

Redundant Group Agency

Lars J. K. Moen¹

Abstract

According to group-agent realism, treating groups as agents with their own intentional states, irreducible to those of the group members, helps us explain and predict the groups' behavior. This paper challenges this view. When groups judge logically interconnected propositions, group members often have incentives to misrepresent their beliefs of propositions they care less about in order to increase the probability of their groups adopting their view of propositions they consider more important. Aggregating such untruthful judgments may lead to the group forming false beliefs. Treating groups as agents will then not help us explain or predict their behavior.

Keywords

group agency, intentionality, judgment aggregation, strategic interaction, true belief.

Received 26 September 2019.

¹ Australian National University, Canberra, Australia.

Corresponding Author:

Lars J. K. Moen, School of Politics and International Relations, Australian National University, ACT 0200, Australia.

Email: lars.moen@anu.edu.au

1. Introduction

‘It is, by now, relatively widely accepted that suitably organized collectives can be intentional agents in their own right, over and above their individual members’.² Christian List may be right about that, but are there good grounds for this wide acceptance? I shall not deny that we may be right to hold certain groups responsible for their actions, either instead of or in addition to holding group members responsible.³ But I shall challenge the view that ascribing agency to groups helps us predict and explain their behavior. Focusing primarily on List and Philip Pettit’s recent works on group agency, I argue that they do not fully appreciate the implications of strategic interaction between group members.⁴ On List and Pettit’s account, we should ascribe agency to a group to understand how it can reliably make consistent judgments of logically interconnected propositions. Doing so means ascribing to the group a capacity to make judgments most of its members reject. Thus emerges an autonomous group agent, List and Pettit argue. But the group’s mechanism for ensuring consistent judgments may give the group members incentives to misrepresent their personal judgments of propositions they deem less important so as to increase the probability of the group making the judgments they desire on more important propositions. A consequence of such strategic behavior is that the group will sometimes make false judgments we would not expect from an agent. It forms false beliefs that undermine its agency.

This argument is based on an account of agency endorsed by List and Pettit as well as other group-agent realists.⁵ This is the interpretationist account of agency, according to which an object, or a system, is an agent if and only if we can better explain and predict its behavior by treating it as an agent. I explain this approach in Section 2. In Section 3, I consider the

² List (2018, 295).

³ List and Pettit (2011, ch. 7).

⁴ Their most comprehensive work on group agency is List and Pettit (2011).

⁵ Especially Deborah Tollefsen. See Tollefsen (2002a; 2002b; 2015).

discursive dilemma, a social choice paradox showing how aggregating group members' judgments of a set of interconnected propositions can lead to inconsistent majority judgments. List and Pettit argue that we must treat groups as agents to understand the mechanisms the group applies in response to the discursive dilemma. But these mechanisms may give the group members incentives to vote untruthfully, and as I show in Section 4, such strategic behavior can lead to the group making false judgments in a way we cannot understand by treating it as an agent. Exploring how the individual group members behave as agents, however, will explain their strategic behavior and, consequently, the group's false beliefs. In Section 5, I show why ascribing agency to groups is a reliable strategy for predicting and explaining their behavior only under ideal circumstances.

The paper's argument against group-agent realism depends on the view that only groups, and not individuals, consist of multiple agents. If this group-individual distinction is false, and we successfully ascribe agency to individuals, then we might also justify group agency. In Section 6, however, I defend this distinction.

2. The basis for group-agent realism

Non-redundant group-agent realism (henceforth group-agent realism) is the view that groups can be agents in their own right with autonomous minds irreducible to their individual human constituents.⁶ In this section, I introduce a non-mystical and non-metaphorical account of group agency, according to which agency is not embedded in an object's physical make-up, but in its behavior.

Group-agent realists obviously do not think groups can have minds with a physical structure like human beings. So they reject an identity theory of mind, according to which a mental event correlate with a specific type of physical event. It would make little sense to say that the same mental event of a group mind and a human mind can correlate with the same

⁶ List and Pettit (2011, 4-7).

type of physical event. Functionalism, on the other hand, defines an object, or system, according to how it functions, without specifying its physical components.⁷ If one part, *a*, of a system performs the same function as a physically different part, *b*, of another system, we may give *a* and *b* the same definition, in spite of their physical differences. So if a physical event, *a*, in the human brain performs the same function with respect to the human mind as some human action, *b*, does within the structure of a group, we may say that *a* is a function of the human mind, while *b* is a function of the group mind. The idea of group agency is that there are functional analogues of individual-level processing at the group level.

The next step in creating a basis for group-agent realism is interpretationism. List and Pettit, as well as other group-agent realists, such as Deborah Tollefsen, follow Daniel Dennett's interpretationist account of agency.⁸ On this approach, we ascribe agency to a system without information about its brain processes.⁹ We do so as a way of explaining or predicting its behavior. We assume the system is rational, that is, it holds beliefs according to a pattern governed by truths and consistency.¹⁰ Recognizing this pattern enables us to successfully explain and predict the system's behavior. And when we can reliably interpret the group's behavior by assuming that it is an agent, then it is an agent. The pattern we recognize consists of intentional states. More precisely, the system's holds beliefs about its environment on which it bases its desires and motivations for action. Agents characteristically scan their environment before they form beliefs and desires. Change an agent's environment, and it will notice it and respond by changing its intentional states.¹¹

⁷ List and Pettit (2011, 28); Tollefsen (2015, 68-69).

⁸ List and Pettit (2011, 6, 23); Tollefsen (2002a; 2015, ch. 5).

⁹ Dennett (1987b, 48).

¹⁰ Dennett 1991.

¹¹ Dennett (1987a, 31).

By ascribing intentional states to the system, we assume it is rational and can form intentions based on its beliefs and desire. We have detected a pattern in its beliefs and desires, which suggests it can act on its own reasons. This pattern presupposes that the system has reflective states, which enables it to process its beliefs and desires so as to make them consistent. We assume the system is capable of making its beliefs coherent. If it believes that ‘*p*’ and that ‘*q*’, it also believes the implication of these beliefs, ‘*p* and *q*’.¹² We shall see that List and Pettit consider consistency especially important in their account of group agency.¹³

A group, then, is an agent if we can explain and predict its behavior by ascribing intentional states to it. It is an agent from our point of view when we set aside non-intentional possibilities for understanding its behavior, and instead assume that it is an agent with the intentional and reflective states we expect to find in an agent. For Dennett, something is an agent ‘only in relation to the strategies of someone who is trying to explain and predict its behavior’.¹⁴ This is the intentional strategy, or the ‘intentional stance’, as Dennett calls it. From the intentional stance, we attribute intentional states to the system without knowledge about its inner processes. So we take the intentional stance towards a group without information about its members’ intentional states.¹⁵ We begin by assuming it is a rational agent, we then determine what beliefs and desires it ought to have, and finally, how it ought to behave to further its goals based on its beliefs and desires. We can then say what the group ought to do, and that is what we predict that it will do.¹⁶ The intentional stance is justified only if it gives us a predictive power no other strategy can provide.¹⁷

¹² Dennett (1971, 94-95).

¹³ List and Pettit (2011, 24).

¹⁴ Dennett (1971, 87).

¹⁵ Tollefsen (2002a, 397).

¹⁶ Dennett (1987a, 17).

¹⁷ Dennett (1987a, 22-23).

Dennett distinguishes the intentional stance from the ‘design stance’ and the ‘physical stance’. From the design stance, we assume the system will function according to its design.¹⁸ If a computer is designed to produce an ‘L’ on the screen when we press the ‘L’ key on the keyboard, that is what we expect it to do from the design stance. From the physical stance, we explain and predict the system’s behavior according to its physical constitution and our knowledge about the laws of nature.¹⁹ The computer may be designed to produce an ‘L’ on the screen every time we press the ‘L’ key on the keyboard, but from the physical stance we possess knowledge of the computer’s physical components, and we may know that because we have spilled water on the computer, the ‘L’ key no longer works, and pressing it will no longer instruct the computer to produce an ‘L’ on the screen. Only from the physical stance can we predict that the ‘L’ key will not work. As Dennett writes, ‘the physical stance is generally reserved for instances of breakdown, where the condition preventing normal operation is generalized and easily locatable’.²⁰ As we shall see, a crucial point of this paper is that we can often make better predictions and explanations by taking the physical stance, and not the intentional stance, towards a group.

The intentional stance is superior to the physical stance when it provides a better or more effective way of explaining or predicting the system’s behavior. Taking the physical stance towards an intentional system, Dennett says, ‘would be a pointless and herculean labor’.²¹ Try to predict the next move of a chess-playing computer, one of Dennett’s favorite examples, by looking at its physical components, and you are in for a time-consuming and perhaps impossible task. Dennett instead recommends taking the intentional stance by seeing the computer as an intelligent opponent thinking rationally for the purpose of winning the

¹⁸ Dennett (1971, 87-88).

¹⁹ Dennett (1971, 88-89).

²⁰ Dennett (1971, 89).

²¹ Dennett (1971, 89).

game. We assume the computer possesses certain information, such as the rules of chess, and that it is motivated by some desire, such as winning the chess game. We are then, at least if we are decent chess players, in a position to predict the computer's next move.²² Alternatively, Dennett says, we 'can always refuse to adopt the Intentional stance toward the computer, and accept its checkmates'.²³ Analogously, humans successfully take the intentional stance towards one another in everyday interaction, which is a more effective strategy than examining each other's physical make-up.

Group-agent realists take the intentional stance towards groups because they consider it the most reliable and effective way—or perhaps the only way—of explaining and predicting their behavior. As Tollefsen argues, we miss 'real patterns of social behavior' by trying 'to explain the social world by appealing only to individual intentional states'.²⁴ Organizations, she says, 'really have beliefs', and seeing that makes us better equipped for explaining their behavior.²⁵ In List and Pettit's view, not seeing groups as mere collections of individual agents, but instead recognizing them as agents in their own right, 'parallels the move from taking a 'physical stance' towards a given system to taking an 'intentional stance''.²⁶ To understand the group, we should not look at its physical constituents—that is, its human members—but rather at how it interacts as an intentional agent with other agents in its environment. We should 'try to ascribe representations and motivations to it that makes sense of its actions'.²⁷ If this strategy works, we are justified in considering groups as agents in their own right.

²² Dennett (1971, 90).

²³ Dennett (1971, 91).

²⁴ Tollefsen (2002b, 43).

²⁵ Tollefsen (2002b, 43).

²⁶ List and Pettit (2011, 6).

²⁷ List and Pettit (2011, 23).

Dennett says the intentional stance ‘works with people almost all the time’.²⁸ And he adds that ‘[t]he strategy also works on most other mammals most of the time’.²⁹ It can also work on other animals, and even plants wanting to blossom in spring, and certain artifacts, such as the chess-playing computer.³⁰ The central question in this paper is whether or not group-agent realists are right to think the strategy can also work on groups.

3. Autonomous group minds

For a group to be an agent in its own right, it needs the capacity to form intentional states that are irreducible to those of its individual group members. A group governed by a dictator is not an agent, as its attitudes are fully reducible to the dictator’s. The same goes for groups governed strictly by a certain sub-group, such as a majority, as its attitudes are then reducible to the members of that sub-group. Autonomous group minds form attitudes that need not reflect the group members’ attitudes.

The group-agent realists I focus on here rely on a sensible view of groups as consisting of nothing more than their individual members organized according to the group’s structure.³¹ But as List and Pettit explain, although a group derives its agency entirely from the individual group members, we cannot always understand a group’s beliefs and actions by studying the intentional states of its human constituents.³² A group agent comes into existence when each individual group member intends to form and enact a group system capable of forming irreducible intentional states.³³ The group agent’s existence and functioning depend on its

²⁸ Dennett (1987a, 21).

²⁹ Dennett (1987a, 22).

³⁰ Dennett (1987a, 22).

³¹ List and Pettit (2011, 4).

³² List and Pettit (2011, 9).

³³ List and Pettit (2011, 34).

individual constituents, but are not fully understandable at the individual level, since individual-level intentional states will not always explain the actions and underlying intentions of such an intentional system. Group agency thus emerges from, but is nonetheless irreducible to, the group’s human constituents.

List and Pettit take a group’s response to the discursive dilemma to show how it forms irreducible intentional states, and why we need to see groups as agents.³⁴ As they explain, this social choice paradox may arise in groups such as legislatures, committees, multi-member courts, and expert panels advising on complex issues.³⁵ The discursive dilemma shows how aggregating individual group members’ judgments of a set of logically interconnected propositions cannot reliably lead to consistent group-level judgments of these propositions. Although all of the group members make consistent judgments, aggregating their judgments may result in the majority making inconsistent judgments. The problem is illustrated in figure 1, where three group members vote on three propositions, ‘*p*’, ‘*q*’, and ‘*p* and *q*’. Each of the group members makes consistent judgments, but the majority does not, as it believes that ‘*p*’ and that ‘*q*’ but not that ‘*p* and *q*’. For the group to function properly, it needs a reliable way of making consistent judgments.

	<i>p</i>	<i>q</i>	<i>p</i> and <i>q</i>
Group member A	True	True	True
Group member B	True	False	False
Group member C	False	True	False
Majority	True	True	False

Figure 1. A group makes inconsistent judgments of three interconnected propositions.

³⁴ List and Pettit (2011, ch. 2). The discursive dilemma was first presented formally as a general problem of social choice in List and Pettit (2002, 89-110).

³⁵ List and Pettit (2011, 43).

To solve this inconsistency problem, the group must reject a simple majoritarian voting rule, and instead adopt an aggregation function that delivers complete and consistent judgments of interconnected propositions for any possible combination of the group members' complete and consistent sets of judgments.³⁶ And that, List and Pettit argue, requires seeing the group as an agent in its own right.³⁷ We must treat it as a system with its own reflective states correcting inconsistent judgments. Looking at the group members' reflective states will be of no use, since each of them has contributed consistent judgments. To see how the group can achieve consistency in response to the discursive dilemma, we must view it as capable of departing from the group members' judgments so as to reinforce its own rationality.³⁸ We thus ascribe to the group 'an important sort of autonomy'.³⁹

List and Pettit suggest different procedures through which the group can achieve consistency. One solution is to apply a functionally explicit 'sequential priority rule', which specifies the propositions' order of priority.⁴⁰ The judgments of propositions of a higher priority will constrain the judgments of less prioritized propositions. The 'premise-based procedure' gives the premises priority over the conclusion (figure 2). If applied to the example above, the propositions '*p*' and '*q*' have higher priority than '*p* and *q*'. So if a majority of the group members support both '*p*' and '*q*', the group will judge '*p* and *q*' to be the case even should most of the group members reject this proposition.

³⁶ List and Pettit (2011, 67).

³⁷ List and Pettit (2011, 76).

³⁸ List and Pettit (2011, 5, 30-31).

³⁹ List and Pettit (2011, 76).

⁴⁰ List (2004, 495-513).

	p	q	p and q
Group member A	True	True	True
Group member B	True	False	False
Group member C	False	True	False
Group	True	True	False True

Figure 2. The premise-based procedure.

The ‘conclusion-based procedure’ works in the opposite direction by giving the conclusive proposition priority over the premises (figure 3). The group may then reject the group members’ judgments of the premises so as to make its judgments of the premises consistent with its judgment of the conclusion. Note, however, that the conclusion-based procedure will not always generate decisive judgments of the premises underlying the conclusion. The group thus makes an incompletely theorized decision, and therefore fails to form beliefs we expect an agent to make.⁴¹ The premise-based procedure therefore provides firmer grounds for the idea of group agency. But the more important point here is that the group adopts a belief that most of its members reject. The group forms judgments that are irreducible to its members’ judgments.

	p	q	p and q
Group member A	True	True	True
Group member B	True	False	False
Group member C	False	True	False
Majority	True False?	True False?	False

Figure 3. The conclusion-based procedure. Note that this procedure is indecisive on which of the first two propositions to reject, as it can reject either of them to achieve consistency at the group level.

An alternative procedure, especially advocated by Pettit, is deliberation among the group members.⁴² Unlike a functionally explicit sequential priority rule, this functionally *implicit*

⁴¹ List and Pettit (2011, 126-127).

⁴² Pettit defends this procedure in several of his works, for example Pettit (2007, 512; 2009, 81-88; 2012, 193-194).

structure will not apply mechanically but instead leave it up to the group members to decide how to proceed. When the group makes inconsistent judgments, the members get feedback from the group—that is, information about the group members’ voting—which gives them a holistic view of the majority’s judgments. They then deliberate and change their votes for the sake of making the group’s judgments consistent. By doing so, they turn the group into a rational agent correcting its own inconsistent judgments. As Tollefsen writes, ‘[w]hen individuals deliberate in an organizational setting, they adopt the rational point of view of the organization. It is from the point of view of the organization, rather than their own personal point of view, that deliberation ought to take place’.⁴³ Group members may vote against their personal views on propositions they consider less important to make sure the group’s judgments are consistent and in favor of their views of propositions they consider most important. With feedback from the group, the members can change their votes in response to other members’ votes. The members then vote not as separate individuals, but as parts of a group mind working towards making their group’s judgments consistent.⁴⁴

According to group-agent realism, then, the best way of explaining how the group forms its judgments is to treat it as an agent reflecting on its judgments so as to make them consistent. We must elevate the rational point of view from the individuals and up to the group level. We thus take the intentional stance towards the group, from which we can best make sense of the group’s judgments. We attribute intentional states to the group that may be the intentional states of none of the individual group members.⁴⁵

⁴³ Tollefsen (2002a, 401).

⁴⁴ List and Pettit (2011, 64).

⁴⁵ Tollefsen (2002b, 28).

4. Group beliefs and strategic interaction

Having now presented the case for group-agent realism, I shall in the remainder of this paper challenge this view. Taking the intentional stance towards groups, like we do towards individual persons, is more problematic than group-agent realists think.

List and Pettit identify three kinds of standards of rationality they require an agent to satisfy: ‘attitude-to-action’, ‘attitude-to-attitude’, and ‘attitude-to-fact’.⁴⁶ These are conditions a system must satisfy to be interpretable, or to be an agent. ‘Attitude-to-action’ standards require there to be group members who can act on behalf of the group. Although important, I do not consider these standards relevant for present purposes. ‘Attitude-to-attitude’ standards demand consistency in the agent’s beliefs and desires. Procedures successfully responding to the discursive dilemma, as discussed in the previous section, make sure the group meets these standards.

The problem with group-agent realism I identify in this paper concerns the ‘attitude-to-fact’ standards. To meet the ‘attitude-to-fact’ standards of rationality, List and Pettit explain, ‘the group must ensure, as far as possible, that its beliefs are true about the world it inhabits’.⁴⁷ It must form beliefs based on relevant evidence gathered from scanning its environment. From the intentional stance, we expect the group to form its beliefs in this way. We expect it to believe what it ought to believe given facts about its environment. If the group fails to meet the ‘attitude-to-fact’ standards, we cannot make sense of its beliefs, and we consequently cannot predict and explain its behavior from the intentional stance. As Dennett notes, ‘all there is to really and truly believing that p (for any proposition p) is being an intentional system for which p occurs as a belief in the best (most predictive)

⁴⁶ List and Pettit (2011, 36-37).

⁴⁷ List and Pettit (2011, 36-37). See also Tollefsen (2002a, 399-400).

interpretation'.⁴⁸ Forming beliefs that are obviously false and unpredictable from the intentional stance is to fail the agency test. From the intentional stance, the group will appear to malfunction, which suggests we should instead view it from the physical stance. On an interpretationist account, we must then conclude that the system is not an agent.

One way to test whether a group meets the 'attitude-to-fact' standards is to consider whether or not its beliefs are aligned with the rational group members' beliefs. If there is overwhelming evidence for a proposition being either true or false, and the competent group members unanimously either support or reject it, then we have good reason to believe that the group members' beliefs are true. From the intentional stance, we will in such cases expect the group's beliefs to be aligned with the group members' beliefs.

But consider a case in which a group applies a 'distributed premise-based procedure', and makes a judgment all of its members reject.⁴⁹ The group is divided into two subgroups responsible for judging different propositions. Subgroup *A* has a special right, due to its members' expertise, to decide the group's judgment of one proposition, '*p*'. And subgroup *B* has a special right to decide on another proposition, ' $p \rightarrow q$ ' (which reads 'if *p* then *q*'). *A*'s judgment of ' $p \rightarrow q$ ' is irrelevant for the group's belief formation, while *B*'s judgment of '*p*' is irrelevant. The members of *A* unanimously support '*p*', while the members of *B* unanimously support ' $p \rightarrow q$ '. But when the two subgroups (*AB*) both vote on '*q*', they unanimously reject '*q*'. When the premise-based procedure leads the group to nonetheless infer that '*q*' is true, the group forms a belief all of its members reject. The problem is presented in figure 4. In the third column (*q*), the group goes against both subgroups to achieve its consistency. The well-informed group members unanimously reject '*q*'. From the intentional stance, we do not see the group members' behavior, and the group's behavior may

⁴⁸ Dennett, (1987a, 29).

⁴⁹ I have taken this example from Dietrich and List (2007, 60).

therefore appear unpredictable and unexplainable. This suggests we should instead take the physical stance, from which we can see the individuals' behavior. The group's behavior will then make sense.

	p	$p \rightarrow q$	q
Subgroup A	True	False	False
Subgroup B	False	True	False
Group	True	True	False True

Figure 4. Subgroup A decides the group's judgment of proposition ' p ', while subgroup B decides its judgment of ' q '. Following the premise-based procedure, the group makes a judgment of the third proposition, ' q ', that the subgroups unanimously reject.

But this example is a very special case. Group-agent realists might say that such cases are rare, and therefore cannot challenge the view that the intentional strategy towards groups will be sufficiently reliable. But there are also other, and more common, cases in which a group must adopt beliefs its members unanimously reject. These are cases of strategic interaction between group members. By strategically misrepresenting their true beliefs in response to one another's judgments, individual group members can undermine the group's capacity to form true beliefs. Strategic behavior in response to others' (expected) behavior is a common way for individuals to behave. In Dennett's view, it is what makes us agents: we form our intentions according to what we believe other agents' intentions to be. Game-theoretical predictions about how humans will behave towards one another, Dennett says, 'achieve their accuracy in virtue of the evolutionary guarantee that man is well designed as a game player, a special case of rationality'.⁵⁰

Any interaction between agents can be modeled in game theory. Each agent's utility function determines its payoffs from any outcome of the interaction. The agent acts to maximize its own rewards given how others act, or how the agent believes others will act.

⁵⁰ Dennett (1971, 100).

The predictive power of game theory lies in the game's Nash equilibrium, or equilibria, at which each agent maximizes its payoffs given the rational behavior of the other agents. The agents' interactions are in equilibrium when it is irrational for any of them to deviate from their current behavior provided that the other agents are also rational. Equilibrium behavior therefore differs from a dominant strategy, which a rational agent follows regardless of how other agents behave.

In strategic interaction, each agent must anticipate the behavior of other agents before deciding how to act, and believe the other actors also approach the interaction in this way.⁵¹ Strategic interactions are a subset of game-theoretic interaction. Game-theoretic situations are non-strategic when there is a dominant strategy. Rational agents' actions are then perfectly predictable.⁵² Behavior is also non-strategic in parametric situations in which an agent believes, perhaps mistakenly, there is no need to take into account other agents' behavior because it makes no difference to the agent's payoff structure.⁵³ The agent will therefore not make assumptions about other agents' behavior, which is a necessary condition for a strategic situation.

I shall argue that group-agent realism relies on the highly questionable assumption that it is always, or at least very probable, that voting truthfully is rational for group members. That is, truthfulness is either an equilibrium or a dominant strategy. Although this assumption holds in many cases, it is often false. The group members will sometimes try to increase the probability of the group making their desired judgment of a prioritized proposition by misrepresenting their beliefs about other connected propositions.

⁵¹ Elster (1983, 77).

⁵² Elster (1983, 78).

⁵³ Elster (1983, 75).

List and Pettit are aware of the problem of strategic voting, and show how a group can make a judgment, say ‘not p ’, that all of the group members individually reject if each of them believes that the other group members will vote ‘not p ’, and each of them prefers conformity to truthfulness.⁵⁴ Since each of them expects all the others to vote ‘not p ’, none of them believes her or his vote will be pivotal. The group member then has an incentive to vote untruthfully if she or he desires conformity. With a slight preference for conformity, conformity around ‘not p ’ is an equilibrium even should each group member individually believe that ‘ p ’ is the case.⁵⁵

Group members may have a further reason to vote untruthfully if they have outcome-oriented preferences, that is, they try to increase the probability of the group, with consistency, making the decision they desire by misrepresenting their beliefs about underlying reasons.⁵⁶ We shall see how such individual strategies may lead the group to form false beliefs. The group, as an intentional system, thus malfunctions, which means we should approach it from the physical stance. This is the stance from which we make sense of interaction between the group members. In cases of group members strategically voting untruthfully, the intentional stance towards the group will be less reliable than towards individuals. Treating groups as agents consequently becomes problematic, and perhaps even pointless. We will therefore often be better off approaching groups from the physical stance as collections of agents rather than agents in their own right.

Consider a case in which a political party fails as a true believer by making a false negative—that is, it denies a true proposition—due to its members’ outcome-oriented preferences. The party decides on whether or not to support policies designed to reduce

⁵⁴ List and Pettit (2011, 119-121).

⁵⁵ Of course, everyone voting truthfully in support of ‘ p ’ is also an equilibrium.

⁵⁶ List and Pettit (2011, 111).

greenhouse-gas emissions. The party members know that the premise-based procedure will produce consistent group-level judgments if their voting causes an inconsistency at the group level. The members vote on three interconnected propositions: (p) ‘greenhouse-gas emissions cause global warming’, ($p \rightarrow q$) ‘if greenhouse-gas emissions cause global warming, then the party should support policies to reduce greenhouse-gas emissions’, and finally, (q) ‘the party should support policies to reduce greenhouse-gas emission’. All available facts point in the direction of ‘ p ’ being true. So as competent agents, the party members unanimously believe that ‘ p ’ is true. If they vote truthfully, the group’s judgment will be highly predictable from the intentional stance. But as we see in figure 5, two-thirds of the group members do not think ‘ $p \rightarrow q$ ’ is true. And it is important for the members of this majority that the party rejects ‘ q ’. They therefore decide to reject ‘ p ’ to make sure the party can reject ‘ q ’ without being inconsistent. This problem applies equally to cases of false positives, in which false propositions are believed to be true.

	p	$p \rightarrow q$	q
A	True	True	True
B	True False	False	False
C	True False	False	False
Majority	True False	False	False

Figure 5. Due to their outcome-oriented preferences, group members B and C misrepresent their beliefs about ‘ p ’ to increase the probability of the group adopting their beliefs about ‘ q ’.

By taking the intentional stance towards the party members, we understand that they misrepresented their beliefs of ‘ p ’ because of their outcome-oriented preferences. The premise-based procedure induced their untruthful judgment of ‘ p ’. It made them think that ‘Since I firmly believe the group should reject ‘ q ’, I should vote untruthfully against ‘ p ’, even though I believe ‘ p ’ is true’. We cannot explain the party’s behavior by pointing at the premise-based procedure, since it is only a part of the party member’s environment, and not

the party's. So while a behavioral pattern is traceable at the individual level, no pattern can be found at the group level. The party thus violates the 'attitude-to-fact' standards of rationality. It has malfunctioned, and we need to go down to the physical stance to make sense of its false and unpredictable judgment. The group members' untruthful voting does not compromise their status as true believers, as we have good reason to think that their voting does not reflect their actual beliefs. But we have no good reason to think that the group makes a judgment that does not reflect its sincere beliefs. The group members' strategic voting therefore compromises the group's status as a true believer.

A conclusion-based procedure is more reliable when the group members' preferences are outcome-oriented. But as already noted, this procedure may generate no decisive group-level judgments of the premises. To function as an agent, a group must reliably give reason for its decisions, and therefore cannot apply the conclusion-based procedure.

Finally, we have no reason to think the functionally inexplicit deliberation procedure will make a group more interpretable from the intentional stance. It will make truthfulness neither an equilibrium nor a general strategy. Suppose each party member with an incentive to misrepresent her or his judgment of '*p*' receives information from the other party members that voting truthfully will not lead to the group supporting '*q*'. These party members may still rationally decide to vote against '*p*' because it is a small price to pay for enhanced protection against the party supporting '*q*' if the other party members should change their minds, or if they lie. The party's judgments will then be just as unintelligible from the intentional stance as with a functionally explicit decision mechanism.

5. Group-agent idealism

In this section, I argue that viewing groups as agent in the same way as individuals are agents must rely on the idealization that group members will not strategically make untruthful judgments in spite of any reason they may have for doing so. List and Pettit are aware of the

possibility of strategic voting manipulating group judgments. They do not, however, fully appreciate that group agency is an idealized notion that does not give us the predictive and explanatory power we expect when we ascribe agency to a system.

List and Pettit acknowledge that their group-agent realism must deal with the group members' strategies to make their group adopt their desired judgments. 'A well-functioning group agent must ... cope with the basic fact that individuals are themselves rational agents', they say.⁵⁷ To do so, the group needs an organizational structure that is 'incentive compatible' with truthfulness, which means it will make each group member always prefer, or be indifferent to, the group-level judgments resulting from expressing her or his judgments truthfully, compared to the result of misrepresenting her or his judgments.⁵⁸ In short, the group's organizational structure is incentive compatible with truthfulness if it makes it rational for each individual to express her or his sincere judgments.⁵⁹ It makes truthfulness an equilibrium or dominant strategy for each group member. We can then expect each individual to function as a cognitive mechanism in a group mind predictably forming true beliefs. As List and Pettit write, '[i]f truthfulness is incentive compatible under a social mechanism involving the expression of individual judgments or preferences, we can be confident that rational individuals will indeed reveal their judgments or preferences truthfully'.⁶⁰ The intentional stance towards the group will then work again.

Failing to make truthfulness incentive compatible, List and Pettit write, means 'a group's ability to achieve its goals may be seriously compromised'.⁶¹ More precisely for present purposes, this failure means the group's ability to form true beliefs may be compromised.

⁵⁷ List and Pettit (2011, 104).

⁵⁸ List and Pettit (2011, 111).

⁵⁹ List and Pettit (2011, 105-106).

⁶⁰ List and Pettit (2011, 108-109).

⁶¹ List and Pettit (2011, 109).

Avoiding this failure is a great challenge for group-agent realism. List and Pettit show that a group's aggregation function is incentive compatible with truthfulness whenever two conditions are satisfied simultaneously.⁶² First, the propositions must be treated independently. That is, the group members must vote on each proposition as if they were not interconnected. Second, a positive group judgment can never change into a negative judgment if some group members who reject the proposition change their judgments towards accepting the proposition. We may call the former the 'independence condition', and the latter the 'monotonicity condition'.

The obvious question now is how the group can meet these two conditions. List and Pettit identify two different ways. The first is to assume that the group members' individual preferences are fixed, and then apply an aggregation function that satisfies the two conditions. This strategy works with 'a suitable premise-based or conclusion-based procedure', List and Pettit explain, '[i]n the lucky scenario in which the propositions of concern are mutually independent and fit to serve either as premises or as conclusions'.⁶³ This independence makes the outcome-oriented preferences of the judges in the example above insignificant, and both the premise-based and conclusion-based procedures will be incentive compatible with truthfulness.⁶⁴ But since the propositions of concern are in many cases logically interconnected, List and Pettit acknowledge that this strategy fails. It will only work reliably with degenerate procedures that violate unanimity or are dictatorial.

List and Pettit therefore turn to their other strategy, which is to 'try to change the individuals' preferences by persuading or convincing them that they should care about a

⁶² List and Pettit (2011, 112-113).

⁶³ List and Pettit (2011, 113).

⁶⁴ List and Pettit (2011, 113-114).

different set of propositions of concern'.⁶⁵ The strategy is to make the group members' preferences less outcome-oriented and more reason-oriented, or to make them care more about the relevant premises and less about the conclusion. List and Pettit admit it is difficult to see how this goal can actually be achieved.⁶⁶ It seems achievable, they suggest, in a society where people are esteemed for thinking and acting virtuously in we-terms, and people crave one another's esteem.⁶⁷ Furthermore, a required ethos makes individuals prefer stating their sincere judgments to conforming to what they believe to be the majority judgments.⁶⁸

The conclusion we should draw here is that making truthfulness incentive compatible depends on idealizations, that is, assumptions that are false in the real world, but ought to become true to meet the conditions of a theory. But the intentional stance is not meant to work reliably only under ideal circumstances. As List and Pettit themselves say, their account of group agency should not depend on group members being 'maximally well-behaved or idealized'; it should work with 'real people, who behave strategically when this is expedient'.⁶⁹ But since we know of no way of structuring groups so as to make truthfulness robustly incentive compatible, we must, base group agency on idealizations about human behavior. If we often fail to understand the reasons for a group's behavior from the intentional stance, we must instead take the physical stance towards it, which means seeing it not as an agent but as a collection of agents. We are consequently not justified in ascribing agency to the group. After all, interpretationism ascribes agency to systems to predict and explain real social phenomena. The intentional stance is of little use if it works only under

⁶⁵ List and Pettit (2011, 113).

⁶⁶ List and Pettit (2011, 128).

⁶⁷ List and Pettit, (2011, 128). See also Brennan and Pettit (2004).

⁶⁸ List and Pettit, *Group Agency*, 127.

⁶⁹ List and Pettit, *Group Agency*, 104.

ideal circumstances. The concept of group agency must be useful for real-world predications and explanations. Otherwise it is redundant.

6. Individuals

I do not argue that the intentional strategy can never explain or predict a group's behavior. Nor do I argue that individuals only hold true beliefs. 'Most agents are fallible', as List and Pettit correctly note.⁷⁰ What I argue is that groups face problems with belief formation in addition to those individual persons may face. And the intentional stance is consequently less reliable towards groups than towards individuals.

I do not argue against what Pettit and Frank Jackson call 'explanatory ecumenism', according to which higher-level and lower-level explanations are complementary, so if we ignore one of them, we miss out on relevant information.⁷¹ Whether to prefer 'a smaller grain or coarser grain explanation', they argue, 'depends on what one's perspective or purpose is'. Higher and lower levels of explanation 'provide complementary bodies of information on one and the same topic'.⁷² But regarding the interpretation of group behavior, we have seen that the intentional stance will often fail to provide relevant information because of strategic interaction within the group. We therefore need a more fine-grained analysis than group-agent realism can provide to reliably understand the group's behavior.

But perhaps the problem of strategic behavior applies to individuals as well. Perhaps individual agents consist of parts strategically trying to control their individual's behavior. That would be grist for the group-agent realists' mill, since it suggests the intentional strategy works just as reliably towards groups as towards individuals. And since it seems to work well enough to justify individual agency, it will then also justify group agency.

⁷⁰ List and Pettit (2011, 82).

⁷¹ Jackson and Pettit (1992, 1-21).

⁷² Jackson and Pettit (1992, 16).

Let us now consider whether it makes sense to think individuals, like groups, consist of parts with their own beliefs and desires capable of undermining individuals' capacity to form true beliefs. We may again think of the climate change example. A person, Linda, ignores the evidence for greenhouse-gas emission causing global warming because of her desire to maintain her current level of resource consumption. Like the political party in the example above, Linda states that she believes this proposition is false even though the evidence overwhelmingly suggest it is true. To continue the group-individual analogy, let us further suppose that Linda would have been persuaded by the evidence and considered the proposition to be true had she treated it independently of the question of whether or not she should reduce her resource consumption. But since the two are interconnected, she prioritizes the latter, and rejects the former for the sake of her own consistency. She expresses her beliefs strategically, thus undermining her own status as a true believer. But to complete the group-individual analogy, we must assume that also individuals consist of different agential parts with intentional and reflective states. Are there intentional agents within Linda that believe greenhouse-gas emissions cause global warming, while also believing Linda ought not to reduce her resource consumption? If there are, then these agents will have an incentive to misrepresent their views of the first proposition in so far as they consider the latter more important. The analogy will then be complete.

But this is not how individuals form their beliefs. Linda may act strategically by misrepresenting her beliefs, but her misrepresented beliefs are not reducible to the strategies of sub-individual intentional agents. In 'egonomics', which studies the tensions between what you can do or believe and what you feel you ought to do or believe, no such sub-individual agents are identified. Thomas Schelling, the founder of egonomics, says there is nothing 'mysterious or philosophically profound' in strategic self-management.⁷³ There are no agents

⁷³ Schelling (1984, 63).

within you behaving strategically in ways that may make you believe falsities. Richard Thaler and H. M. Shefrin understand self-control and intertemporal choice by modeling an individual as an organization with a farsighted planner and several myopic, or selfish, doers.⁷⁴ You make strategies, such as closing your eyes, to enable yourself to do what you think you ought to do, such as not giving in to the temptation of having another piece of cake. But this is the individual's strategy; there are no sub-individual agents developing strategies to gain control of the individual. We can interpret the individual's behavior, then, by ascribing intentional states to the individual, and not her or his parts. At the group level, however, the group members have strategies we cannot interpret by ascribing intentional states to the group. We can only make sense of these strategies by ascribing intentional states to the group members.

7. Conclusion

We ascribe agency to a system to reliably understand and predict its behavior. If doing so does not give us this explanatory and predictive power, we have no grounds for believing the system to be an agent. As Tollefsen says, '[i]f our best efforts to make sense of an individual fail, then there is no reason to believe that we are dealing with a rational agent and hence there is no reason to believe we are dealing with an intentional agent'.⁷⁵ This paper's argument against group-agent realism is based on a crucial difference between individuals and groups: only the latter consist of multiple agents interacting strategically with one another. Within a group structure, individuals will sometimes strategically misrepresent their personal judgments in order to bring about an outcome they desire. We have seen how such strategic behavior can lead a group to form beliefs that are obviously false. Taking the

⁷⁴ Thaler and Shefrin (1981, 392-406).

⁷⁵ Tollefsen (2015, 102).

intentional stance towards the group is no way of understanding its mistaken attitude towards its environment.

We must instead take the physical stance towards the group, from which we see the group members as intentional agents. From the physical stance we see how group members' strategic behavior can lead groups to form false beliefs in a way individuals do not. In such cases of false belief formation, ascribing intentional states to the group is of no use; we cannot interpret it as an intentional system. Only on the idealization that group members will reliably make truthful judgments will the intentional stance towards group work reliably. Group-agent realism, however, promises predictive and explanatory power in the real world if we ascribe agency to groups. It consequently fails to deliver the benefits it promises.

Acknowledgment

I would like to thank Keith Dowding and Philip Pettit for their helpful comments on earlier versions of this paper, and Justin Bruner and David Schweikard for constructive conversations around the ideas expressed in this paper. The paper was presented at the 2018 Pavia Graduate Conference in Political Philosophy, and I am grateful for the comments I received there.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

- Brennan, G. and P. Pettit. 2004. *The Economy of Esteem: An Essay on Civil and Political Society*. New York, NY: Oxford University Press.
- Dennett, D. C. 1971. "Intentional Systems." *The Journal of Philosophy* 68 (4): 87-106.
- Dennett, D. C. 1987a. "True Believers: The Intentional Strategy and Why It Works." In *The Intentional Stance*, 13-35. Cambridge, MA: The MIT Press.
- Dennett, D. C. 1987b. "Three Kinds of Intentional Psychology." In *The Intentional Stance*, 43-81. Cambridge, MA: The MIT Press.
- Dennett, D. C. 1991. "Real Patterns." *The Journal of Philosophy* 88 (1): 27-51.
- Dietrich, F. and C. List. 2007. "A Liberal Paradox for Judgment Aggregation." *Social Choice and Welfare* 31 (1): 59-78.
- Elster, J. 1983. *Explaining Technical Change: A Case Study in the Philosophy of Science*. Cambridge: Cambridge University Press.
- Jackson, F. and P. Pettit. 1992. "In Defense of Explanatory Ecumenism." *Economics and Philosophy* 8 (1): 1-21.
- List, C. 2004. "A Model of Path-Dependence in Decisions over Multiple Propositions." *The American Political Science Review* 98 (3): 495-513.
- List, C. 2018. "What Is It Like to Be a Group Agent?" *Noûs* 52 (2): 295-319.
- List, C. and P. Pettit. 2002. "Aggregating Sets of Judgments." *Economics and Philosophy* 18 (1): 89-110.
- List, C. and P. Pettit. 2011. *Group Agency: The Possibility, Design, and Status of Corporate Agents*. Oxford: Oxford University Press.
- Pettit, P. 2007. "Rationality, Reasoning and Group Agency." *Dialectica* 61 (4): 495-519.

- Pettit, P. 2009. "The Reality of Group Agents." In *Philosophy of the Social Sciences: Philosophical Theory and Scientific Practice*, edited by Chris Mantzavinos, 67-91. Cambridge: Cambridge University Press.
- Pettit, P. 2012. *On the People's Terms: A Republican Theory and Model of Democracy*. Cambridge: Cambridge University Press.
- Schelling, T. C. 1984. "The Intimate Contest for Self-Command." In *Choice and Consequence*, 57-82. Cambridge, MA: Harvard University Press.
- Thaler, R. H. and H. M. Shefrin. 1981. "An Economic Theory of Self-Control." *Journal of Political Economy* 89 (2): 392-406.
- Tollefsen, D. 2002a. "Organizations as True Believers." *Journal of Social Philosophy* 33 (3): 395-410.
- Tollefsen, D. P. 2002b. "Collective Intentionality and the Social Sciences." *Philosophy of the Social Sciences* 32 (1): 25-50.
- Tollefsen, D. P. 2015. *Groups as Agents*. Cambridge: Polity Press.

Author biography

Lars J. K. Moen is a Ph.D. candidate in the School of Politics and International Relations at the Australian National University. His research is in political philosophy and philosophy of social science. He currently focuses on group agency, political concepts of freedom, and methodological issues in political philosophy.