Constructivist and Ecological Modeling of Group Rationality

Gerald Gaus

1 The Constructivist and Ecological Approaches

Following F A. Hayek, Vernon L. Smith distinguishes constructivist and ecological forms of rationality, or perhaps we should say, forms of modeling rational agents (Smith, 2008: ch. 2). Each of these forms is itself multifaceted, but critical to constructivist models is an engineering perspective (Hayek, 1979: ch. 10), or what Christian List and Philip Pettit, following Daniel Dennet, call the "design stance" (13).¹ We might distinguish two phases of an extended design stance: (i) we identify, as it were, the basic engineering principles and blueprints that are required to create group agents, seeing how they work, or determine whether they are possible and (ii) with these in hand, we examine current instances of group agents and employ our principles to see how to improve their functioning.² Axiomatic analyses of rational agency, as employed by traditional economics and axiomatic social choice theory, are quintessentially constructivist: the aim is to define a set of formal principles (axioms) that, along with inputs (at best, minimally substantively constrained) will generate rational, consistent, outputs.³ Armed with these formal principles the constructivist seeks to develop blueprints for rational agents, or at least a set of design specifications. Constructivist modeling of rational agents has viewed rationality as a more-or-less general, all-purpose capacity: it is based on formal principles such as *modus ponens* and Arrovian-like consistency requirements that are assumed to hold across all domains of reasoning, or at least across very large swaths (say all representation-based reasoning, all preference-based reasoning) (see Gigerenzer, 2002, chs. 9-10). In the careful yet bold constructivist analysis of List and Pettit, the principles not only range over all domains of reasoning, but with modest modifications, over different types of agents — groups as well as individuals. Indeed, they are especially daring: their principles of rationality range over a sweeping variety of groups (39-41).

Hayek sometimes wrote as if constructivism was simply an error, but as Smith (2008: 2) quite rightly points out, that itself would be an error. Constructivist analysis — such as traditional utility theory or axiomatic social choice theory — can be tremendously enlightening in a wide range of cases. However, an ecological analysis is apt to manifest three departures from constructivism. (i) In place of the design stance, it will often be more enlightening to take up the evolutionary perspective, and see how ways of reasoning — which may well violate the constructivist's formal axioms — are adaptive in some environments. The core of the ecological idea is that good reasoning is in an important sense defined by adaptive responses to the environment in which reasoners of a certain type find themselves. This evolutionary analysis may appeal to biological evolution, social evolution, or gene-culture co-evolution.⁴ For example, while the constructivist is apt to interpret Amos Tversky and Daniel Kahneman's important body of findings concerning cognitive heuristics as "quick and dirty" ways of reasoning that constitute (perhaps excusable) violations of the norms of good reasoning, ecological analyses such as Gerd Gigerenzer's stress how they are ways of thinking that "make us smart" because they are effective context-dependent adaptive responses to our environment. (ii) The ecological approach stresses that good reasoning is often domainspecific: the norms of good reasoning in, say, social judgments may be very different than those in science.⁵ In social domains, as I will suggest later on (§4), there is good reason to suppose that moral judgments and reasoning involve an emotional repertoire;

those who lack it simply cannot reason well about these matters, even though they are flawless in the use of propositional logic and even if they (in some sense) endorse the true relevant propositions. (*iii*) Following from this, while reasoning has a fundamental procedural component, we will often find that only by attributing substantive concerns to agents can we grasp basic modes of reasoning. For example, deontic reasoning in interpersonal social contexts seems driven by the goal of cheater detection: we excel at deontic reasoning, and it is so important to do it well, because it is crucial to human social contexts to detect cheaters on cooperative norms. On the other hand, we are not very good at reasoning on the basis of material conditionals, which are in may ways analogous forms of reasoning.⁶

In these brief remarks I shall highlight three aspects of List and Pettit's analysis that I think well illustrate its constructivist nature: (*i*) its stress on the discursive dilemma as a primary challenge to group rationality and reasoning; (*ii*) its general though qualified support for premise-based decision-making as the preferred way to cope with the problems of judgment aggregation; and (*iii*) their account of rational agency and moral responsibility. I wish to stress that in my estimation theirs is a powerful constructivist analysis: they are cognizant of all the main problems, and do their best to cope with them. *Group Agency* is original, carefully argued and just plain smart. And constructivist.

2 JUDGMENT AGGREGATION AND THE DISCURSIVE DILEMMA

List and Pettit are famous for their work (both jointly and singularly) on problems of judgment aggregation and the "discursive" dilemma — an application of Arrow-theoremlike reasoning to the problem of judgment aggregation. No doubt most of the contributions to this symposium will rehearse it, so as not to be an outlier, consider a problem with a clear logical structure, such as r if and only if p and q. Suppose as in Display 1 (adapted from p. 45) that a three member group employs majority voting, and they take a vote on both premises as well as the conclusion.

	р	q	r
Alf	TRUE	TRUE	TRUE
Betty	TRUE	FALSE	FALSE
Charlie	FALSE	TRUE	FALSE
Majority	TRUE	TRUE	FALSE

DISPLAY 1: A DISCURSIVE DILEMMA

Only Alf deems both premises to be true, so only he affirms r; yet the group (the majority) affirms both p and q, so given the logical structure of the problem they (the majority) are committed to r; but the majority deems r to be false. List and Pettit generate a theorem showing that there is no possible aggregation function — no way of combining individual judgments (attitudes) into a group judgment (attitude) — that is guaranteed to meet four conditions: universal domain, collective rationality, anonymity and "systematicity" (which requires that group attitudes on each proposition depend only on the attitudes towards it, an independence condition) (49-50).

The practical power of any impossibility theorem depends on two factors: the degree to which the axioms are intuitively compelling and, for a given context, how likely it is that the relevant inconsistency will arise. In particular contexts (e.g., democratic decision making, developing a social welfare function to guide policy, or expert panels) we may have enough information about the context to determine how compelling are the axioms. However, List and Pettit's defense of their axioms is considerably blunted insofar as we do not know a critical contextual factor: the type of group under consideration.

Consider, for example, List and Pettit's universal domain condition: "the aggregation function admits as inputs any possible profile of individual attitudes towards the proposition on the agenda, assuming that individual attitudes are consistent and complete" (49),⁷ Focusing from now on simply on judgment aggregation, my worry is that as general matter, there is no reason to expect all or even most group agents to endorse universal domain. A variety of groups are defined by restricting the domain of admissible beliefs, such as many churches and other groups committed to a set of common beliefs. And even when the restriction on domain is not explicit, the point of many groups is joint action by like-minded people — not simply alike in their values, but in their outlook on the world in general. To be sure, members of such group agents will not agree about everything, but they will confront a wide variety of problems in which some of the propositions in the deliberative problem are set by the homogeneity that characterizes the group. As List and Pettit remark, their universal domain condition is an acknowledgement of a (deep) pluralism of views in our society,⁸ but they wrongly claim that we should expect to find such pluralism to be a feature of group agents in general (52). Pluralistic societies contain many homogenous groups: those that are highly pluralistic in their beliefs are a distinct subset of group agents.

The problem, though, is deeper. It seems reasonable to conjecture that the conditions under which the universal domain axiom is most compelling will tend to be precisely those circumstances in which the judgment aggregation model of collective reasoning

will not apply. The judgment aggregation/discursive dilemma analysis is only of relevance if a group regularly confronts decisions where they disagree about the truth of the relevant propositions but the relations among the propositions have an agreed-upon logical structure. Call this *proposition pluralism/structure monism*. We know that this combination sometimes occurs (in various courts and expert bodies), but it is hard not to wonder whether it is a general feature of group agents. A hypothesis: as we introduce pluralism due to what John Rawls called the "burdens of judgment,"⁹ disagreement about the truth values of propositions and disagreement about the relation between the propositions will tend to increase together. But if there is extensive disagreement about the relation between various propositions the group will disagree whether they confront a possible discursive dilemma. (Of course it may still be a fact that they confront such a dilemma because there is a certain structure to the propositions which a demigod would know; but if we could bring in a demigod we might just as well consult it directly about the truth of the propositions rather than taking a vote.) As we relax the "burdens of judgment," we would also generally expect to move towards agreement on both the relation among propositions as well as their truth values (in which case the judgment aggregation model makes sense, but the universal domain axiom is not compelling). Of course these two types of agreement will not be perfectly correlated, and we would sometimes expect the proposition pluralism/structure monism that gives the discursive dilemma its bite. The empirical questions are how often, and for what sorts of groups.

3 PREMISE-BASED DECISION MAKING AND GROUP RATIONALITY

Given the impossibility theorem, we know that there is no judgment aggregation function that is guaranteed to always conform to all the axioms. This, of course, is different than saying "a group cannot satisfy" all the axioms (59) in the sense that it must confront inconsistencies of judgment or modify the axioms. Neither does the impossibility theorem show that "majority voting on all propositions...does not generally work" (58). Whether a group confronts inconsistencies or violates axioms will depend on the sorts of judgment profiles that in fact arise in it even if many more profiles are in principle admissible. This is, I think, of considerable significance. A constructivist design approach seeks, at least in the first instance, formal principles that are fully consistent under all possible configurations of acceptable inputs; an ecological approach will concern itself with the extent to which problems actually arise, or are likely to arise, with which the agent must cope. The response of the design approach to an impossibility theorem is, as it were, to alter the design specifications by relaxing one or more the axioms so as to avoid the possibility of inconsistency.

List and Pettit generally endorse "giving up systematicity" (58) as the best response. To, unfortunately, ignore much of a subtle and complex analysis (apologies here to List and Pettit), an upshot of giving up systematicity is to allow premise-based aggregation functions, in which, say, the group votes on each of the premises; the conclusion is generated by the accepted premises and their logical structure (thus in Display 1, the group holds that r is true). In a pure form of premise-based aggregation, the group does not directly give its view on the conclusion (r). Obviously, premise-based aggregation functions avoid the discursive dilemma. Moreover, employing Condorcet jury theorem reasoning, List and Pettit argue for the general epistemic superiority of premise-based aggregation over conclusion-based aggregation (in which the voters directly register their judgments only on the conclusion). Their analysis of the epistemic case (chap. 4) is

thorough, and they are well aware of the complexities and problems. For now, let me simply stress that (roughly) the jury theorem only provides an epistemic case for group decision making if, on the proposition she judges true or false, the average judger has a better than .5 chance of being correct.

The fundamental problem with premise-based aggregation is its tendency to induce disagreement between the judgments of members' rationality and the group's rationality. As List and Pettit rightly note, the basic fact that must be taken into account when analyzing the rationality of a group agency is that its constituents are themselves rational agents (104). This, in my view, produces a discontinuity between the principles of individual and group rationality. Consider again Display 1: if we accept premise-based aggregation, the group judges that r, yet two-thirds of the group, exercising their own reason, judge that not-r. Indeed, in some ways the whole point of adopting the premise-based procedure is that in circumstances such as these the group will judge r even though most individual members do not. Let us say that that under such conditions group rationality fails to be congruent with individual rationality. Most of the members cannot see how r is the correct answer; exercising their own capacities the best they can, they conclude not-r.

Reliance on jury theorem type analysis, I think, considerably aggravates the problem, for the case that r better tracks the truth than not-r is based on a statistical analysis of the benefits of judgment pooling rather than substantive reasons for r. Of course after the vote the members might take the pooling information into account, and this piece of evidence could lead them to change their own conclusions to r.¹⁰ Since the problem has a certain logical structure, they would have to go back and see what premise they need to accept. The question is whether the statistical information would, or indeed, should, be determinative for them. If it is generally determinative, group diversity on those propositions will decline, and the group will move towards homogeneity. But this requires that individuals adopt views towards propositions not basically on what they see as the substantively relevant evidence, but on the law of large numbers and the assumption that the average competency of the group is better than .5.

However, there are good grounds for members to reject the conclusion that the Condorcet jury provides a strong case for them to revise their views. Suppose in Display 1 Betty's use of her individual reason still leads to the conclusion that r is false; as far as she can see, that must be because the group has wrongly voted q as true. If she still cannot see q as true, she must conclude that the group is wrong; and that is only a plausible conclusion if the competency of the voters on the q question was, after all, less than .5 (see further Estlund, 2008: ch. 12). Now once Betty draws this conclude that the group *must* be wrong on q (when average competency drops below .5, the probability that the group is wrong pretty quickly approaches unity as the group gets moderately large) thus reinforcing her dissent on q; but that means that it must be also wrong on r, as she thought.

When we adopt an ecological view, we better appreciate that a group faces two adaptation problems. The first is that upon which List and Pettit focus: as an agent in an environment it must successfully pursue its goals, and its group rationality specifies the ways that it processes information and members' attitudes to yield effective group actions. However, it also faces an internal adaptive problem: the procedures and programs it employs in interacting with its external environment must be consistent with maintaining its existence as a rational group agent. It must ensure that sufficient members continue to play their part in the group.¹¹ Different groups may handle this adaptive problem differently: some through sheer threats to comply with group conclusions, others with monetary incentives. But for any effective group there must be some core members for whom the group's judgments are congruent with their own, so they can confirm that this is indeed a rational thing to do. If we are talking about voluntary groups in a free society that are not simply held together by monetary incentives, then I suggest that a general desideratum for group rationality is that its judgments display significant congruence with the individual rational judgments of its members. Otherwise, members will decide that it is acting queerly and irrationally and exit. It has been said that justice must not only be done, but must be seen to be done; in a rational group agent, rationality must not only guide its decisions, but these decisions must be seen as rational. An adaptive norm of rationality for a group will thus be determined by a combination of the demands of its external agency and the desideratum that the norm and its outputs are generally congruent with members' individual rationality: if the norm fails badly on either count the group's constitution will be maladaptive. In this light, one must wonder about stressing premise-based aggregation — especially insofar as it relies on jury theorem reasoning — as a more-or-less general recipe for group rationality.

Now as I said, List and Pettit are well aware of the critical problems confronting their analysis: the fifth chapter of *Group Agency* is devoted to the "incentive compatibility" problem. They see the problem arising from the fact that members are also rational agents as the possibility that members may act strategically, and so defect from the

pursuit of collective goals and "seriously compromise the group's performance as a whole" (104). Interestingly, they consider a case of judges on a panel each of whom is "convinced of the truth of his or her own judgment" and who have strong outcomeoriented commitments, and so are tempted to vote strategically on the premises in order to achieve what the see as the correct outcome (111). List and Pettit depict this as an incentive problem: how do we get the constituents elements to act on the rational group norms? In principle, those who defect from the group judgment because they see it is unjustified or irrational are on all fours with those who defect to insure higher personal payoffs. However, to put aside one's interests for the sake of the group strikes me as significantly different than being self-effacing in the use of one's reason. Forms of group rationality that make heavy demands on its members be self-effacing in the use of their own reason simply will not be confirmed as rational by most members who compose the group. By taking up the design perspective, the problem for group agency is seen in terms of pesky members who insist on relying on their own reason: the constructivist's planning problem is how to arrange the incentive structure to mitigate this pesky behavior.

If, however, congruence with individual rationality is part and parcel of group norms of rationality, we are led to seeing group rationality as an issue within social and political philosophy. Rather than starting the analysis with a simple robot's agency and working up from there (chap. 1), we might see selection of the norms of group rationality as itself a group deliberative problem. Rawls's two-stage argument in the original position is a helpful model (see Gaus, 2013). The members of a group, reflecting on their values and ends, would first endorse a set of norms of group rationality to guide their group agency

and then, at the second stage, would check whether the expected results tend to be congruent with the exercise of the individual rationality of most members. Persistent significant lack of congruence must lead them to reevaluate the rationality of their shared norms.

4 RATIONAL AND MORAL AGENCY

In chapter 7 of *Group Agency* List and Pettit move from simply rational group agency to a version of moral agency — the sort of agent that can held responsible for its actions. On their account a group must meet three conditions to be hqeld morally accountable: (*i*) it must face normatively significant choices; (*ii*) it has to understand and have access to the relevant evidence for deciding these normative questions; (*iii*) it must have control in the choice among the options (158). Because at least many groups can meet these three conditions, they can be held responsible for their judgments (and so actions). List and Pettit invoke Strawsonian reactive attitudes: when we hold agents responsible we blame, are indignant and so on (154). When we are responsible we feel guilt.

There is very strong evidence that full-fledged adult moral judgment is a cognitiveemotional complex.¹² Four-to-six year olds typically form correct normative judgments: they know what actions are right and wrong, and can even provide justifications for the rules and norms. But because they have not developed guilt responses; they expect to be happy if they successfully violate the rule to their own advantage. In older children normative judgments are accompanied by emotional reactions of guilt; violators (self or others) are now expected to be unhappy, even if they benefit from the violation. Psychopaths are less competent than even young children, being unable to distinguish moral and conventional rules,¹³ but like young children they do make normative judgments: they can judge what is right and wrong. However, again because of lack of affect, they do not experience guilt or remorse. So in one respect very young children and psychopaths are similar: for both populations normative judgments are propositions divorced from their emotional life. Not only are such agents themselves unable to experience the reactive attitudes, but their inability to grasp the internal oughtness of moral judgment undermines our reactive attitudes toward them. Not really grasping what it is to make a moral judgment, and (at least in the psychopath's case) lacking empathetic ability, their failure does not manifest that ill-will towards us that Strawson placed at the heart of the reactive attitudes. This idea that competent judgment is affect-dependent is, of course, not a general feature of judgment, but one specific to domain of moral reasoning.

Reflect once again on Display 1, and assume that: p = all bribery payments are wrong, q = our operations in Australia involve bribery payments, and r = our operations in Australia involve wrongdoing. On a premise aggregation model, only Alf would have the relevant moral emotion, for only Alf has formed the judgment. The group will have made the normative judgment, but in ways approaching the young child and the psychopath, the relevant emotional response is absent; Betty and Charlie do not experience any such emotion, as they do not believe both premises. Under these conditions it looks as if the Strawsonian reactive attitudes and so moral responsibility is at least severely undermined toward the group. The group acting against the moral judgment does not display ill-will by turning its back on the internal ought. Unless one is willing to hold six-year olds responsible, it is hard to see how the group can be held responsible under premise-based aggregation.

CONSTRUCTIVIST AND ECOLOGICAL MODELING OF GROUP AGENTS/14

The difficulty here for the List-Pettit model does not, I think, simply derive from their general (and certainly qualified) advocacy of premise-based judgment aggregation, but their overall belief/preference model which, more broadly, is presented as domain general. Given this model they can handle for any question q, the aggregation of judgments based on beliefs or aggregation of preferences. In my view, the domain of moral reasoning does fit well into this picture. Emotional states are neither simply beliefs nor preferences nor even a combination of the two; and in mature moral agents certain judgments and emotions go hand-in-hand.

5 CONCLUSION

In their analysis of epistemic practices for groups, List and Pettit quite rightly note that there is no "one size fits all" optimal design (102). But yet at the heart of their analysis is that "robust group rationality" — the requirement that the group gives consistent answers on *all questions* confronting it — seems "non-negotiable" (141). Completeness is an especially controversial requirement: actual bounded individuals often have incomplete preference orderings, and cannot decide some issues. As Amartya Sen (2002: ch. 4) has shown, rational choice does not require completeness. Some groups may find that the benefits of securing completeness are worth the costs (and there are always costs in designing systems to achieve it), others not.¹⁴

When I think about the vast array of groups — from The Bill and Melinda Gates Foundation to Al Qaeda, the BBC to the CIA, the Crips to the FBI, General Motors to Rocko's Pizza on Broadway, the Royal Academy of Arts to the Scientologists to "Vakashas Vampires Coven Inc.," Microsoft to the Jimi Hendrix Experience, the Russian Federation to the Russian Ballet, the University of Arizona to the Jackson Square Tarot Card Readers Association, the Roman Catholic Church to OPEC — I can't help but be skeptical that there is a single general story to be told about the norms of reasoning that are adaptive to these radically different internal and external environments

Notes

¹ Unless otherwise noted, all page references in the text refer to List and Pettit, 2011.

² List and Pettit (2011: 13) only explicitly identify the latter as the design stance.

³ On the constructivist nature of traditional economic modeling, see Smith, 2008 and Gigerenzer, 2000: 202.

⁴ Hayek (1973: ch. 1) was adamant both that the evolutionary perspective was crucial, and that it concerned social, not biological evolution. I examine his account of evolution in Gaus, 2006; a general overview of social evolutionary accounts, see Gaus and Thrasher, 2013.

⁵ See Gigerenzer, 2000: chps. 10-11. I consider domain specific reasoning in relation to deontic social morality in Gaus, 2011: 126-31.

⁶ I consider the literature on this question in Gaus, 2011: §8. For an overview see Manktelow and Over, 1995. The work of Denise Cummins (1996) is especially enlightening on this question.

⁷List and Pettit point out that if a certain proportion of individual attitudes are "unidimensionally aligned" the relevant judgmental inconsistency will not occur. Suppose we have a dimension in which all individuals (we do not need this strong of an assumption) can be arrayed over a dimension, such as left/right in politics. Suppose further that those on the left favor p and those on the right favor not-p, and similarly for q. Epistemic positions are thus strongly correlated with some underlying dimension, such as ideology, a dispute between different schools, factions, or whatever. In this case the relevant inconsistency cannot occur.

⁸ "Deep," because unidimensionally aligned judgments display pluralism (see previous note), but that shallower sort of pluralism does not generate the dilemma.

⁹ For example, we disagree on many matters because the evidence is often conflicting and difficult to evaluate and even when we agree on the relevant considerations, we often weigh them differently; because our concepts are vague we must rely on interpretations that are often controversial; the manner in which we evaluate evidence and rank considerations seems to some extent the function of our total life experiences, which of course differ. See Rawls, 1996: 56-7.

¹⁰ Rousseau (1973: 250) famously suggests that this may occur. Those in the minority should say: "If my particular opinion had carried the day I should have achieved the opposite of my will...."

¹¹ I consider this idea of two levels of adaptive pressures for groups in Gaus, 2006.

¹² This paragraph draws on Gaus, 2011: §§11-12 where the relevant studies are discussed. What I say here is regrettably sketchy and didactic.

¹³ There is dispute about this; see Shoemaker, 2011.

¹⁴ Certainly *consistency* must be "non-negotiable!" Michael Freeden (1996: chs. 1-2), the foremost theorist of ideologies, stresses that political ideologies, while not "dispensing with reason," seek only some degree of internal coherence. The success of some groups may depend on their ability to maintain some inconsistencies

References

- Cummins, Denise. 1996. "Evidence for the Innateness of Deontic Reasoning." *Mind & Language*, vol. 11 (June): 160–90.
- Estlund, David. 2008. Democratic Authority. Princeton: Princeton University Press.

Freeden, Michael. 1996. Ideologies and Political Theory. Oxford: Clarendon Press.

- Gaus, Gerald. 2006. "The Evolution of Society and Mind: Hayek's System of Ideas." In *The Cambridge Companion to Hayek*, edited by Ed Feser. Cambridge: Cambridge University Press: 232-58.
- -----. 2011. The Order of Public Reason. Cambridge: Cambridge University Press.
- 2013. "The Turn to Political Liberalism." In A Companion to Rawls, edited by Jon Mandel and David Reidy. New York: Wiley-Blackwell.
- and John Thrasher. 2013. "Social Evolution." In *The Routledge Companion of Social and Political Philosophy*, edited by Gerald Gaus and Fred D'Agostino. New York: Routledge.
- Gigerenzer, Gerd. 2002. *Adaptive Thinking: Rationality in the Real World* (Oxford: Oxford University Press.

Hayek, F. A. 1973, Rules and Order. Chicago: University of Chicago Press.

- 1979. The Counter-revolution in Science: Studies on the Abuse of Reason.Indianapolis: Liberty Fund.
- List, Christian and Philip Pettit. 2011. Group Agency: the Possibility, Design, and Status of Corporate Agents. Oxford: Oxford University Press.

- Manktelow, K. I. and D. Over. 1995. "Deontic Reasoning" in Perspectives on Thinking and Reasoning: Essays in Honor of Peter Wason, edited by Stephen E. Newstead, and Jonathan St. B. T. Evans. East Sussex, UK: Lawrence Erlbaum: 91–114.
- Rawls, John. 1996. *Political Liberalism*, paperback edition. New York: Columbia University Press.
- Rousseau, Jean-Jacques. 1973. *The Social Contract* in *The Social Contract and Discourses*, translated by G. D. H. Cole. London: Dent.
- Sen, Amartya. 2002. "Maximization and the Act of Choice." In his *Rationality and Freedom*. Cambridge, MA: Belknap Press of Harvard University Press: ch. 4
- Shoemaker, David W. 2011. "Psychopathy, Responsibility, and the Moral/Conventional Distinction." *The Southern Journal of Philosophy*, vol. 49, Special Supplemental: 99-124
- Smith, Vernon L. 2008. *Rationality in Economics: Constructivist and Ecological Forms*. Cambridge: Cambridge University Press.