelan and his colleagues (2013) asked participants to replace the second occurrence of a group name in a sentence like “After Boeing lost the Army contract, Boeing expected to have to lay off workers” with either 'it' or 'they' (and they treat this decision as evidence that people prefer collective or distributive readings respectively). Their participants chose to use 'they' just over half of the time; and in a carefully structured follow-up task, where participants were asked to choose between two paraphrases of a sentence, people once again chose the distributed reading just over half of the time. Phelan and his colleagues argue that these data show that we typically interpret sentences ascribing mentality to groups as claims about the mental states that individuals have *qua* members of groups. But even if they are right that this is how we interpret such sentences, their data do not provide conclusive support for this conclusion.

To begin with, their participants *sometimes* preferred collective readings of the sentences. Unfortunately, Phelan and his colleagues did not examine the extent to which these differences tracked perceived differences in deliberative unity. In this case, it is unclear whether the people who chose a collective reading over a deliberative reading saw Boeing as more deliberatively unified. Some people may be inclined to assume that a corporation is a loose aggregate, while others may be more willing to see a corporation as a unified and persisting entity. So it would be worth doing a follow-up study in light of existing data and philosophical arguments. To my ear, it seems plausible to say of a deliberatively unified philosophy department that *it* was hasty in making a hiring decision, or of a deliberatively unified corporation that *it* made a bad choice in outsourcing manufacturing. Of course, it doesn't follow that I thereby treat these groups as *entities*, who are proper targets of intentional explanation. Phelan and his colleagues may be right that even when I adopt the collective reading, I may still be using the group-term as instrumentally useful shorthand to refer to things that the members of a group *do together*. I doubt that judgments about the acceptability of particular sentences will be able to discriminate between instrumental and realistic uses of intentional ascriptions. The converging evidence discussed in the next section may help to ameliorate these worries. But whether they do is a complicated matter to which I return in the final section of this chapter.

3. Cultural differences in the interpretation of collective behavior?

A second approach to the study of group minds in commonsense psychology builds on insights from gestalt psychology, where it was often claimed that people perceive objects as persisting *entities* when they have clear perceptual boundaries and a high degree of internal structure. Donald Campbell (1958) famously argued that we also perceive groups as entity-like when their members are psychologically *similar* and *deliberatively unified* in ways that allow them to carry out coordinated actions. This sits well with the fact that people appeal to a group's reasons for acting where it is seen as deliberatively unified (O'Laughlin and Malle 2002). And perhaps this is what lies behind Carson McCullers (1943) description of a group of men whose thoughts converge so perfectly that “all together they will act in unison, not from thought or from the will of any one man, but as though their instincts had merged together so that the decision belongs to no single one of them, but to the group as a whole.” Perhaps it is also the guiding insight behind Herbert Spencer's (cited in Campbell 1958) claim that we treat *society* as an organism “because, though formed of discrete units, a certain concreteness in the aggregate of them is implied by the general persistence of the arrangements among them throughout the area occupied”. When we observe the darting and diving of a flock of birds, it is hard to escape the assumption that *its* goal directed behavior could emerge, were it not under the control of a single mind.

For those raised in modern Western societies, it may seem obvious that individuals must have ontological priority over the groups to which they belong, and it may seem obvious that group behavior can always be explained by reference to the beliefs, desires, and intentions of individual group members. But as Hazel Markus and Shinobu Kitayama (1991) argue, this may not be so obvious to East Asians who have been raised to interpret individual behavior in relational and context-dependent terms. If East Asians typically see both groups and individuals as locations in social space (Ames 1994; Hall and Ames, 1998; Rosemont, 1991), then perhaps they will also be less inclined to see group behavior as an expression of individual mental states.

Of course, some mental states only make sense as the mental states of individuals. Only individuals have biologically-based desires for food, drink, and sex, but it is unclear how far this gets us. Yoshihisa Kashima and his colleagues (2005) also report that people across cultures see individuals as more entity-like than groups when they focus on internal structure and the unalterability of behavioral traits. And they report that people across cultures assume that individual beliefs and attitudes are sometimes constrained by group membership, just as the data from Phelan and his colleagues (2013) suggests. But where people focus on deliberately structured mental states like beliefs, desires, and intentions, it seems that individuals are only perceived as more entity-like than groups in the West; East Asians seem more willing to treat groups as unified agents with the capacity to act intentionally (Kashima et al. 2005, 162). In line with this hypothesis, and building on an analysis of newspaper headlines containing ascriptions of mental states, Tanya Menon and her colleagues (1999, 702) contend that while “prevailing American theories hold that persons have stable properties that cause social outcomes and groups do not, the theories prevailing in Confucian influenced East Asian cultures emphasize that groups have stable properties that cause social outcomes.”

These data are consistent with the possibility that cultural differences modulate the perception of deliberative unity, and yield an increased willingness to treat groups as unified entities with minds of their own. So another set of researchers examined the effect of these cultural differences on explicit judgments about group mental states (Huebner, Sarkissian, and Bruno 2009). They presented sentences ascribing mental states to individuals and groups to participants in both North Carolina and Hong Kong. In both cases, participants saw the ascription of intentional states as more acceptable than the ascription of phenomenal states to groups—but the difference was significantly reduced in participants from Hong Kong. Unfortunately, our data were hard to interpret. These differences hovered around the midpoint of the scale, and they had wide standard deviations. Furthermore, this study was conducted in English, in a westernized city. So the Huebner and colleagues were apprehensive about the interpretation of the data, and conducted a follow-up study in Shanghai, using Mandarin-language stimuli verified using a back-translation technique. Here, Boston-based participants agreed that a corporation could be happy, but were reticent about the claim that a corporation could feel upset; by contrast, participants in Shanghai showed the opposite effect! They seemed more willing to see a corporation as a locus of experience. This result is perplexing, to say the least.

Huebner and his colleagues initially argued that cultural factors modulate the willingness to ascribe mental states to groups, and they suggested that this was because cultural difference in perceived entitativity played an important role in how participants conceptualized the groups under consideration. While these data are consistent with this hypothesis, they do not rule out the confounding factors I noted in the previous section. Specifically, the authors did not address the worries about distributive and collective readings of plural noun-phrases—and in the case of these cross cultural data, such problems are likely to be even more pronounced—after all, there are substantial difficulties associated with the interpretation of generics and plurals in Mandarin (Sybesma and Cheng 1999). So many more experiments will be necessary to get a clear sense of the effect of cultural factors on the ascription of mentality. These experiments will have to be

carried in a way that is sensitive to the linguistic and cultural issues that are relevant to the ascription of group mental states—at this point it is hard to predict how the data will turn out.

4. Are groups entity-like

While some philosopher are likely to be familiar with the research I have discussed thus far, most seem to be less familiar with research on impression formation and groups agency; and in this section I turn to this literature to flesh out the hypothesis that people are intuitive holists. Psychologists have long known that we see people as agents with whom we can reason and collaborate, and that we rapidly form impressions of them, and use these impressions to draw inferences about their mental lives. Indeed, we tend to generalize from initial impressions in ways that ignore or discount the subsequent information we learn. Tasks requiring evaluative inferences about the socially-salient traits of an individual tends to reveal a *primacy effect*, with the information that is learned first dominating judgments. But similar tasks targeting groups tend to reveal a *recency effect*, with the information presented most recently dominating people's judgments (Manis and Paskewitz 1987). This suggests that people rely on *working memory representations* when they make judgments about groups, whereas they seem to rely on reflexive strategies of impression formation in explaining individual behavior. This is why people are more successful at recalling dispositional traits ascribed to individuals than at recalling dispositional traits ascribed to groups; and it is why people are more accurate and faster in remembering socially significant information about individuals (McConnell, Sherman, and Hamilton 1997). These data suggest that we do not perceive groups as agents that have stable psychological dispositions, and that we have to work to remember what we have learned about a group.

Across a broad range of tasks, people see individuals as possessing dispositional traits, and their explanations of individual behavior display a sensitivity to dispositional information in a way that explanations of group behavior do not. For example, Joshua Susskind and his colleagues (1999) primed participants with expectations about a social target, and asked them to complete sentences containing partial descriptions of expectancy-consistent and expectancy-inconsistent behavior; they found that people completed descriptions of expectancy-consistent behavior by referencing an individual's reasons for acting, while they completed descriptions of expectancy-inconsistent behavior by appealing to causal factors that could resolve the inconsistency. But people didn't attempt to resolve the inconsistencies in the same way when the target was a group. With groups, the inferences were insensitive to the supposed expectations about the group's dispositional properties. Such data support for the hypothesis that socially salient information about groups is not encoded in ways that facilitate spontaneous dispositional inferences—that is, they suggest that we do not see groups as having dispositional traits of the sort that individuals are thought to possess.

Of course, there are distinctions between different types of groups. Beyond differences in deliberative unity, groups are distinguished by member similarity, the presence of shared goals, the significance of intra-group interactions, and the ease of entry to and exit from a group, among other things (Lickel et al. 2000); and in general, as Campbell (1958) argued, “the more that group members spend time together, want the same things, share a similar fate, exhibit similar characteristics, and consider the group to be an important part of life, the more entitative that group is perceived to be” (Sherman and Percy 2010, 153). And in line with the data that I addressed in the previous section, there is reason to suppose that people will form impressions of groups that they perceive as coherent entities—or at least that they will assume that the members of such groups will share essential properties that make inferences about their unobserved properties seem reasonable (McConnell, Sherman, and Hamilton 1997).

When individuals ask challenging and esoteric questions in a quiz task, they are seen as more intelligent than those who fail to answer these questions, even where the assigned role is

known to be arbitrary (Ross, Amabile, and Steinmetz 1977). A similar effect occurs when groups are seen as unified agents: a group that is designated as the Answer-group will be rated as less knowledgeable than a Question-group when both groups are thought to be integrated, stable, and long-lasting (Yzerbyt, Rogier, and Fiske 1998). But the judgments here are judgments about the knowledge of the group members as individuals. Similarly, people assume that a trait that is displayed by one member of a group will be displayed by the other members, when the group is seen as a persisting entity (e.g., when the target is a group of close friends who have spent a lot of time together and who feel very connected; Crawford, Sherman, and Hamilton 2002). And, the members of groups that are perceived as entity-like are often thought to share psychological characteristics (Dasgupta, Banaji, and Abelson 1999). In each case, the data suggest that individual traits are often thought to depend on group membership—which is exactly what would be predicted by Phelan and his colleagues (2013).

When we learn that a group is unified, this reinforces the search for shared properties that keep group members together, and this leads us to suppose that group members will have similar observed and unobserved properties; likewise, when we think that group members are similar to one another, this lead us to assume that the group will exhibit stable properties, in part because the group's behavior becomes predictable in light of similarities between the group members (Yzerbyt, Corneille, and Estrada 2001; Yzerbyt, Rocher, and Schadron 1997). Thus, when a group is seen as highly agent-like, "its members are expected to behave in a more consistent manner, they are thought to be more similar to one another, they are categorized in a more undifferentiated way at the group level, and the discrepant members are assimilated to rather than contrasted from the prototype of the group" (Yzerbyt et al. 1998, 1092). This is especially salient in light of the heuristic strategies we employ in making judgments about stereotype-groups. Here, we assume that the members of a group will share essential qualities, and we focus on these similarities even in the face of differences between members. But this familiar fact generalizes, and commonsense psychology seem to allows for inferences about individual mental states on the basis of our assumptions about group membership wherever we see groups as stable and long lasting.

5. What do the existing data show?

Even the most deflationary approaches to social ontology are likely to recognize that there are important differences between deliberately unified groups, and loose aggregates that happen to act together. John Searle (1990), for example, considers the difference between a bunch of people in a park running toward the same shelter to avoid rain, and a group that displays the same movements as part of an avant garde dance performance. He notes that the intentionality of the dance performance provides no reason to posit a group mind—it is fully explicable in terms of the group-relevant attitudes of individual dancers. Despite deep differences in their approaches to social ontology, most people working on collective behavior acknowledge that these sorts of *shared intentions* play a critical role in coordinating activity when people decide to do things together; and it is commonly agreed that these kinds of shared intentions constrain our actions and attitudes because they cannot be rescinded unilaterally (Bratman 1993; Gilbert 1989). So it is nice to see that the data in psychology and experimental philosophy confirm that non-experts have a similar perspective; they recognize that people often form beliefs, desires, and intentions *qua members of groups* that they would not form independently of their group-relevant commitments (Gilbert 1987; Pettit 1996; Phelan, Arico, and Nichols 2013). But this is where things get complicated.

Once we have recognized the existence of individual mental states that are only manifested socially, there is little reason to assume that most deliberately unified groups act on reasons that are properly ascribed to the group *as such*. So it is surprising that inflationary claims about group minds and collective intentions persist in philosophy and social psychology. Margaret

Gilbert (1989, 18) argues that some groups constitute 'plural subjects', which are properly understood as targets of intentional explanation; but her arguments depend on the socially manifested attitudes that individuals take up *qua* group members: "One is willing to be the member of a plural subject if one is willing, at least in relation to certain conditions, to put one's own will into a 'pool of wills' dedicated, as one, to a single goal (or whatever it is that the pool is dedicated to)". Similarly, Christian List and Phillip Pettit (2011) deploy resources from social choice theory to demonstrate that individual and group decisions sometimes come apart, as they do when a court's decisions diverge from the decisions that each member prefers. But here too, "the construction of the court's opinion, its legitimacy, and its effect on society can all be explained without invoking anything beyond the conservative ontology of individuals and their states" (Rupert 2005, 6). Finally, the psychological data discussed in the last section are often thought to support exciting claims about our ability to track the dispositional character traits of deliberatively unified groups. In a recent review, Steven Sherman and Elise Percy (2010, 151) argue that "when a group is perceived as entitative, all the principles that apply to impression formation for individuals apply to the group as well". But the data only show that people sometimes see group members as sharing essential traits, and then draw inferences about group members and future behavior on the basis of their essentialist assumptions.

In each case, the inflationary claim about group minds requires downplaying the extent to which individual traits are expressed in the context of group membership (Pettit 1996; Phelan, Arico, and Nichols 2013; Wilson 2001). There is little evidence that commonsense psychology downplays these commitments. In *most* cases predictions about group behavior, whether they are made by experts or non-experts, depend on assumptions about the mental lives of group members, *qua* group members (cf., Huebner 2013). But as I noted above, there are some situations where people treat deliberatively unified groups as proper targets of intentional explanation, and there are sometimes situations where they explain the behavior of such groups by appealing to a *group's* reasons for acting. Such data speak more directly to the group mind hypothesis, especially since people sometimes interpret claims about group mental states literally, and do so in a way that relies on some of the same neural systems that are active when they interpret the behavior of individuals. But we must tread cautiously in interpreting these data. These data cannot distinguish between the hypothesis that group minds have a place in commonsense psychology, and the hypothesis that group agency is posited as instrumental shorthand when the mental states of group-members are unknown.

The neuroscientific data reveal that different sources of information are used in interpreting individual and group behavior, but they do not establish what these differences amount to (Contreras et al. 2013). As I noted briefly above, ascriptions of mental states to groups do tend to evoke activity in regions that are often active in mentalizing tasks targeting individuals. However, However, a more targeted analysis revealed that these regions were differentially responsive to groups and to individual people. This suggests that people in these tasks are representing groups and individuals differently during tasks that require mental state inference, but it is hard to know whether this is a difference that makes a difference relative to person-level behavior. That said, the data on explicit judgments outside of the scanner do provide some guidance on this issue. They suggest that people *often* prefer individualistic interpretations of collective behavior when they are available, though they don't do so in every case. At this point, the most that we can say is that there is a lot more work to do in determining whether people ever posit group minds—and I hope that experimental philosophers and their allies will continue to collect the relevant data. But even when the data are in, I doubt that they will answer all of the interesting questions about the plausibility of group minds, and I would like to close by explaining why this is the case.

Decades of research in developmental psychology have shown that we have the capacity to *detect* and *attribute* agency on the basis of minimal cues, in ways that reveal an approximation of the 'intentional stance' (Csibra and Gergely 2013 and Arico et al. 2011 for reviews).² People

also seem to adopt the intentional stance to explain even the movement of animated groups of objects (Bloom and Veres 1999). But the reliance on the intentional stance to explain these sorts of behavior tells us little about the commonsense ontology of minds. After all, as Daniel Dennett (1989, 2) has long noted, our use of the intentional stance to explain goal-directed behavior reveals a “motley assortment of 'serious' belief attributions, dubious belief attributions, pedagogically useful metaphors, *façons de parler*, and perhaps worse, outright frauds”. Dennett has never been sanguine about the possibility of pulling apart these different uses of intentional language, but I think it is clear that people do not see the behavior of circles and squares as evidence of genuine mindedness; at best, these experiments lead them to use cognitive systems that evolved to track goal-directed behavior in a derivative way. But no matter how broadly these systems can be applied, they are not much use in tracking or explaining most kinds of group behavior. We rarely (perhaps never) *observe* group behavior *as such*. When a crowd is rioting, we see individuals setting things on fire, breaking window, and carrying TVs out of stores. When a labor union strikes, we see union members and their allies on the streets. So even if we reflexively detect agency in purposeful behavior, the systems that make this possible are unlikely to be recruited to make sense of group behavior outside of the lab.

So what are we to make of the fact that people routinely interpret the behavior of courts, churches, states, hiring committees, and corporations in ways that invoke intentional ascriptions? I maintain that in these case, they rely on more abstract interpretive strategies, and that the data from psychology and experimental philosophy support the claim that commonsense psychology has a broadly functionalist character: people willingly ascribe beliefs and desires to human and non-human animals, robots, supernatural agents, and groups (Arico et al. 2011; Gray et al., 2007; Haslam et al. 2008; Huebner 2010; Knobe & Prinz 2008; Phelan and Buckwalter 2012; Sytsma and Machery 2009). In these cases, people do seem to adopt the intentional stance, treating rational intelligibility as evidence that a social target has beliefs and desires. But the use of intentional idioms does not show that people are committed to treating *groups as agents*, or as possessing beliefs and desires. Indeed, the use of such idioms is perfectly compatible with the possibility that they are using the group as a useful stand-in that allows them to make predictions about future behavior on the fly.

Even if we make relatively reliable predictions about group behavior by positing mental states, this shouldn't persuade anyone who isn't already a dyed in the wool behaviorist or instrumentalist about mentality that there are group minds—and I doubt that non-experts have such commitments. Adopting the intentional stance in a way that is useful in the long run requires being able to answer the question, “what is it about an entity that makes it rational for *that entity* to behave in a particular way?” And framing the question in this way makes it possible to see how a complex nexus of causal relations 'hangs together' as a unified, intelligent, and intelligible agent (Hornsby 1997).

Consider an example where the intentional stance works well. I treat my cat, Nutmeg, as an intentional system, and doing so helps me to explain her behavior with a high degree of accuracy. In her case, this doesn't take much. I suppose that Nutmeg has beliefs about where her food and treats are kept, that she has desires for food, treats, and affection, and that she is afraid of vacuum cleaners and street sweepers. I know that each such description is anthropomorphic, and that the language I employ is somewhat misguided. But at the same time, I find it difficult to resist adopting an intentional stance, and her ongoing goal-directed behavior is rationally explicable (even if it is best explained by appeal to reinforcement learning systems rather than more complex states like beliefs and desires). Here is the upshot. Because Nutmeg is a biological entity, things matter to her, and this allows me to ask what makes it rational *for her* to behave in the ways that she does. Because I care about her wellbeing, and because I work to satisfy the needs and interests that I attribute to her, I am forced to treat her as a complex nexus of causal relations that hangs together as a unified, intelligent, and intelligible agent. But perhaps more interestingly, my assumptions about her needs and interests can be

calibrated against the underlying processes that make her behavior rationally intelligible; and at the same time, my interpretations can shape her behavior in ways that help to make her a plausible target of intentional ascriptions (cf., Zawidzki 2013). But things are different in the case of group behavior.

Philosophical claims about the implausibility of group minds often depend on intuitions about the impossibility of subjective experiences in a group (Block 1978). I cannot defend this possibility in detail here, but I believe that part of the reason for this is that we have a hard time imagining that things *matter* to corporations (and we have a fairly easy time imagining that things *matter* to the individuals that compose those corporations). There is some evidence that our willingness to treat an entity as morally considerable modulates our willingness to see it as minded (Robbins 2008; Robbins and Jack 2006). I contend that this is likely to be part of the reason why it is so hard for us to hold on to the possibility that groups can be minded in any interesting sense when we reflect on our use of mentalistic language to describe group behavior. A person shows no moral failing if she disassembles a chess-playing computer or deactivates a robot she has built; but a person who disassembles her housemate or deactivates her neighbor is morally deranged (Dennett, 1976). This is because housemates and neighbors have their own concerns and projects, and we make a moral mistake if we are not concerned with them. While it would take more experimental work to show this, anecdotal evidence suggests that the lack of subjective experience is part of what makes it hard for people to treat artificial mentality as a real possibility. While there is room for dispute over which entities care how things go for them, my guess is that most people do not see groups as having concerns of their own, above and beyond the concerns of the individuals that compose them. People don't tend to see it as objectionable to frustrate *a group's* goals without permission, even though they may see it as problematic to harm the individuals that compose a group. For example, people may see a hostile takeover as morally acceptable even when it severely modifies the deliberative structure of a corporation, and they may even see it as morally preferable if it benefits individual group members; but people will be more likely to see it as morally objectionable to carry out a hostile takeover of an individual that severely modifies her deliberative structure, no matter how it affects various neural assemblages, or so I predict.

The fact that a group cannot give a damn about how things go for it makes it hard for us to treat a group as a minded entity. But it is unclear whether even the most robust intuitions of this sort are useful for addressing ontological questions about minds. Perhaps research in the psychology and experimental philosophy of group minds simply reveals interesting failures of imagination (Dennett 1988; Huebner 2013). To demonstrate that something of ontological import follows from even the most widely shared intuition, philosophical arguments must be marshaled. Even if most people do in fact reject the possibility of group minds, it will not follow that there are no group minds—that is a matter to be settled by philosophers and cognitive scientists. This is why debates in philosophical action theory must push beyond claims about the rational intelligibility of group behavior to demonstrate that their appeals to plural subjects and collective intentions are not explanatorily superfluous. This is also the reason why cognitive scientists interested in collective memory and collective decision making need to make it clear how their data support claims about group behavior as such, as opposed to claims about how individuals behave *qua* group members. These are difficult issues, and they force us to stretch our intuitions beyond the threshold of common sense. Debates over the possibility of group mentality in the philosophy of mind and cognitive science turn on issues of cognitive integration, and they call for architectural arguments that depend on specialized tools from psychology and cognitive science. I maintain that arguments about the existence of group minds are revisionary in nature, and that they require re-thinking what it is to think. That said, I think it would be quite surprising if claims about collective mentality didn't conflict with commonsense intuition (Huebner 2013).

Works Cited:

- Allport, Floyd. 1924. *Social Psychology*. Boston: Houghton Mifflin
- Ames, Roger. 1994. "Reflections on the Confucian self" in M. Bockover (ed), *Rules, Rituals, and Responsibility*, LaSalle: Open Court.
- Arico, Adam. 2010. "Folk psychology, consciousness, and context effects." *Review of Philosophy and Psychology*, 1(3), 371–393.
- Arico, Adam, Brian Fiala, Robert Goldberg, and Shaun Nichols. 2011. The Folk Psychology of Consciousness. *Mind and Language*, 26, 3, 327-352.
- Block, Ned. (1978). "Troubles with functionalism," in C.W. Savage (ed.), *Minnesota Studies in the Philosophy of Science*, 9 (261–325). Minneapolis: University of Minnesota Press.
- Bloom, Paul and Csaba Veres. 1999. "The perceived intentionality of groups," *Cognition*, 71, B1–B9
- Bratman, Michael. 1993. "Shared intention," *Ethics*, 104: 97-113.
- Campbell, Donald T. 1958. "Common fate, similarity, and other indices of the status of aggregates of persons as social entities," *Behavioural Sciences*, 3, 14–25.
- Contreras, Juan, Jessica Schirmer, Mahzarin Banaji, and Jason Mitchell. 2013. "Common brain regions with distinct patterns of neural responses during mentalizing about groups and individuals." *Journal of cognitive neuroscience* 25, 9: 1406-1417.
- Crawford, Matthew, Steven Sherman, and David Hamilton. 2002. "Perceived entitativity, stereotype formation, and the interchangeability of group members." *Journal of Personality and Social Psychology*, 83, 1076-1094.
- Csibra Gergeley and György Gergely. 2013 "Teleological understanding of actions." In Mahzarin Banaji and Susan Gelman (eds). *Navigating the Social World: What Infants, Children, and Other Species Can Teach Us*. Oxford: Oxford University Press: 38-43.
- Dasgupta, Nilanjana, Mahzarin Banaji, and Robert Abelson. "Group entitativity and group perception: Associations between physical features and psychological judgment." *Journal of Personality and Social Psychology*, 77, 991-1003.
- Dennett, Daniel C. 1976. "Conditions of Personhood," in A. Rorty (ed) *The identities of persons*. Los Angeles: University of California Press.
- Dennett, Daniel C. 1988. "Quining Qualia," in A. Marcel and E. Bisiach (eds), *Consciousness in Modern Science*, Oxford: Oxford University Press, 42-77.
- Dennett, Daniel C. 1989. "True believers," in *The Intentional Stance*. Cambridge: MIT Press, 13-36.
- Dittrich, W. & S. Lea (1994). "Visual perception of intentional motions," *Perception*, 23, 253-268.
- Durkheim, Emile. 1982. *The rules of sociological method*. S. Lukes (ed) and W.D. Halls (trans). London: Free Press.
- Durkheim, Emile. 1995. *The elementary forms of religious life*. K. Fields (trans). London: Free Press.
- Fauconnier, Giles and Mark Turner. 1998. "Conceptual Integration Networks," *Cognitive science*, 22(2), 133-187.
- Fauconnier, Giles and Mark Turner. 2003. *The Way We Think*. New York: Basic Books.
- Freud, Sigmund. 1990. *Group psychology and the analysis of the ego*. J. Strachey (trans). London: W. W. Norton & Company.
- Gardner, Beth. 2013. "Iceland aims to seize opportunities in oil exploration." *New York Times*, 1 October 2013, <http://www.nytimes.com/2013/10/02/business/energy-environment/iceland-aims-to-seize-opportunities-in-oil-exploration.html> (accessed 10 January 2014).
- Gilbert, Margaret. 1987. "Modeling Collective Belief," *Synthese*, 73, 185-204
- Gilbert, Margaret. 1989. *On Social Facts*. New York: Routledge.

- Gilbert, Margaret. 1996. *Living together: rationality, sociality, and obligation*. New York: Rowman and Littlefield.
- Gladwell, Malcolm. 2000. *The tipping point*. Boston: Back Bay Books.
- Gray, Heather, Kurt Gray, and Dan Wegner. 2007. "Dimensions of mind perception," *Science*, 619, 315.
- Hall, David & Roger Ames. 1998. *Thinking from the Han: Self, Truth and Transcendence in Chinese and Western Culture*. Albany, SUNY Press.
- Haslam, Nick, Yoshihisa Kashima, Stephen Loughnan, Junqi Shi, and Caterina Suitner. 2008. "Subhuman, inhuman, and superhuman: contrasting humans and nonhumans in three cultures," *Social Cognition*, 26, 248–258.
- Heider, Fritz and Marianne Simmel. 1944. An experimental study of apparent behavior. *American Journal of Psychology*, 57, 243–259.
- Heinlein, Robert. 1941. *Methuselah's Children*. New York: Gnome Press.
- Hornsby, Jennifer. 1997. *Simple Mindedness: A Defense of Naïve Naturalism in the Philosophy of Mind*. Cambridge: Harvard University Press.
- Huebner, Bryce. 2010. "Commonsense concepts of phenomenal consciousness: Does anyone care about functional zombies?" *Phenomenology and the cognitive sciences*, 9, 133-155.
- Huebner, Bryce. 2013. *Macrocognition: A theory of distributed minds and collective intentionality*. New York: Oxford University Press.
- Huebner, Bryce, Hagop Sarkissian, and Micheal Bruno. 2009. "What does the Nation of China think about phenomenal states?" *Review of philosophy and psychology*, 1, 2, 225-243.
- Jung, Carl. 2001. *Psychology of the unconscious*. Princeton: Princeton University Press.
- Kashima, Yoshihisa, Emiko Kashima, Chi-Yue Chiu, Thomas Farsides, Michele Gelfand, Ying-Yi Hong, Uichol Kim, et al. 2005. "Culture, essentialism, and agency," *European Journal of Social Psychology* 35, 147-169.
- Kershner, Isabel (2013). Israel keeps a quiet eye on turmoil in Egypt. *New York Times*, 16 August 2013, http://www.nytimes.com/2013/08/17/world/middleeast/israel-turmoil-in-egypt.html?_r=0 (accessed 10 January 2014).
- Knobe, Joshua and Jesse Prinz. 2008. "Intuitions about consciousness: experimental studies," *Phenomenology and the cognitive sciences*, 7, 67–85.
- Le Bon, Gustave. 2002. *The crowd: A study of the popular mind*. London: Dover Press
- Lickel, Brian, David Hamilton, Grazyna Wieczorkowska, Amy Lewis, Steven Sherman, and A. Neville Uhles. 2000. "Varieties of groups and the perception of group entitativity," *Journal of Personality and Social Psychology*, 78, 223-246.
- List, Christian and Phillip Pettit. 2011. *Group Agency: The Possibility, Design, and Status of Corporate Agents*. New York: Oxford University Press.
- Malle, Bertram. 1999. "How people explain behavior: A new theoretical framework," *Personality and Social Psychology Review*, 3, 1, 23-48.
- Malle, Bertram. 2010. Intentional action in folk psychology. In T. O'Connor and C. Sandis (Eds.), *A companion to the philosophy of action* (pp. 357-365). Chichester, UK: Wiley-Blackwell.
- Malle, Bertram and Joshua Knobe (1997). "Which behaviors do people explain? A basic actor-observer asymmetry," *Journal of Personality and Social Psychology*, 72, 288-304.
- Manis, Melvin and Joan Paskewitz. 1987. "Assessing psychopathology in individuals and groups," *Personality and Social Psychology Bulletin*, 13, 83-94.
- Markoff, John. 2011. "Google Lobbies Nevada to Allow Self-Driving Cars," *New York Times*, 10 May 2011, <http://www.nytimes.com/2011/05/11/science/11drive.html> (Accessed 10 January 2014).
- Markus, Hazel and Shinobu Kitayama. 1991. "Culture and the self," *Psychological Review*, 98, 224-253.

- McConnell, Allen, Steven Sherman, and David Hamilton. 1997. "Target entitativity: Implications for information processing about individual and group targets," *Journal of Personality and Social Psychology*, 72, 750–762.
- McCullers, Carson. 1943. "The ballad of the sad café," *Harper's Bazaar*, 77, 140-61.
- McDougall, William. 2005. *The group mind*. Whitefish, MT: Kessinger Publishing.
- MacKay, Charles. 1980. *Extraordinary popular delusions and the madness of crowds*. New York: Harmony Books.
- Menon, Tanya, Michael Morris, Chi-yue Chiu, and Ying-yi Hong. 1999. "Culture and the construal of agency," *Journal of Personality and Social Psychology*, 76, 701-717.
- O'Laughlin, Matthew and Bertram Malle. 2002. "How people explain actions performed by groups and individuals," *Journal of Personality and Social Psychology*, 82, 33-48.
- Pettit, Phillip. 1996. *The common mind*. Oxford: Oxford University Press.
- Phelan, Mark, Adam Arico and Shaun Nichols. 2013. Thinking Things and Feeling Things: On an Alleged Discontinuity in Folk Metaphysics of mind. *Phenomenology and the Cognitive Sciences*, 12: 4: 703-725
- Phelan, Mark and Wesley Buckwalter (2012). Analytic Functionalism and Mental State Attribution. *Philosophical Topics*, 40 (2): 129-154.
- Robbins, Phillip. 2008. "Consciousness and the social mind," *Cognitive Systems Research*, 9, 15–23.
- Robbins, Phillip and Anthony Jack. 2006. "The phenomenal stance," *Philosophical studies*, 127, 59–85.
- Rosemont, Henry. 1991. "Rights-bearing individuals and role-bearing persons," In Henry Rosemont and Mary Bockover (eds). *Rules, Rituals, and Responsibility*. LaSalle: Open Court.
- Ross, Lee, Teresa Amabile, and Julia Steinmetz. 1977. "Social roles, social control, and biases in social perception," *Journal of Personality and Social Psychology*, 35, 485-494.
- Rosset, Evelyn. 2008. "It's no accident" *Cognition*, 108, 771-780.
- Rupert, Robert. 2005. Minding one's own cognitive system: When is a group of minds a single cognitive unit. *Episteme: A Journal of Social Epistemology*, 1 (3): 177-88.
- Searle, John. 1990. Collective intentions and actions, in P Cohen, J Morgan, and ME Pollack (eds), *Intentions in Communication*. Cambridge: Bradford Books.
- Sherman, Steven and Elise Percy. (2010). "The psychology of collective responsibility," *Journal of Law and Policy*, 19, 137-170.
- Spencer, Herbert. 1897. *The principles of sociology* (3 vols.). New York: Appleton and Co.
- Susskind, Joshua, Kristin Maurer, Vinita Thakkar, David L. Hamilton, and Jeffrey W. Sherman. 1999. "Perceiving individuals and groups: Expectancies, dispositional inferences, and causal attributions," *Journal of Personality and Social Psychology*, 76, 2, 181-191.
- Cheng, Lisa Lai-Shen, and Rint Sybesma. 1999. "Bare and not so bare nouns and the structure of NP", *Linguistic Inquiry* 30 (4): 509-542.
- Sytsma, Justin and Edouard Machery (2009) "How to study Folk Intuitions about Phenomenal Consciousness," *Philosophical Psychology*, 22. 1, 21-35.
- Surowiecki, James. 2004. *The wisdom of crowds*. London: Doubleday press.
- Tollefsen, Deborah. 2002. "Organizations as True Believers." *Journal of Social Philosophy*, 33, 3, 395-411.
- Tollefsen, Deborah. 2004. "Collective intentionality," *Internet Encyclopedia of Philosophy* (accessed 14 May 2014), <http://www.iep.utm.edu/coll-int/>
- Velleman, David. 1997. "How to share an intention," *Philosophy and Phenomenological Research*, 57, 29-50.
- Wilson, Robert. 2001. Group-level cognition, *Philosophy of Science*, 68 (supp): S262-S273.

- Yzerbyt, Vincent, Olivier Corneille, and Claudia Estrada. 2001. "The interplay of subjective essentialism and entitativity in the formation of stereotypes," *Personality and Social Psychology Review*, 5, 141-155.
- Yzerbyt, Vincent, Steve Rocher, and Georges Schadron. 1997. "Stereotypes as explanations: A subjective essentialistic view of group perception," in Spears, Russell, Penelope Oakes, Naomi Ellemers, and Steven Haslam (eds.), *The psychology of stereotyping and group life*. London: Basil Blackwell, pp. 20–50.
- Yzerbyt, Vincent, Anouk Rogier, and Susan Fiske. 1998. "Group entitativity and social attribution: On translating situational constraints into stereotypes. *Personality and Social Psychology Bulletin*, 24, 1090–1104.
- Zawadzki, Tadeusz. 2013. *Mindshaping: A New Framework for Understanding Human Social Cognition*. Cambridge: MIT Press.

¹ As Felipe de Brigard reminds me, such data do not discriminate between the hypothesis that people commonly assume that every group member entertains the same mental state, and the hypothesis that people commonly assume that each of the group members merely entertain complementary mental states. I cannot address these possibilities here, but it is worth noting that each of these positions has been defended in the philosophical literature on shared intentions, typically on the basis of an appeal to commonsense intuition (cf., Tollefsen 2004).

² Here's the strategy: "first you decide to treat the object whose behavior is to be predicted as a rational agent; then you figure out what beliefs that agent ought to have, given its place in the world and its purpose. Then you figure out what desires it ought to have, on the same considerations, and finally you predict that this rational agent will act to further its goals in the light of its beliefs" (Dennett 1989, 17).