ATTITUDE AGGREGATION FOR DEMOCRATICALLY STRUCTURED GROUPS

By

CASEY HART

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

(Philosophy)

at the

UNIVERSITY OF WISCONSIN-MADISON

2017

Date of Oral Examination: December 18, 2017

This dissertation is approved by the following members of the Final Oral Committee:

Michael G. Titelbaum	Associate Professor, Philosophy
Elliott Sober	Hans Reichenbach Professor & William F. Vilas Research Professor, Philosophy
Daniel M. Hausman	Herbert A. Simon & Hilldale Professor, Philosophy
Alan Sidelle	Professor, Philosophy
Paul Kelleher	Associate Professor, Bioethics, Philosophy

For Nicole, Lizzie, and Juliette—

Abstract

Groups are all around us: families, teams, corporations, committees, nations, and so on. We interact with groups, are members of groups, and we blame or praise groups for how they act. Despite being so familiar, unpacking the nature of groups and their constitutive group members presents serious philosophical puzzles. In response to these puzzles, some take the route of dismissing groups; groups are merely shorthand for something that reduces to the individuals that make up the 'group'. I reject this eliminativism and opt for a robust realist conception of some groups. But this raises the task of showing how groups are built up out of individuals in such a way that explains our experience of groups and does not just reduce to the individuals themselves.

My core thesis in the dissertation is that a democratically structured group—a group wherein every member has equal standing, broadly construed—has propositional attitudes that can be determined by the Steward Approach. The Steward Approach posits a Steward, a rational agent whose evidence is all and only the honest testimony of the group members. The group attitude is identical to the set of rational attitudes that the Steward can have. This approach has several virtues. First, it is a positive and solvable approach. Epistemologists generally agree that there is an answer to the question of what is rationally permissible in any given evidential situation, despite the difficulty of determining what response is rational in some cases. Second, this approach bears fruit in sidestepping the impossibility theorems that indicate such a solution is not to be found. Third, and most importantly, this approach gives intuitive results that jive with our common-sense understanding of groups.

I conclude the dissertation by highlighting applications of my Steward Approach to different domains, notably the problem of peer disagreement and voting norms. In developing the Steward Approach, I develop tools that can be used to solve the problem of peer disagreement. And, whenever the consent of groups plays a moral role, the Steward Approach is important. It can be used to ground certain group voting structures: a just voting structure is one that, among other considerations, will respect the propositional attitudes of the democratically structured group. There is considerable fertile ground for fleshing out the implications of a solution to the attitude aggregation problem.

Acknowledgements

I am incredibly thankful to the many who have contributed to my development as a philosopher and person while I have worked on this dissertation. Mike Titelbaum is always ready with incisive questions and has taught me to see big picture, interesting narratives over focusing on true but boring details. Elliott Sober, despite using too many semicolons and converting everything into likelihood ratios, has a passion for good philosophy that is contagious and a love for his students that is heartwarming. Dan Hausman gives the quickest and most helpful written comments; he pushes me to be better. If even a tiny part of these mentors has rubbed off on me in my time in Madison, I am much better for it. That extends outside of my dissertation committee to other faculty, staff, fellow grad students, and undergraduates I have had the pleasure of working with.

If my acknowledgement of my philosophy influences is inadequate—and it is—then so much more inadequate is my acknowledgment of my family. My parents: your generosity has made my pursuit of a PhD possible. My wife Nicole: your patience, love, and support have given me the confidence and drive to be better. You have also taught me not to put the rest of my life on hold while I'm in school. My daughters: though I have just met you, you have put life in a new perspective for me. My in-laws, friends and extended family: thank you all for the incalculable contribution you have made to my life and this dissertation.

Table of Contents

Chapter 1: A Robust, Common-Sense Realism About Groups

The central aim of this dissertation is to offer a solution to the attitude aggregation problem: what, if any, relation holds between the attitudes—beliefs, desires, credences, and so on—of the group and the attitudes of the individuals that comprise the group? Pursuing that task presupposes that groups are worthy of such an investigation. Are groups real, or at least useful fictions? If so, are groups the sort of things that have propositional attitudes? In this first chapter, I will sketch a common-sense view of groups that will serve to motivate the rest of the project. To provide this sketch, I will offer several arguments that groups are (perhaps fictional) entities worth discussing. Then, I will flesh out the common-sense features of groups.

In this chapter, I will borrow from Pettit and List's (2011) taxonomy of views about groups. They distinguish four major views on groups as agents. First, there are eliminativists who hold that all talk of groups is metaphorical. Second, there are eliminativists who are error theorists; they hold that, while not all talk of groups is metaphorical, 'group-agency talk is always misconceived'. Then, there are two realist positions. Third are the 'thin' realists, who hold that group-agency talk is neither metaphorical nor misconceived. Thin realists hold that group facts are reducible to individual level facts. Fourth, the robust realists hold that group-agency talk is non-metaphorical, often true, and irreducible to individual-level talk. I favor robust realism, though for my purposes in this dissertation, either of the realist positions will do. In the arguments I offer below on behalf of groups, I will note when an argument favors the robust realist position over the thin one.

Though List and Pettit do not make this point explicitly, it is worth noting that one can maintain a variety of positions if one makes distinctions between various domains. One could be

a robust realist with respect to some class of groups but be a thin realist with respect to another. As formulated, their taxonomy is applied to the groups that one takes most seriously. That is, if you think there is any group-talk that is non-metaphorical, true, and irreducible to talk of individuals, then you are a robust realist. This is so even if you think there is other group-talk that is metaphorical or reducible to talk of individuals.

This chapter proceeds in two parts. First, I will argue for the existence of groups as entities. Our ontological picture will be incomplete without reference to groups. This alone, however, is insufficient as a defense of robust realism, since groups could be real but fall far short of being agents. Second, I will unpack some of the arguments in the first section to develop the common-sense characteristics of groups, which includes robust enough features for agency.

§1: Arguments for Groups as Entities

I will put forward three main types of arguments, a natural language argument, and argument from theories of natural selection, and an argument from our treating groups as objects of moral evaluation.

§1.1 Our Natural Language About Groups

Consider the following statements:

> "The team desperately wanted to win."
>
> "The company prefers American to foreign labor."
>
> "The New York Times believes that Trump committed treason."
>
> "The philosophy department thinks it's quite likely that funding will be cut."

Each of the above statements, taken at face value, indicates that there are groups—teams, companies, newspapers, philosophy departments. Moreover, these groups have propositional attitudes—desires, preferences, beliefs, and credences. Because these statements are so natural, the default position should be to treat groups as real entities capable of having propositional attitudes. In what follows, I will respond to a few objections to this as the default position.

The first objection to consider is that the default position should be the simplest one. Ockham's Razor might favor the eliminativist view. Positing group agents seems to be a more complex theory, and the Ockham's Razor-style approach suggests that complex theories are worse than simpler ones.[1] I have three responses to this line of objection. First, I don't put too much weight on arguments about default positions or who has the burden of proof. In the end, I will have to muster my best arguments for the conclusion that groups are worth discussing, and if those arguments are unconvincing, any time spent arguing that it should be the default position was wasted. So, even if you generally subscribe to Ockham's Razor and think eliminativism about groups is simpler, I hope my arguments will persuade you that the simpler conclusion is false in this case. Second, the eliminativist position is *not* simpler, certainly not obviously so. While eliminativists, or even thin realists, can get away without positing groups as real ontological entities, they must advance more complex theories to deal with the sorts of utterances listed at the start of this section. Plus, they will have to present alternative theories to the others that I discuss in the following arguments. To show that eliminativism is simpler, it must be simpler in the context of one's whole worldview rather than in a single dimension. When one can trade simplicity in one respect for simplicity in another, it will be difficult to determine which

[1]Sober (2015) would wish me to distinguish the Razor of Silence from the Razor of Denial here. The former would tell us to be silent about whether groups exist if a theory that didn't reference groups was simpler and adequate, the latter would tell us to deny—treat as false—the more complex theory.

approach is simpler overall. My third response to this objection is that simplicity is not always a guide to truth. Even if I grant that the robust realism about groups is the most complex theory, it does not follow that this theory is false. It might turn out that the world is just more complicated in this respect than it could have been.

Simplicity aside, why think that our natural language use favors a realist view over an eliminitavist one? We often speak metaphorically of non-agents having agency. We use language of desires about plants (e.g. "That tree wanted to be watered") without really meaning that plants have desires. A car that is slow to start on Monday might be jokingly said not to want to wake up after a restful weekend. Given that we don't want to take these sorts of assertions literally, you might argue that the group-talk listed at the start of this section is not literal either. But there is an important disanalogy between these cases. If you asked someone who had just said "Trees desire water" whether trees really have desires, I'd expect them to say 'no' and describe what they mean.[2] The person who says the car doesn't want to wake up will also deny that cars really sleep if they are pressed. But it is not so clear that people will give up claims about groups: if asked me whether a team really wants to win, I would say yes, or perhaps consider if the team had other motives,[3] before I gave a second thought to whether teams really have something like desires.

Let me consider another line the eliminativist could take, one inspired by Lakoff and Johnson' s *Metaphors We Live By* (2008). It may satisfy the constraints of eliminativism while

[2] Perhaps: "Trees require water to flourish."

[3]By this I mean that when the question is posed again I would take the emphasis to be on whether they wanted to win rather than do something else. Asking a "Did you really mean what you just said?" question has the effect of highlighting what the speaker is least confident about asserting, or perhaps what they expect their interlocutors to find most controversial. In the tree and car case, the speaker is violating a norm by personifying a non-agent. In the team case, I would be least confident about what the team's motives were, not whether teams had motives.

predicting that agents would not give up assertions about groups as agents. According to Lakoff and Johnson, metaphor plays an essential role in defining our concepts and shaping our understanding of the world. One of their central examples is the framing of argument as war: we defend our position, attack our interlocutor's views, and attempt to win the dispute. This is metaphorical insofar as we know that armed conflicts really aren't the same as verbal disagreements, but nonetheless we would not retract our assertions that you defend your position in an argument. In the counterfactual world where the metaphor for argument was dance, arguments would be very different. In such a world, taking a step back in response to an objection to your view would not have such a negative connotation.

You could raise a parallel metaphor about groups: perhaps 'groups as agents' is a metaphor that we live by in shaping our views of groups. There might be alternative conceptions of groups—maybe groups as tools,[4] watches,[5] or ships[6]—but those conceptions would lead to a very different understanding and set of practices in groups, or so the story would go. This way of thinking about groups is interesting, but it is not inherently eliminativist. Take the 'argument is war' metaphor. Lakoff and Johnson's account does not entail that there are no arguments. To the contrary, metaphors of this sort "create realities for us, especially social realities" (page 126). So, even if groups are only agents contingent on us using metaphor, it may still be the case that groups are agents. It will at least be worth exploring the metaphor more carefully. Lakoff and Johnson note that President Carter's treatment of the energy crisis as a war brought with it all sorts of entailments: the crisis then had an enemy that required us to 'gather intelligence' about

[4] Committees are made for a purpose. Groups have uses.

[5] For a group to function, all the pieces must fit together. They are functionally defined.

[6] The leader is the captain of the ship; the uniting purpose the wind in their sails; the difficulties a group faces are rough waters.

and ‘marshal our forces’ and ‘make sacrifices’ in our attempt to win the war. Similarly, if ‘groups as agents’ is a metaphor that characterizes groups, there will be entailments for social realities that are worth pursuing: how do groups act, how do they make commitments, and how do they persist through time?

This is a critical point to underscore as I argue for robust realism in this chapter. While I think robust realism is true, the chapters that follow depend on a much more modest claim: groups can be fruitfully modelled as agents whose attitudes are built up out of the attitudes of the group members. Thin realism, or perhaps even an eliminativism that treats groups as useful fictions, is compatible with this modelling thesis. So, even if you find yourself bristling at my arguments in favor of robust realism, keep an open mind about the more modest claim.

To summarize, our normal assertions about groups, taken at face value, indicate that groups are agents. This generates a defeasible preference for realism. Ockham’s Razor does not favor eliminativism over realism. It’s not clear that realism is more complex than eliminativism, and even if it is, that is not a reason to prefer the simpler theory. Then, I showed that one could hold that ‘groups as agents’ talk was metaphorical without being an eliminativist. I have not yet offered much of a positive argument from the premise of our natural language use to the conclusion of robust realism about groups. I will do that now.

My positive argument is a Quinean (1948) one: “a theory is committed to those and only those entities to which bound variables of the theory must be capable of referring…” I will be expanding on this argument throughout this chapter, as I will be providing theories from other domains that quantify over groups. However, even just interpreting the common utterances listed at the start of this section generates a case for groups as entities. The most straightforward interpretation of these assertions is to take them literally, and their literal interpretation entails

that, if the assertions are true, robust realism is true. It may turn out that such utterances are uniformly false or that we have good reason not to take them literally, but otherwise our language about groups counts against eliminativism.

In addition to the Quinean strategy, one can adopt a pragmatic strategy in favor of robust realism. If it turns out that it would be exceedingly useful and unifying for our otherwise well-supported theories for robust realism to be true, then there is at least practical reason to accept robust realism. Even if one denies that there is positive evidence for robust realism, one may nonetheless think that it's worth treating groups as agents. While I adopt a realist view, someone who treats groups as useful fictions may still accept most of what follows in this dissertation.

§1.2 Theories of Evolution

Let me extend the Quinean argument by looking at theories of natural selection that, if true, require realism. We should accept the entities entailed by our best theories. 'Best' is a tricky term, but I don't need to take too strong of a stand here. I will merely argue that there are some compelling theories of the world that use groups as an integral part, and this gives us strong reason to think that groups exist. My argument is even stronger for the claim that models of groups as agents are useful.[7]

An important question in the philosophy of science is the units of selection problem: what things are subject to natural selection? There are many possible answers: genes, individual organisms, groups of organisms, species, ecosystems, and so on. For my purposes here, I will focus on theories of natural selection in which groups play an irreducible role.

[7] Robust realism doesn't entail the claim about modelling, since it could be that groups are real but very uninteresting.

Altruism poses serious problems for theories that posit individual organsims as the sole units of natural selection. Altruists benefit other organisms at a fitness cost to themselves. But, this generates a free-rider problem wherein selfish organisms will dominate the altruists in any given population, and this dominance would send selfishness to fixation. But, we *do* seem to see altruism in a variety of populations. This is a puzzle. Some have responded to this puzzle by holding to versions of egoism—that individuals are only selfishly motivated—and then they deny that apparent cases of altruism are actually cases of altruism.

Sober and Wilson (1999) propose a solution to the problem of altruism that treats groups as centrally important. They claim that, while altruists will always fare worse than selfish individuals in a given population, groups with higher concentrations of altruistic organisms may fare better than groups with higher concentrations of selfish organisms. Suppose there are two groups, G_A and G_S, where G_A is 90% altruistic and G_S is 90% selfish. After a few reproductive cycles, the selfish individuals in each group will be doing better than the altruists in each group. However, on the whole, the altruist group G_A will be doing better than the selfish group G_S, and this means that the frequency of altruism is globally increasing. This is an instance of Simpson's (1951) paradox. Even though altruists fare worse than selfish organisms in every group, they can do better overall because they are overrepresented in the more successful groups.

Natural selection requires inheritance and variation of fitness. So, to the extent that groups play a role in either of these features, groups are part of any complete natural selection explanation for the biological diversity that we see today. In the case of altruism, the way individuals combine and recombine into groups is essential. Even though altruistic groups fare better than selfish ones, selfishness will rise to fixation if the group is held fixed (i.e. all of an organism's descendants remain in the same group as the organism). This doesn't always happen,

though. Often, a group's membership will change, and in these cases altruism can thrive when altruists group together.

In the conclusion of their book, Sober and Wilson talk of the importance of groups:

> From ancestors who were at best only moderately well-adapted at the group level, our lineage has evolved so that individuals participate in social groups that sometimes invite comparison to bee hives and single organisms…Human groups are like single genomes, which achieve their unity by being organized to prevent subversion from within as much as possible. It is a great irony that the language of human social control—sheriffs, police, parliaments,…,etc.—has been borrowed to describe the social behavior of genes, without the reciprocal conclusion being drawn that human social groups can be like genomes…Group selection has not been the only force in human evolution…but in all likelihood it has been a tremendously powerful force.

I will return to these ideas later in the chapter when I talk about what features groups have on our common-sense conception of them. For now, I am arguing that groups are entities on their own, irreducible to the individuals that comprise them. Sober and Wilson make a strong statement in line with robust realism: not only are they clear that group selection is explanatorily powerful, but they also argue that purely individual level selection cannot adequately shed light on 'the entire stage' of biological traits that emerge from natural selection.

Another view that examines the importance of groups in human evolution is dual-inheritance theory. Dual-inheritance theorists hold that human behavior is the product of *both* individual-level genetic and group-level cultural evolution.[8] To borrow and example from Henrich and McElreath (2007), the Bari people of Venezuela believe that a child can have multiple fathers. This is based on the belief that paternity is proportional to the number of sexual encounters with a 'father'[9] after the mother is pregnant. This leads many pregnant women to pursue extra-marital encounters, since having more than one 'father' is advantageous to the

[8] See Henrich, Joseph, and MCElreath (2003), Boyd (1985), and Richerson and Boyd (1976).

[9] I will use quotes when I am using the Bari conception of father.

child. Indeed, it turns out that "the optimal number of fathers for a child's survival is more than one". Having more than one 'father' is a boon to one's fitness (note: more fathers need not always be better, but clearly two fathers are better than one in Bari culture). Thus, over time, natural selection has favored Bari people with more than one 'father'. Or, alternatively, natural selection has favored Bari people who pursue multiple sexual partners when pregnant. But, it is not a purely genetic process: contra the Bari's beliefs, every child still has the genetic input of only one male. The cultural situation in which the Bari find themselves is essential to the story we tell about this selective process.

There are at least two ways to characterize the role culture is playing in cases of natural selection. First, we can characterize the culture as a unit of selection. Culture is inherited, provides a selective advantage for groups of individuals, and thus can change over time in ways analogous to the ways individuals change. This would be supported in the Bari case if Bari cultural practices and beliefs made Bari people fitter than other neighboring cultures. Second, we can characterize the culture as the environment in which individual Bari organisms find themselves. Thus, the practice of beliefs about paternity and social practices that go with this (e.g. giving food and support to one's 'children') is analogous to harsh temperatures that may promote the evolution of a thick fur coat. In either case, groups play a significant role from the perspective of dual-inheritance theory. In the Bari example, culture at minimum plays the role of environment; I do not know whether the cultural practice also favors the Bari over other competing cultures.

Henrich and McElreath (2003) argue that culture itself emerged as a result of natural selection. They discuss the importance of distinguishing markers for ethnic and social groups. The social nature of humans leads us to participate in many coordination problems. Coordination

is more successful when participants share norms and values. Over time, repeated interactions lead to well-defined and enforced groups. On this view, social groups themselves have an evolutionary explanation. Hencrich and McElreath sum up the need for culture in a theory of human psychology as follows:

> "...explaining many important aspects of human psychology and behavior will require examining how genes under the influence of natural selection responded to the regularities produced by culture. This means that understanding the behavior of a highly cultural species like humans will sometimes demand a culture-gene coevolutionary approach."

This framing emphasizes the environmental nature of culture. Cultural groups provide a consistent backdrop in which certain traits emerge.

In this section I have looked at the role that groups play in a few theories of natural selection, those of Sober and Wilson and dual-inheritance theory. In both cases, a purely individual-level analysis fails to provide the full explanation for the explananda, be it the emergence of altruism or the behavior of the Bari. This gives us strong reason to treat groups as real entities. In §2, I will discuss what these natural selection arguments about groups tell us about the nature of groups.

§1.3 Groups as Moral Agents

One common argument against groups being real entities trades on the intuitive absurdity of treating corporations as legal people. Take the following quotes, for example:

- "I don't care how many times you try to explain it," Barack Obama said on the stump. "Corporations aren't people. People are people."[10]

[10] See Greenfield (2015).

- "Most people who hear [the] phrase ['corporations are people'] for the first time scratch their heads, look at you funny, and maybe wonder what you've been smoking. The absurdity of corporate personhood has that effect on people — it just doesn't make sense!"[11]
- Stephen Colbert: 'Mitt Romney claims to be pro-corporations, but would you let him date your daughter's corporation?'

It is hard to deny that there is something jarring about the idea of corporations as people. But, on the other hand, there is an even more plausible claim that corporations are subjects of moral evaluation. And it follows from this latter claim that groups are moral agents.

Before I proceed, two clarificatory issues. First, while this section resembles the natural language argument in §1.1, my argument here is not limited to the most plausible interpretation of our talk of groups. Rather, I will make use of natural-sounding expressions to underscore that there is widespread (and proper) recognition of groups as morally evaluable. Second, in this section I will move freely between talk of corporations and talk of groups. One might wish to separate out legal entities like corporations and businesses from groups in the philosophical sense that we have considered so far. However, the business ethics literature on group agency frames the question as the problem of "corporate moral agency,"[12] and corporations/businesses are merely taken as paradigm instances of groups. I will revisit this distinction in one of the objections I consider below, but otherwise my arguments do not hang on any equivocations with respect to 'corporation'.

[11] See Morgan (2001).

[12] See Moriarty, Jeffrey (2017).

Let's begin with the legal status of corporations. In *Citizens United v. Federal Election Commission*, the Supreme Court reversed the decisions of *McConnell* and *Austin*, which permitted limited restriction of corporate speech. Justice Kennedy delivered the opinion of the court, and his discussion of groups and corporations clearly puts them on a par with individual persons: "The disclaimers required by §311 'provide the electorate with information'…and 'insure that the voters are fully informed' about the person or group who is speaking".[13] This equation of persons and groups received backlash in both the dissenting opinions and public opinion. To many, it is ludicrous to say that corporations are people. This might suggest that such dissenting views take groups less seriously, but that is not the case. Consider: "Suppose that General Motors Corp., troubled that a candidate for Congress from Michigan was too favorable to the United Auto Workers, decided to do everything in its corporate power to defeat that candidate."[14] Even here, the group General Motors Corp. is "troubled" and "decided" to act in light of that attitude. While some are clearly skeptical of the role that groups should play in the political process, the common view of groups is still that they are agents.

Still, legal practices need not perfectly reflect best ontological practices. If the Supreme Court referred to 'grue' in a brief, that would not make the predicate any more natural.[15] In the case of groups, however, there is strong moral reason to follow the Supreme Court's suit in awarding them agency. Corporations, and groups more broadly, are often treated as subjects of moral evaluation. If we are to take these evaluations seriously, then groups must be awarded the status of agents.

---

[13] Page 52 of *Citizens United* (2010).

[14] Lyle (2010).

[15] See Goodman (1955).

To support my claim that groups are subjected to moral evaluation, consider the following claims:

- Exxon is to blame for the oil spill in Alaska.
- The church was praiseworthy for its services to the homeless.
- The Red Cross has violated an implicit agreement with donators by improperly managing the funds donated for hurricane relief.

Prima facie, these sentences indicate that groups can be subject to moral evaluations in the same way that individuals can be. It is just as natural to say the following individual-level claims:

- Elizabeth is to blame for the harms Juliette sustained.
- Nicole is praiseworthy for her service to the homeless.
- Juliette has violated her implicit agreement with me.

In what follows I will unpack some necessary conditions for moral agency, and then I will defend my argument for robust realism from objections.

Moral agency entails that the agents are able to make commitments and act in such a way to fulfill them (or violate them). It would be absurd to attribute blame to someone or something incapable of making commitments or intentional action. If a tree falls on your house, you might feel anger, but you certainly shouldn't blame the tree for falling; the tree did not intentionally fall onto your house. Similarly, my wife was furious with the squirrels who chewed through our Christmas lights, but it would have been an error for her to accuse the squirrels of acting immorally. Squirrels might be capable of intentional actions, but they are incapable of making morally binding commitments.[16] Are we making a similar sort of mistake when we blame Exxon,

[16] In contrast, George Constanza firmly believes that there can be morally binding commitments with pigeons in *The Merv Griffin Show* episode of Seinfeld. Jerry agrees, explaining that "Of course we have a deal: They get out of the way of our cars; we look the other way on the statue defecations."

praise churches, and criticize charities? Robust realists can say 'no': groups are moral agents. As such, it is proper to criticize (or celebrate) their failure (or success) to act in accordance with their commitments. Eliminitivists must argue that moral assessments of groups are metaphorical or false. Let us consider two ways eliminitivists might attempt this.

The eliminitivist may adopt expressivism with respect to moral appraisals of groups. Thus, they might interpret "The church is blameworthy" as expressing a negative emotion towards the church, its members, or some closely related topic.[17] The problems with expressivism with respect to the moral agency of groups carries over from problems with expressivism in the moral domain more broadly. On an expressivist view, moral claims are neither true nor false, but instead express the speaker's feelings. This runs into a number of problems. Moral discourse does not seem to be merely emotive: there are cases where one can be convinced of moral arguments dispassionately. And, if expressivism is correct, then moral claims are neither true nor false, and thus cannot be the premises in an argument. If you find expressivism implausible generally, you should find it implausible in the case of groups.

However, even if you found expressivism palatable more generally, this is still not a problem for my argument for robust realism. I am claiming that groups are argents in roughly—I will qualify this in §2—the same way that individual persons are agents. So, if one is an expressivist about morality at large, then being an expressivist about moral appraisal of groups still puts groups on a par with individual moral agents. The only serious objection to my view would be someone who held a non-expressivist view about most moral claims, but was an expressivist about moral claims concerning groups. I am not sure how one would motivate this

[17] E.g. one might criticize the church in the context of it being tax exempt to show their displeasure for the current tax code.

distinction other than by begging the question in this context (that is, by arguing that since groups are nothing like moral agents, moral evaluation of groups must be importantly different from moral evaluation of individual moral agents). It certainly seems to me that blaming Exxon for an oil spill is the same type of thing as blaming a criminal for robbing a bank; both feel like moral evaluations of agents who have violated their moral obligations. If so, our moral treatment of groups entails something very close to robust realism, and it certainly undercuts eliminativism.

The second objection is that moral judgments of groups reduces to moral judgments of at least some of the individuals that comprise the group; call this the *reductionism objection*. For example, (1) can be understood as a moral judgment of the CEO, or the captain of the boat that spilled the oil, or, more generally, simply shorthand for "Those at Exxon who are responsible for the oil spill are to blame." This would accommodate my claim at the end of the last paragraph: clearly our moral blame of individuals and groups will feel the same if group-blaming is just individual-blaming in disguise. This objection falls prey to counterexamples.[18]

The first set of counterexamples to the reductionism objection is that we often blame or praise groups for things that the individuals that comprise them are incapable of performing.[19] For example, we might praise the NAACP for advancing civil rights in the United States of America over the last 108 years. But, no single individual has been a part of the NAACP for that entire period of time. Since a group may persist despite membership changes, groups are capable

18 Boston College (2011) points to a negative correlation between the cohesion of a group and the blame we assign to group members. This seems to be evidence for my claim here, though it could just be a moral illusion analogous to a sensory illusion.

19 Cooper, David, 1968, "Collective Responsibility," *Philosophy*, 43: 258–268.

of things that individuals are not. French (1998)[20] extends this line of argument, pointing out that groups can elect a president, win or lose a football game, and perform other group level actions. It is much more difficult to reduce these actions to the individual level. To compare, the reductionism objection, if successful, would need to show that blaming Americans for electing Trump would have to reduce to blaming individuals for voting (or not voting) as they had in the last election.[21]

While I have rejected the objections that call for a different treatment of moral evaluation of groups, I am open to there being minor differences. As pointed out in the second objection, there are things that groups are capable of that individuals are not, and the other direction holds as well. Individuals can, as Colbert pointed out, get married, whereas groups are incapable of doing this. So, it is possible that we will need at least slightly different conceptions of what is going on when we praise a group. For this reason, I hold the more conservative view that groups are *roughly* moral agents. Groups must still be capable of making commitments, representing the world, and acting based on those commitments and that representation. Otherwise, moral evaluation would be absurd. But this doesn't mean that group agents are exactly the same as individual agents. I will continue this discussion in §2.

So far, I have considered the role that groups may play as moral agents, but groups may also play roles as moral subjects. Now I will turn to whether we have duties regarding groups. Weijer, Goldsand, and Emanuel (1999)[22] discuss the proper guidelines for research protection for

---

20 French, Peter, editor, 1998, *Individual and Collective Responsibility*, Rochester, VT: Schenkman.

21 This is perhaps too quick. The reduction might also include individuals acting as they have to leave the country in the state in which the individual votes being what they were would contribute to Trump's election. In any case, the reduction must be solely executed on the individual level.

22 Weijer, Charles & Goldsand, Gary & J. Emanuel, Ezekiel. (1999). Protecting Communities in Research: Current Guidelines and Limits of Extrapolation. Nature genetics. 23. 275-80. 10.1038/15455.

communities in *Protecting Communities in Research: Current Guidelines and Limits of Extrapolation*. Weijer et al criticize existing methodologies, notably the Belmont Report, for only focusing on individuals rather than communities. The goal of their paper is to extend the theory of ethical treatment of groups beyond the ones that were used by their contemporaries in 1999. While they take it as a given that groups are objects of moral concern, they note there will be great difficulty in determining the consent procedures for certain groups: "There are substantial problems with applying protections developed for aboriginal populations to other less cohesive communities, especially ones without legitimate political authorities." I will leave this as an additional condition of adequacy for my theory of groups developed throughout this dissertation; my account will be more successful if it can provide resources for ethicists to determine consent (among other such ethical concerns) for groups that lack clear governmental structures.

Hausman (2007) also considers the ethical repercussions for causing harms to groups. He distinguishes two types of harms: harm to individuals in virtue of social group membership ("group mediated harms to individuals"), and harms to structured groups ("harms to groups"). An example of the former is violence done to individuals in virtue of their religious affiliation. If Muslims are subjected to additional searches at airports, this directly harms individuals, but it only does so in virtue of their membership to a group. Thus, Hausman terms these harms group-mediated harms to individuals. On the other hand, certain groups may be harmed directly: when Shaquille O'Neal and Kobe Bryant feuded, it harmed the Los Angeles Lakers basketball organization. While this harm also trickled down to other individuals on the team, it hurt them in virtue of the team being harmed. Contrast this with the case of airport searches: there individuals

are the primary sufferers of harm. Hausman unpacks the required structure that a group must have to suffer such harms, and I will return to this in the next section.

For now, I want to focus on how his account rejects eliminativism. He shows that there are at least two ways in which groups are relevant to moral concerns. In the case of group mediated harms to individuals, only individuals are directly harmed. This may seem to support an eliminativist view, but it does not. Even in these cases, the individual suffers harm because of some belief on behalf of those committing the harmful actions that attributes group membership to the individual harmed. In the case of additional searches at airports, those performing the searches are doing so based on a perception of the harmed individual's belonging to the Muslim community.[23] There is an even stronger case to be made for groups in the second sort of harm, harm directly to structured groups. If groups are indeed the sorts of things that can be morally wronged, then they must have interests and moral standing. While this does not entail agency, it is sufficient to reject an individual-only conception of reality. Hausman treats eliminativism as a silly position:

> "As sports fans everywhere will attest, sports teams exist….The philosophical task is to understand how groups such as teams can exist and suffer harms, rather than supposing that their existence is some weird metaphysical doctrine defended only by philosophical cranks." (page 361)

He takes group existence as a given, and then proceeds to consider the best overall theory into which that fits.

I have argued that groups are both moral agents and moral subjects. We rightly praise and blame groups. This means that groups must be at least roughly like individual agents. Moreover,

[23] Hausman points out that this perception need not be veridical. He recalls his father being removed from a "Whites-Only" restroom in the Miami airport. His father was, in fact, white, but he had tanned quite a bit on their Florida vacation.

groups are subjects that can be harmed, and they can feature in the explanations for how individuals are harmed. Fleshing out the necessary conditions on groups to fulfill these roles will be an important task in the next section for building a common-sense picture of groups.

§2. A Common-Sense View of Groups

Just how similar to individuals are groups? If groups are *roughly* agents, then how rough is the approximation? In this section, I will combine features of the various arguments in favor of realism to support robust realism. The resulting common-sense view of groups will serve as a test for my solution to the attitude aggregation problem.

The arguments from natural language support groups being the sorts of things that have propositional attitudes. The expressions at the start of §1.1 feature groups that have propositional attitudes. Moreover, these propositional attitudes bear some relation to the attitudes of the individuals that comprise the group. It would be silly to say that "The team wants to win, even though nobody on the team wants to win." But, these propositional attitudes are not just universally applied to each group member; it is not absurd to say that "The team wants to win, even though the backup quarterback doesn't want the team to win." We treat groups as agents that are built up out of, but not reducible to, individual group members.

Similarly, it does not follow from every group member having some attitude that the group has that attitude. Suppose, for instance, that every member of the NAACP preferred chocolate ice cream to vanilla. It does not follow that the NAACP prefers chocolate to vanilla. In fact, some would find it very strange to say that the NAACP had any ice cream preferences at

all.[24] Groups need not have attitudes toward every proposition that its members have attitudes toward. Instead, there will be some fact about the nature of the group that will determine which propositions are 'considered' by the group rather than merely by its members. In many cases the purpose of the group will help fix these propositions. The NAACP will have beliefs related to the advancement of racial equality. A baseball team will have beliefs about whether they will make the playoffs, and about their preference of playing in cold weather.

The arguments from theories of evolution showed that groups are sometimes units of selection and sometimes environments in which selection can occur. As units of selection, groups must be capable of inheriting and passing on traits. This capacity entails certain things about a group's persistence conditions, but we will look more at those later. In the evolutionary context we learn the most about groups by seeing how a purely individual-level story is insufficient. Sober and Wilson defend group selection by considering the selfish gene theory. The selfish gene theory is championed by Dawkins in his 1976 book. This is the theory that posits genes as the sole explanatory unit of selection, as they are the (only) things capable of faithfully reproducing themselves. This line of argument was used against group selection: groups are unable to replicate themselves accurately, and genes are driving the bus anyway. But, as Sober and Wilson point out, individual level selection is subject to the same argument: individual organisms, like groups and unlike genes, are not replicators. So, to the extent that one thinks individuals can coordinate their parts as vehicles of selection, one can make a parallel argument that groups can, too. Groups play roles as units of selection insofar as they are

[24] Dan Hausman, to whom I owe this example, finds it incredibly strange to say that the NAACP likes chocolate ice cream in this case. I do not have as strong an intuition, but I do think there is an important difference between the stance of the NAACP on, say, tax policy compared to whether it has preferences with respect to desserts.

harmonious enough to serve as ‘vehicles of selection’. This rings true: collections of individuals that are not unified are not groups. This holds for collections of individuals that were never unified (pick a random person from each continent: they will likely not form a group) or groups that were once unified but have fractured (like a political party that splits over a core disagreement).

Dual-inheritance theorists maintain that groups can have mechanisms for identifying and enforcing membership, as well as a series of norms that facilitate advantageous interactions for its members. Such groups will be characterized by their practices. This is how you would observe the dominance of a cultural group: their practices spread. Members of a certain group, then, must bear some relationship to a group’s indicative practices. This is a why a non-practicing Catholic may not be considered a Catholic at all, but perhaps would be better attributed as belonging to the group of non-practicing Catholics. This also highlights that groups are different from their members, in that a group member need not practice all of the characteristic activities of the group. A church member may not believe every tenet of his church’s theology, an athlete may miss a practice, and a member of the executive board may disagree with some of his company’s actions. Groups are coherent wholes, but their parts need not always pull in the same direction.

The argument from groups as moral agents entailed that groups are capable of making commitments and acting in accordance with those commitments. This means that groups must exist in time, at least long enough to satisfy the sorts of commitments that they are capable of engaging in. Suppose the NRA promises to bring it about that legislation is passed that closes the gun show loophole. If we take this promise seriously, this entails that we think the NRA will

exist for long enough to do this, and it also entails that we think the NRA is capable of doing things that might result in such legislation being passed.

Lastly, let us return to Hausman's conception of structured groups. The central claim is that in order to be harmed, groups must have interests that can be satisfied or thwarted. He provides following four necessary conditions for being a structured group:

1. Structured groups must exist over time and be reidentifiable.
2. Structured groups must not be merely identified by its membership or some common feature(s) that each member shares.
3. The organization of the parts of the group is an important part of the group's identity.
4. Individuals acknowledge and interact with structured groups.

As necessary conditions, these serve to show which collections of individuals do not form structured groups. Condition (1) eliminates fleeting indexical groups, such as "the people on this bus right now". Such a group will fail other conditions as well, but it will fail (1) by disappearing too quickly to possess any lasting interests that can be satisfied or violated. Condition (2) precludes collections like "the brunettes in the northern hemisphere" from being considered a structured group. Were we to consider that a group, harms to the group would easily reduce to harms to individuals, and thus the group is unnecessary. Similarly, Hausman thinks condition (3) prevents structured group harms from reducing to harms to individuals. I find condition (3) to be less than perfectly clear, but I take it that form of government is a paradigm case: shifts from democracy to socialism would radically change the nature of a structured group. If a group were not characterized by something like the organization of its parts, there would be a danger of the group again just reducing to its members. Lastly, condition (4) is more of a highly reliable heuristic for identifying structured groups than it is a necessary condition for them. Hausman

admits that he has no argument for (4) being a necessary condition, rather he cannot think of a counter-example. I can't either. Structured groups are the sorts of things that I have been referencing throughout this chapter: teams, companies, nations, committees, religions. Individuals don't take seriously ad hoc collections of individuals like lefthanded people that live in your neighborhood.

One might choose to call any collection of individuals a 'group', and I have no principled objection to this. However, in this dissertation I will be interested in characterizing the attitudes of a special subset of these collections that I have sought to pick out by my common-sense vie of groups. From here on out, I will be using 'group' only to refer to the collections of people that have the following properties. Groups:

- persist through time, though are subject to change,
- have propositional attitudes,
- are capable of acting,
- are subject to moral evaluation,
- and can be harmed.

Further, these groups are not:

- reducible to the individuals that comprise the group, or
- determined by just any set of individuals.

Thus, the common-sense view of groups is a robust realist one. It treats assertions about groups as non-metaphorical, true, and not reducible to facts about the individuals that comprise the group. On this conception of groups, the collection of all five-year-old children does not constitute a group: they do not form an agent capable of action, subject to moral evaluation, and so on.

On this common-sense picture, a tenure committee can be said to have views about whether some professor should receive tenure. And these views can be said to be rational (or irrational). It turns out, though, that there are problems facing group rationality that are not present on the individual level. I turn to these problems in Chapter 2.

Chapter 2: The Attitude Aggregation Problem for Democratically Structured Groups

In this section I will lay out the problem of attitude aggregation. Roughly the problem is to find an aggregation function, a function that takes as inputs the attitudes of a set of individuals, combines these attitudes, and yields a group attitude as an output. Once this problem is on the table, I will present some of the most famous impossibility results, which show that one cannot satisfy a set of plausible constraints on the aggregation function. Then, I will argue that these results do not seriously threaten the common-sense picture of groups laid out in chapter one. The impossibility results do, however, provide a motivation for tying the attitude aggregation problem to the traditional epistemological task of combining evidence. I will conclude by refining the problem for democratically structured groups, what I take to be the most interesting case.

Call F(.) the function that takes as inputs group member attitudes and yields as output group-level attitudes. For precision, here are some definitions:

- We will be interested in a group **G**, composed of members g_1, g_2, …, g_n.
- We will be interested in some set of propositions, p_1, . . . p_n; call it **P**. **P** is also called the **agenda**.
- We will be interested in some set of attitude types **A**, composed of A_1, A_2,…, A_n.
- Attitude type A is **binary** if it has two values. I call these values "yes" and "no".[25]
- Attitude type A is **ternary** if it has three values. I call these "yes", "no", and "neutral".[26]

[25] E.g. judgment: One can judge a proposition to be true ("yes") or false ("no").

[26] E.g. belief: One can believe ("yes"), disbelieve ("no"), or withhold belief ("neutral").

- Attitude type A is **n-ary** if it has n values, where n is a natural number.
- Attitude type A is **continuous** if its values can be placed in correspondence with [0, 1].[27]
- Call a group member g's set of attitudes towards each proposition in the agenda g's **attitude set.**
- An attitude set is **complete** just in case there are no empty entries, **incomplete** otherwise. A set would be incomplete if a group member failed to have some attitude (in **A**) towards some proposition (in **P**).[28]
- An attitude set of some member is **consistent** just in case it meets some rationality constraints for that attitude. More will be said on this later. But, for instance, a set of beliefs must be such that it is possible for them to be all true. A set of credences must satisfy the probability axioms. And a set of preferences[29] must be transitive, asymmetric, and irreflexive.
- Call the collection of attitude sets of the members of **G** the **profile**. You can organize the profile by placing the attitude sets of each group member in rows of a table like so:

Profile	p_1	p_2	…	P_n
g_1	$<A1_{g1}(p_1)$, $A2_{g1}(p_1)$ …>	$<A1_{g1}(p_2)$, …>	…	$<A1_{g1}(p_n)$, …>
g_2	$<A1_{g2}(p_1)$, …>	$<A1_{g2}(p_2)$, …>	…	$<A1_{g2}(p_n)$, …>

[27] E.g. credence: one can assign any credence in [0, 1] to a proposition.

[28] This might occur if some group member just hadn't considered some proposition, for instance.

[29] One can represent preferences by writing them into propositions and considering agent's judgments about those propositions. I will pursue this strategy in chapter 5.

…	..	..	…	…
g_n	$<A1_{gn}(p_1)$, …>	$<A1_{gn}(p_2)$, …>	…	$<A1_{gn}(p_n)$, …>

Note that you could just as easily construct new profiles for each attitude, in which case you could have singletons in each cell. I just prefer to pack more elements into a single table rather than having multiple simple tables. That said, most of the examples I consider will be for just one attitude, so this choice won't have much impact.

- The **aggregation function F(.)** takes as input some profile and yields as an output the **group attitude set**, the attitudes of the group towards every proposition in the agenda.

It will be useful to look at subsets of the profile:

- Call the **p-subset** of a profile set of attitudes that its members hold toward the proposition p. For instance, the p1-subset of ***Profile*** above is the set of entries in the p_1 column.
- Call the **A-subset** of a profile the profile that only includes the entries for attitude A. For instance, you might have a profile that includes beliefs and credences over some agenda. The credence-subset of that profile would be the profile that only includes the credences.
- Call the **M-subset** of a profile the profile that only includes group members M, where M is a subset of G. If M is a singleton set {g}, then the M-subset will just be g's attitude set.
- Any two (or more) of these subsets **match** just in case the subsets are permutations of each other. For instance, a p-subset and p*-subset match just in case there are the same number of each attitude value in the columns on the profile table underneath p and p*, respectively.

With these definitions on the table, let's see how this works in an example. Consider a committee composed of three members: Abe, Ben, and Carol. They are tasked with determining whether

Darrin is making satisfactory progress in the graduate program. They are considering the following set of propositions:

d: Darrin is making satisfactory progress in the graduate program.

c: Darrin is completing his coursework satisfactorily.

t: Darrin is teaching satisfactorily.

(c&t) →d : If Darrin is teaching satisfactorily and completing his coursework satisfactorily, then he is making satisfactory progress in the graduate program.

The agenda is {d, c, t, (c&t) →d}. Further, let's suppose we're only interested in one attitude, belief. The following is a profile table for this profile, call it Profile ABC:

Profile ABC	c	t	(c&t) →d	d
Abe	<yes>	<no>	<yes>	<no>
Ben	<yes>	<yes>	<yes>	<yes>
Carol	<yes>	<no>	<yes>	<yes>

Profile ABC indicates that Abe disbelieves d, but Ben and Carol believe d. You might also notice that Carol believes that d even though she does not think both c and t are true. It is easy to add more attitudes to this table. If for instance, we wanted to add credences to the profile, then we would insert an extra entry after the "yes/no" for belief. For instance, suppose that Abe believed that c was true and had credence .85 that c was true. Then the Abe/c cell—the cell in row 2 and column 2—would read <yes, .85>. The aggregation function takes the information in the profile table and returns a group attitude. It can often be helpful to present this data in an **augmented profile table**, which just adds an extra row at the bottom of the profile table that includes the group attitudes across the agenda. Like so:

Aug-Profile ***ABC***	c	t	(c&t) →d	d
Abe	<yes>	<no>	<yes>	<no>
Ben	<yes>	<yes>	<yes>	<yes>
Carol	<yes>	<no>	<yes>	<yes>
Group	<yes>	<no>	<yes>	<yes>

In this case, I used an intuitive aggregation function, majority voting. I will say more about this function and some alternatives shortly.

The aggregation function for G, $F_G(.)$, takes a profile as its input and yields a set of attitudes to every proposition on the agenda. Since profiles can include multiple attitudes—credences *and* beliefs, for instance—it may be helpful to decompose $F_G(.)$ into $f_{GA1}(.)$, $f_{GA2}(.)$, and so on, where each $f_{GAi}(.)$ is the aggregation function for G with respect to just attitude A_i. Note that this decomposition only makes sense if the aggregation functions are independent; I will argue in chapter 4 that this is not always the case. This is a more general treatment than elsewhere in the literature, since most aggregation theorems are with respect to a single binary attitude, usually judgment or preference. In the short term, there will be no real difference, since I will be presenting interpretations of these single attitude theorems. However, as my dissertation progresses, I will consider more robust profiles that include multiple attitudes, at which point this notation will prove useful.

At this point, let me flag a problem that I will not be tackling in this dissertation: how do we restrict the agenda to all and only the propositions relevant to the group. As discussed in chapter one, it does not follow from every NAACP member's liking chocolate that the NAACP

likes chocolate. Instead, we should limit the agenda to only the propositions that the group can be meaningfully said to have attitudes about. This is a difficult and interesting problem, but how we restrict the agenda in this manner turns out not to affect the details of the Steward Approach that I advocate in chapter four. So, I will leave this problem to the side.

In *Aug-Profile ABC*, I made use of the intuitive aggregation function of majority voting. To get a feel for aggregation functions, consider how each of the following functions would aggregate the beliefs of group ABC[30]:

- **Majority Voting:** The group attitude toward any proposition is whatever attitude the majority of group members take toward it. In the case of ties, the group expresses a 'neutral attitude', such as withholding belief.[31]
- **Unanimity:** The group attitude toward every proposition is neutral unless the group members unanimously agree, in which case the group matches the unanimous attitude.
- **Always Believe:** No matter what the attitudes the group members take, the group attitude is to affirm every proposition—that is, believe/prefer/desire/etc. it.
- **Dictatorial Rule:** This rule singles out one member of the group and sets the group attitudes to perfectly match that member.

Aug-Profile ABC	c	t	(c&t) → d	d

[30] The following functions mostly apply to binary or ternary attitudes. Later I will consider functions that are more general.

[31] I treat belief as a ternary attitude: believe, disbelieve, withhold belief. Similarly, you could treat preference as ternary: prefer, disprefer, withhold preference. If we wanted to maintain a binary version of majority voting, we could adopt a "tie-goes toward belief" (or disbelief) policy, or some other way to break the tie. Or, we could restrict the rule to odd-numbered groups of individuals with complete attitude sets.

Abe	<yes>	<no>	<yes>	<no>
Ben	<yes>	<yes>	<yes>	<yes>
Carol	<yes>	<no>	<yes>	<yes>
Majority	<yes>	<no>	<yes>	<yes>
Unanimity	<neutral>	<neutral>	<yes>	<neutral>
Always Believe	<yes>	<yes>	<yes>	<yes>
Dictator (Abe)	<yes>	<no>	<yes>	<no>

The above table shows how each of these aggregation functions would combine the attitudes of Abe, Ben, and Carol. Each function yields a different set of group attitudes over the agenda.

The number of aggregation functions, even for single binary attitudes on small agendas, is staggering. In the ABC case that we are considering, there are four propositions in the agenda, and the group will take one of three attitudes (yes, no, or neutral) to each. That leaves $3^4 = 81$ possible group attitude set outputs. But, the number of aggregation functions is much greater than this, since multiple functions can yield the same output. Even if there were just a single proposition, a ten-person group would have 2^2^10 aggregation functions on a binary attitude (List and Pettit, 2011). This number only grows as we consider profiles covering multiple attitudes, larger agendas, larger group sizes, and non-binary attitudes.

Fortunately, we do not need to deal with all of these possibilities when tackling the problem of attitude aggregation. Some functions are obviously poor candidates for determining group attitudes. For instance, the Always Believe function does a terrible job, both descriptively and normatively as an aggregation method. This function fails in at least two ways. First, the

Always Believe function is not responsive to the individuals that comprise the group. The standard way to think about the problem of aggregation is to consider how we build the group attitude out of the individual ones, but Always Believe makes no reference at all to the individuals in the group. Second, the Always Believe function guarantees inconsistent group attitudes whenever the agenda includes any set of inconsistent propositions. For, if there is such a set of inconsistent propositions, the group will believe all of them, and thereby the group will have beliefs that cannot all be true. We might ultimately decide that groups have some relaxed standards of rationality from those of individuals, but surely the level of irrationality entailed by the Always Believe function is too high. So, the lessons to take from Always Believe? Our aggregation function should take seriously the attitudes of the individuals in the group, and it shouldn't lead to inconsistency on a large class of agendas.

The Dictator rule, and Dictator (Abe) in particular, suffers from one of these two problems. It fails to take seriously the attitudes of all of the members of the group. Of course, some groups do function as dictatorships, in which case a dictatorial rule will be successful in determining the group attitude.[32] This is a fairly uninteresting case, though: not all group structures are such that the group attitude is so easily determined. As I will lay out at the end of this chapter, I am interested in groups where every member has equal standing. Graphically, this means that swapping the names or rows on a profile should not affect the aggregation function of that profile.

But what about the majority rule? This is a very plausible view, especially given that I am interested in democratically structured groups. What is a democratic group if it isn't

[32] Though it won't matter for determining a group's democratically structured attitudes, as defined at the end of this chapter.

characterized by majority voting of its members? Unfortunately, things are not so easy. This is where the impossibility theorems come in. There is a long tradition, dating back at least to Condorcet (1785), made famous by Arrow (1963), and having its contemporary roots in List (2001). List's original paper has been extended and generalized several times, and the version that I adapt here combines threads of many of those proofs, incorporating my terminology. All of these impossibility results aim to show that a set of plausible constraints on the aggregation function cannot be met. List calls this jointly inconsistent set of constraints the Discursive Dilemma (or sometimes the Doctrinal Paradox). To motivate the inconsistency, consider the following augmented profile:

Disc. Dilemma	c	t	(c&t) ↔ d	d
Abe	<yes>	<no>	<yes>	<no>
Ben	<yes>	<yes>	<yes>	<yes>
Cecil	<no>	<yes>	<yes>	<no>
Group	<yes>	<yes>	<yes>	<no>

This augmented profile is just a modification of the views of group ABC. Here, all three members agree that if Darrin meets his coursework and teaching goals if and only if he is making satisfactory progress. However, Ben finds fault with Darrin's teaching and Cecil finds fault with his coursework. This means that two of the three find Darrin to not be making satisfactory progress. But, this poses a problem for the majority voting rule that's been applied in the augmented profile: a majority of the committee finds his coursework to be satisfactory, and a majority also finds his teaching to be satisfactory. This means that majority voting yields an inconsistent group attitude: the majority view is that he is making satisfactory progress in coursework and teaching, that this would entail that he is making satisfactory progress in the

graduate program, but that nonetheless he is not making satisfactory progress in the graduate program. The majority scheme is in trouble because of cases like this.

The Discursive Dilemma generalizes this problem to show that no aggregation function will be able to satisfy a set of plausible constraints. Consider the following constraints on the aggregation function:

> **Universal Domain (UD):** The aggregation function should take all and only profiles composed of logically consistent and complete individual attitude sets.
>
> **Collective Rationality (CR)**: The aggregation function should yield as its output a group attitude set that is complete and consistent. Whenever $A \vDash B$ and $A_G(A) = \alpha$, then $A_G(B) = \alpha$.

These first two constraints fix the domain and range such that the aggregation function is consistency and completeness preserving. It would be unsurprising if there were no way to combine irrational group members into a rational group attitude. Instead, we are interested in a "rational in, rational out" aggregation function, and UD and CR restrict the domain and range accordingly. As the augmented profile Discursive Dilemma showed, majority voting can fail to combine complete and consistent individual judgment sets into a consistent group judgment set. So, this task is more easily said than done. Here are a few more constraints:

> **Member Invariance (MI):** Take some group **G** composed of members $g_1, g_2, \ldots, g_n$. Generate a group **G'** composed of the same members in a different order by swapping g_i and g_j but holding fixed the attitudes of g_i and g_j. F(profile **G**) = F(profile **G**'). This rule just means that the order of the members can't affect the aggregation output. Graphically, this means that permuting (any number of) rows on a profile table will not affect the last row on the augmented profile table.

Proposition Invariance (PI): Take some group G's profile P, with an agenda containing at least p and p*. Then take the A-subset of P and call this profile R.[33] If R's p-subset and p*-subset are equal,[34] then $A_G(p) = A_G(p^*)$. PI is a neutrality constraint among propositions: if two propositions have the same set of attitudes held toward them, then the group attitudes of those propositions should match.

The constraints listed so far are not sufficient to generate a contradiction. To do this, let us add two further restrictions:

Binary Attitudes: Every member of **A** is binary. I will discuss the effect of relaxing this after I give the first version of the impossibility theorem, but for now it will be easiest to focus on binary attitudes, as is common to the literature.

Minimal Agenda: The agenda must include two atomic propositions, their conjunction, and the negation of their conjunction. Call these propositions m, n, m&n, and ~(m&n), respectively. Minimal Agendas are closed under conjunction and negation, so whenever **p** and **q** are in the agenda, **p**&**q**, ~**p**, and ~**q** must also be in the agenda. This closure would also entail that **p**&~**p** is the agenda, but I will exclude tautologies and contradictions for simplicity without any meaningful losses.

Following List (2001), I will show that no aggregation function will satisfy these constraints for groups containing at least two members.

[33] I do this because P could contain multiple attitudes, in which case the p- and p*-subsets of R might match even though the p- and p*-subsets of P do not.

[34] Two sets A and B are equal just in case every element in A is in B and vice-versa. That is, A = B iff $(\forall a)(a \in A \rightarrow a \in B) \& (\forall b)(b \in B \rightarrow b \in A)$

Lemma 1: If any of a profile's minimal agendas contains a set of propositions whose subsets all match, and these propositions cannot all be true or all be false,[35] then F(.) will violate Collective Rationality if it satisfies Propositional Invariance.

Suppose there is some profile with a set of propositions ***P*** such that it can't be the case that every proposition in ***P*** is true nor can it be the case that every proposition in ***P*** is false.[36]

First, consider the simplest case, where there are two propositions p and q in the agenda and p and q are contradictory. Since the p-subset and q-subset match by assumption, PI entails that $A_G(p) = A_G(q)$. But this is a problem: since A is binary, either $A_G(p) = A_G(q) =$ yes or $A_G(p) = A_G(q) =$ no. If $A_G(p) = A_G(q) =$ yes, then F fails Collective Rationality for assenting to (believing, preferring, etc.) inconsistent propositions. If both receive "no", then F(.) will also violate CR. Since the agenda is minimal, it will contain ~p and ~q, and CR entails that $A_G(\sim p) = A_G(\sim q) =$ yes. But ~p and ~q are inconsistent, so F(.) violates CR anyway.

Next, consider a more complicated case, where no two members of ***P*** are contradictories. In this case there must be some $\boldsymbol{Q} \subset \boldsymbol{P}$ and proposition $r \in \boldsymbol{P} \cap \boldsymbol{Q}^c$ such that $\boldsymbol{Q} \vDash \sim r$ (where $\boldsymbol{Q}^c$ is the complement of ***Q***). Assuming the agenda is minimal, then the conjunction of all members of ***Q*** is in the agenda; call this proposition q. Clearly q and r are contradictories. PI entails that each member of ***Q*** and r must receive the same group attitude, either yes or no. And, $A_G(q)$ can take either yes or no. So, there are four cases:

	$A_G(q)$	$A_G(\boldsymbol{Q})$	$A_G(r)$
Case I	yes	yes	yes

[35] A set P is such that its members cannot all be true or all be false iff P is inconsistent and ~P = {~p1, ~p2,…} is inconsistent.

[36] This does not run afoul of universal domain as will become clear later. There can be cases where contradictories have equal subsets in a profile.

Case II	no	yes	yes
Case III	yes	no	no
Case IV	no	no	no

All four of these cases violate CR. Cases I and IV are situations already discussed, q and r are contradictories that are assigned the same attitude; both lead to violations of CR. Case II assigns ***Q*** yes but q no, which violates CR since ***Q*** ⊨ q. Case III assigns q yes but ***Q*** no, which violates CR since q ⊨ ***Q***. Since every case violates CR, F(.) violates CR. □

Theorem: There is no aggregation function F(.) that satisfies Universal Domain, Collective Rationality, Member Invariance, and Propositional Invariance.

Suppose for reductio that F(.) is an aggregation function that satisfies UD, CR, MI, PI, and that we are dealing with Binary Attitudes on a Minimal Agenda.

To show a contradiction, I will develop a strategy to build a profile satisfying Binary Attitudes and Minimal Agenda that forces Collective Rationality to be violated when the rest of the constraints are satisfied. Call this profile **Q**, and let us suppose that **Q** has at least two members. We can build a single attitude profile for **Q** using the following recipe:

Tough Crowd	m	n	m&n	~(m&n)	…
Member 1	yes	yes	yes	no	…
Member 2	yes	no	no	yes	…
Member 3	no	yes	no	yes	
Member 4, 6,…(n)	yes	yes	yes	no	…
Member 5, 7,…(n)	no	no	no	yes	…

The "Member 4, 6, …" corresponds to all of the even numbered members greater than 3 up to (and perhaps including n), and likewise the last row to the odd numbered members greater than 3.

Profile Tough Crowd cannot be aggregated according to all of the constraints. There are two cases:

Case 1: n is even. $A_G(m\&n) = A_G(\sim(m\&n))$, since m&n- and ~(m&n)-subsets match. Corollary 1 applies: m&n and ~(m&n) are contradictory, and their subsets of Q match.

Case 2: n is odd. $A_G(m) = A_G(m) = A_G(\sim(m\&n))$, since their subsets match. Since m, n ⊨ m&n, and m&n and ~(m&n) are contradictory, Corollary 1 applies.

Since Corollary 1 applies in either case, it follows that F(.) will violate CR if it satisfies PI. Since F(.) satisfies PI by assumption, F(.) violates CR. □

This result is largely in line with other related proofs in the literature. In the next paragraph I will compare a number of similar formulations. My different formulation has some strengths, however. First, I adopt a general framework that extends to many different types of attitudes, whereas most of the existing literature focuses on judgment or preference. The second advantage will become more evident in the subsequent discussion. It is widely agreed that these impossibility theorems illustrate that groups will be best understood (or formed) by identifying which of the inconsistent conditions is rejected. For instance, one could form a group that rejects propositional invariance and thereby is able to skirt problems highlighted in the above proof. Thus, different formulations of the constraints will dictate different ways to cut up the logical space of group types that are not susceptible to the impossibility theorems. I have formulated the constraints in such a way as to make the categories as sensible—to me, at least—as possible.

Impossibility theorems state something like the following:

- There is no aggregation function F(.) that satisfies Universal Domain, Collective Rationality, Member Invariance, and Propositional Invariance. (mine)
- "There exists no aggregation function satisfying universal domain, collective rationality, anonymity, and systematicity."[37] (List and Pettit 2002)
- When there are more than two alternatives, no social welfare function (which is an aggregation function for preferences) satisfies universal domain, collective rationality, a weak pareto condition[38], and independence[39], unless it is also a dictatorial function. (Arrow 1963)

In light of the common-sense view of groups, what is the right way to interpret these results? The first move might be to question the validity of the proofs. Our initial confidence in the general nature of groups may be stronger than our confidence in the proofs. But, these proofs are airtight. The second move is to reject one of the constraints. If some set of constraints is inconsistent, then at least one of them must be false. We can then investigate the consequences for rejecting each of the constraints, in hopes that rejecting at least one of them isn't so bad. I will do this later.

But first I want to consider an application of this impossibility theorem strategy to an ordinary epistemic practice: processing evidence to form judgments. I am interested in the following argument from analogy:

---

[37] Anonymity roughly tracks with Member Invariance, and systematicity roughly tracks with Propositional Invariance.

[38] Whenever every group member prefers alternative x to y, the group itself prefers x to y.

[39] If two different groups of the same size have matching preference subsets comparing x to y, then their aggregated preference towards x and y should match; if the first F(group one) prefers x to y, then so must F(group two), and vice-versa.

1. The impossibility theorems for attitude aggregation are relevantly similar to an impossibility theorem for evidence aggregation.
2. The evidence aggregation impossibility theorem poses no serious threat to agents' ability to rationally process evidence.

---

3. So, the impossibility theorems for attitude aggregation pose no serious threats to groups' rationally responding to the rational attitudes of their members.

I can then leverage the reasons for not being worried in the evidence case to find the weakness in the attitude aggregation case.

Here is the parallel case for evidence: in an agent's processing evidence, as in a group's aggregating its members' attitudes, various pieces of evidence must be combined to formulate opinions about a variety of propositions. These opinions are supported and undermined by different pieces of evidence, and agents face some difficulty in formulating rational attitudes while responding faithfully to their evidence. With this in mind, consider the following argument.

Suppose an agent S encounters a set of evidence $E = \{e_1, e_2,\ldots, e_n\}$. Further, consider some set of propositions $P = \{p_1, p_2,\ldots, p_m\}$ called the *agenda*. We say that the evidence is *binary* with respect to the agenda iff every e_i either favors or disfavors each p_i. The *level of support* for p from E is the pair <f, d> where f and d are the number of pieces of evidence that favor or disfavor p in E.[40] For the present case, let us suppose that each piece of evidence is equally strong. This is not true in general: there may be a single piece of evidence that is much

[40] For instance, if my three pieces of evidence all favored some proposition, then the level of support of my evidence for the proposition is <3, 0>. Note that f+d = |E|. So we only need to know any two of: the cardinality of the evidence, the number of "favors", the number of "disfavors", in order to determine the third value.

stronger than many other pieces of evidence combined. Later, in chapter four, I will discuss weighing evidence when individual pieces may be of varying strengths. The aggregation function F(.) is a function from E and P to a set of beliefs over P. These resulting beliefs are the aggregated beliefs, and these are the beliefs that we expect (and rationality demands) the agent to adopt based on her evidence. We can organize this set of evidence, propositions, and aggregation in an *evidence table* like so:

	p_1	p_2	…	p_m
e_1	favors	favors	…	favors
e_2	favors	disfavors	…	favors
…	…	…	…	…
e_n	disfavors	disfavors	…	favors
F(P, E)	disbelieve	believe		believe

In the above table, I randomly assigned "believe/disbelieve" in the appropriate cells, so do not read too much into their layout. The above table indicates that both e_1 and e_2 favor p_m (the rightmost cells in the second and third rows), and that S should believe p_m (the bottom right cell).

With these definitions at hand, I can proceed to offering constraints on the aggregation function. Graphically, these constraints are on how the final row of an augmented evidence table should be built out of the other rows. Here are the constraints:

Universal Domain (UD): The aggregation function should take as its input any possible evidence set.

Rationality Constraint (RC): The aggregation function should yield as its output a set of binary attitudes that are complete and consistent.

These two constraints combine to rule out rational dilemmas, situations in which an agent is confronted with evidence that cannot be rationally responded to. By adopting UD and RC, I am supposing that an agent can deal with any possible evidential situation rationally, at least in terms of forming beliefs.

Order Invariance (OI): The order in which the evidence is received is irrelevant. $F(\ldots,\{e_1, e_2\}) = F(\ldots,\{e_2, e_1\})$.

Support Constancy (SC): Any two propositions with the same level of support must have the same aggregated belief value.

OI and SC say that our aggregation function must be evenhanded with respect to evidence and propositions.

Minimal Agenda (MA): The agenda must include two atomic propositions, their conjunction, and the negation of their conjunction. Call these propositions m, n, m&n, and ~(m&n), respectively. Minimal Agendas are closed under conjunction and negation, so whenever **p** and **q** are in the agenda, **p**&**q**, ~**p**, and ~**q** must also be in the agenda.

The aggregation function comes cheaply if the agenda is impoverished. This MA requirement puts enough meat on the bones to show the difficulty in combining evidence.

The Favoring Relation (F)[41]: If evidence e favors p and e favors q, then e favors p&q.

F serves to make sure that our interpretation of a piece of evidence is consistent.

Lemma 2: If two contradictory propositions have the same level of support, then if F(.) satisfies SC, it violates RC.

[41] There's a lot to be said about the favoring relation, but I'm not sure how many commitments I need or want to make here. For instance, Elliott Sober (personal communication) reads "e favors p" as $Pr(p|e) > Pr(\sim p|e)$. I am fine with this interpretation, but I do not wish to accrue more baggage than necessary.

Suppose two contradictory propositions, p and q, have the same level of evidential support.[42] Then, by SC, they will have matching aggregated belief values. There are two cases:

i. $B_S(p)\&B_S(q)$: If S believes both p and q, then S violates RC by believing inconsistent propositions.

ii. $B_S(\sim p)\&B_S(\sim q)$: Similarly, if S disbelieves both p and q, then S violates RC by believing their negations, which are inconsistent.

Both cases lead to a violation of RC. □

Proof: No function F(.) can satisfy UD, RC, OI, MA, and F when there are at least two pieces of evidence.

The strategy here will be to build an evidence profile such that no matter how many pieces of evidence are added, Corollary 2 will apply to guarantee that RC is violated. Consider the following evidence table:

Tough Evidence	m	n	m&n	~(m&n)	…
e_1	favors	favors	favors	disfavors	…
e_2	favors	disfavors	disfavors	favors	…
e_3	disfavors	favors	disfavors	favors	

[42] One might object that this supposition is impossible. Here's the objection: Suppose p and q are contradictories with the same level of evidential support. Let's consider the simple case with two pieces of evidence: e_p for p and e_q for q. Finding evidence for a proposition means that one should boost their credence in the proposition. But, this would mean that upon learning e_p and e_q one should boost their credence in p in q. Since they are contradictories, their probabilities must sum to one, but this is clearly impossible on the assumptions just made. The problem with this objection is the assumption that "finding evidence for a proposition means that the proposition is more likely true than not". I think it is very plausible to say that, if e favors p, that $Pr(p|e)>Pr(p)$, but this doesn't mean that $Pr(p|e)>.5$. For example, my smoking would be evidence that I have (or will develop) lung cancer, but this doesn't mean that I am more likely than not of having (developing) lung cancer.

e_{2i} for $i \geq 2$	favors	favors	favors	disfavors	…
e_{2i+1} for $i \geq 2$	disfavors	disfavors	disfavors	favors	…

Let's consider the possible sizes, N, of evidence:

$N \geq 2$: (m&n) has the same level of support as ~(m&n). They are contradictory, so Corollary 2 applies.

$N > 2$ and is odd: m and n have the same level of support as ~(m&n). Since m, n ⊨ m&n, F(.) will mandate that S believes (m&n) if he believes m and n. But, this will violate RC since SC believes ~(m&n) also. So, F(.) must require that S disbelieve m and n. This also leads to a contradiction: since S will also disbelieve ~(m&n) by SC and ~m, ~n ⊨ ~(m&n).

In either case, F(.) violates RC. □

It might seem we are in a tricky situation here. If you reject the possibility of rational dilemmas, at least in cases of equal bodies of evidence for conflicting propositions, then you are faced with the judgment impossibility theorem just posed. The strongest way to frame this worry is that there is no objective way to rationally respond to evidence. If true, this would be disturbing indeed. However, I think we can safely evade any such serious worries; in showing how, I hope to inspire responses to the analogous aggregation problem for groups attitudes.

One tempting way to solve the problem is to claim that the sorts of situations depicted by Tough Evidence are relatively rare: we often have a balance of evidence in favor of one set of consistent propositions or another. Moreover, our two pieces of evidence rarely have the same evidential weight: testimony from an expert is worth more than testimony from a layperson, so cases of two opposing pieces of testimony often are decided by factors other than the sheer number of pieces. However, this objection does not sufficiently undermine the impossibility

theorem's bite. While it introduces the possibility that cardinality of evidence is insufficient to guarantee a "tie", it also introduces the possibility that bodies of evidence of different sizes might be ties as well. For instance, we might face cases where there are many pieces of testimony in favor of a proposition, but one stronger piece of testimony against the proposition. This threatens violations of RC in a broad range of cases.

Moreover, symmetrical cases of evidence seem to be pretty common, contra the objection. When deciding whether job B is better or job C is better, one might have only the evidence that the location favors B while the pay favors C. Or, when the undergrad asks whether to one box or two box in Newcomb's problem, the first two philosophy professors she asks might have conflicting views. It doesn't seem that such situations are rare. And, even if one thinks they are rare, a normative theory of rationality should tell us what to do in those cases. The rarity objection doesn't undermine this impossibility result

I think there are two relatively clean ways to take the sting out of the judgment impossibility theorem:

1. The range of F(.) is not well-motivated. The claim of RC is that any possible collection of binary evidence in UD will generate a set of binary attitudes that satisfy rationality constraints. But, there is a perfectly natural third option when faced with competing pieces of evidence: withhold judgment. Suppose you are a juror who only hears two pieces of testimony regarding the defendant's guilt. One claims the defendant is guilty, and the other claims the defendant is not guilty. Barring any reason to favor one piece of testimony over the other, and barring any further evidence regarding the defendant's guilt, it seems clear that one is remiss to judge the defendant one way or another.[43] This

[43] Two complications: First, let us assume that merely being the defendant does not offer any reason to presume guilt or innocence. Second, let us assume your judgment is based solely on what you have reason

de-fangs the impossibility result. The worry with the initial result was that there was no rational response, but substituting a tertiary attitude for a binary one makes the problem dissolve.[44]

2. SC is too strong. It demands that we respond to our evidence and propositions evenhandedly, which is certainly epistemically valuable. However, there are other epistemic values, including keeping one's judgments consistent and closed under entailment. So, in cases like Tough Evidence where N = 2 and there is equal evidential support for (m&n) and ~(m&n), consistency might motivate one to pick one of them over the other, in violation of SC.[45] This response is friendly to those who are pluralists about epistemic values, especially those whose pluralism motivates permissivism, where permissivism is the view that there are cases where more than one response is rationally permitted to the same body of evidence. In the context of the impossibility result, this position rejects the existence of a function, since it holds that there can be more than one rational output of F(.) for a given input. Alternatively, you could characterize this view as generating a broader range that includes sets of different permissible attitudes, in which

---

to believe, not based on any other norms of the justice system—such as presuming innocence except when beyond reasonable doubt.

[44] Dan Hausman suggests an objection: there are cases where we are forced to act based on our evidence, and this prevents us from opting out by suspending judgment. To make this objection more concrete, suppose you have to diffuse a bomb, and you are deciding between cutting the blue or the red wire. You have equal but conflicting evidence about which one to cut. I have just argued that you can withhold belief/judgment to avoid the rational dilemma, but you don't have the luxury of withholding judgment in the case of diffusing the bomb: you must act. My response to this objection is that saying one must act is different from saying one must believe. You should definitely cut one of the wires, but that doesn't require your belief that you are cutting the right one.

[45] One might worry that consistency doesn't tell you which proposition to prefer. It only forces you to reject at least one proposition from an inconsistent set, and thus is arbitrary. I will consider a response to this objection, or at least its analogue to the attitude aggregation case, in chapter 5.

case F(.) is still a function, but F(.) need not return a unique rational response to the available evidence.

We do not need to worry about our ability to generate rational judgments in response to bodies of evidence like those depicted in Tough Evidence. We are under no obligation to form judgments towards every proposition that we consider; we can resort to the ternary attitude of belief and withhold belief.[46] Or, we can reject SC in lieu of a more permissive approach that respects a variety of epistemic values.

Do these solutions translate to the problem of attitude aggregation? They do. Here are the analogues. Recall the proof was for the following claim: There is no aggregation function F(.) that satisfies Universal Domain, Collective Rationality, Member Invariance, and Propositional Invariance. We can undermine the severity of this claim as follows:

1. The range is too narrow. A group of individuals that have judgments over an agenda might yield a group that does not have judgments on every item in that agenda. We can broaden judgments to beliefs and allow the group a "withhold belief" value whenever there are ties among the judgments.[47]
2. MI and PI are too strong. While it seems virtuous for a group to treat its members and propositions evenhandedly, there are competing constraints on the group attitudes. As above, we can remedy this problem by either allowing multiple outputs or by giving a ranged output wherein the group attitude is identified as the range of acceptable outcomes.

[46] Taken further, you might "back up" to an even more general and fine-grained attitude. Credences might fit this bill, depending on your views of the relationship between credence and belief.

[47] Again, where groups are forced to act, they can act in spite of withholding judgment about what is true.

Thus, the impossibility theorems in Arrow's tradition do not undermine the possibility of rational group agents, even of the common sense sort. Rather, it shows that aggregating certain sets of attitudes will fail to yield complete judgment sets, or that these judgment sets may be influenced by epistemic values other than the level of support the propositions have from the group members.

In the next chapter, I will articulate my central proposal, which makes further use of the analogy between aggregating attitudes and agents aggregating their evidence: I claim that we can identify the attitudes of a group with the rationally permitted attitudes of a neutral agent whose only evidence is the testimony of the group members. Once that proposal is on the table, I will revisit whether it offers a satisfactory response to the challenges posed by impossibility theorems. In the remainder of this chapter, I will define and argue for the importance of aggregating *democratically* structured groups.

§2 …for Democratically Structured Groups

Some might find the whole problem puzzling because there are many groups that are relatively easy to aggregate. The attitude of the North Korean government is simple: those attitudes are identical to Kim Jong-un's. This aggregation function violates the non-dictatorial condition, but what else can you do for a dictatorship? Similarly, the attitudes of the United States Supreme Court are easy to decipher: the Justices cast votes, and the majority opinion wins.[48] These groups are structured in such a way that the aggregation function falls out of the structure for free. The DPRK is dictatorially structured, and the Supreme Court is, for lack of a better term, Supreme-Court-Structured. Their built-in decision-making procedures allow us to easily determine group

[48] Or in the case of ties, we can think of the court as withholding judgment.

attitudes with respect to some class of propositions. As the title of this section—and the dissertation—suggests, I am interested in the attitudes of *democratically* structured groups.

A group is democratically structured just in case every member has equal standing in the group, and hence contributes equally to the group attitudes. I was somewhat hesitant to pick 'democratic' because that word has a connotation of majority rule, which I do not want. I do not think that a majority-rule aggregation rule is adequate.[49] Rather, I emphasize the 'by the people' connotation of democracy. The paradigm instance for a democratically structured group is a collection of your good friends all deciding where to go out to dinner. You likely do not have an established voting system, or 'president of the friends' office. Instead, you will all share roughly equal responsibility in determining where the group ends up. Note that this does *not* mean that everyone plays exactly the same role. One person may be more knowledgeable about restaurants, and so that person may be more likely to recommend particular restaurants. And perhaps one friend has dietary restrictions, so they might veto any places that are not vegetarian friendly. Still, the restaurant you end up going to is selected by the group in a manner that respects each group member equally.

Groups of friends that are not democratically structured are painfully obvious. We have all had groups of friends like this where at least one friend was not respected equally. Perhaps the unequal friend was too dominant: her views were always more powerful in driving the group's actions than they ought to be. Or, perhaps the unequal person was less respected than the others when it came to shaping group attitudes. This isn't always a bad thing: there are groups of people where having unequal power structures is appropriate. A family with young children would be

[49] I toyed with 'egalitarian group structure' and 'unstructured group attitudes'. The first narrowly lost out, but just didn't roll off the tongue as well. The second seemed to indicate that those entities weren't genuine groups.

terribly structured if the young children had equal power to the adults of the family, for instance.[50] For another example, combat groups in the military must act decisively, and as such there may need to be non-democratic structures that enable such action. In both cases, the contrasting unequal groups should make clear what I mean by democratically structured groups.

You might wonder how common democratically structured groups are. If they are rare, or relatively uninteresting, then my solving the attitude aggregation problem for democratically structured groups is less important. Democratically structured groups are pervasive and more interesting than non-democratically structured ones. Even groups that have explicit non-democratic structures usually are only so structured for a specific type of propositions. Again, consider the case of the United States Supreme Court. This group has explicit structure for determining judgments on court cases, but presumably the court has attitudes about a number of other sorts of propositions. For instance, there are truths about what sorts of cases the court prefers to hear and whether the court believes global warming is occurring. While it is certainly important to know what the Court has decided in some case, it is fairly easy to do so. The interesting aggregations that are highlighted by the impossibility theorems are ones where the group structure doesn't hand the aggregation function to us.

Even non-democratically structured groups can be fruitfully thought of as having democratically structured attitudes. In fact, the democratically structured group attitudes of a group might be morally prior to the group's non-democratically structured attitudes. One criticism of an evil dictator is that the dictator acts in ways that conflict with the interests of their nation. But, if the nation's attitudes just were identical to the attitudes of the nation's dictator,

[50] This doesn't mean that I think children are less important or that their well-being is less important than the well-being of their parents.

this would be absurd. What is meant, I think, is that the dictator's attitudes diverge from the democratically structured attitudes of the nation. You might think of this counterfactually: the dictator is morally bad because they are acting in ways that conflict with how the nation would have acted if it were structured democratically.

Let me preempt a possible misconception of democratically structured groups: democratically structured groups are not groups wherein everyone plays the same role. It is perfectly compatible with a democratic structure that certain group members are experts in some domains, and thus those members have more sway when it comes to group decisions in those domains. Let's suppose that my wife and I form a two-person group. Since my wife is a dental hygienist, I defer to her opinions when it comes to matters about dental care. This does not undermine our group being democratically structured: I freely consent to this deferral. When looking at large sets of propositions, it might appear that a democratically structured group has subsets of the agenda that are aggregated dictatorially, but this is not problematic if the members of the group would willingly defer. This willing deference does not undermine each person's equal participation in the group.

My refining of the problem of attitude aggregation is unique. I have not seen others do this in the literature. However, this move is not a huge departure from the spirit of other views in the literature. We can see this in the sorts of constraints that are suggested as restrictions on the aggregation function. The non-dictatorial condition is only plausible if you want to rule out certain group structures. Similarly, member invariance and propositional invariance also indicate a disposition for groups and agendas being 'evenhanded'. The natural way to motivate these constraints is to distinguish the sort of group we are after from other existing group structures. Arrow (1951) says the following about the non-dictatorial condition: "Since we are interested in

the construction of collective methods of social choice, we wish to exclude dictatorial social welfare collections". The impossibility theorems only get off the ground because of the plausibility of their constraints, and those constraints are motivated by something like a desire to model democratic structures. I will return to the constraints in chapter four. But first, in the next chapter, I will present my solution to the attitude aggregation problem for democratically structured groups.

Chapter 3: The Steward Approach

In this chapter I will argue for my solution to the problem of attitude aggregation: the attitude of a group is identical to what a neutral agent ought to believe when that neutral agent's only evidence is the testimony of that group. I know, that's a mouthful. Put another way, the attitude of the group is equal to the evidential value of the group's testimony. Imagine your only evidence was the testimony of the group: what would it be rational for you to believe? I call this approach the Steward Approach, since I name the neutral individual with no prior evidence "the Steward". In this chapter, I will first clarify the Steward Approach to attitude aggregation by making the thesis more precise and providing a few examples. Then, I will offer three arguments and a promissory note in favor of my solution capturing a group's attitude.

First, a caveat. The Steward Approach, which I formulate below, is only intended to be applied to groups. Strictly speaking, you could apply the Steward Approach to any collection of individuals. However, as discussed in chapter one, not every collection of individual agents forms a group. You, the current President of the United States, and Danny DeVito do not likely meet the minimum requirements of forming a group. It is still possible to apply some principles that I develop in later chapters to that odd collection of three[51] agents, but the Steward Approach is there to tell us about the nature of group attitudes. If there is no group to begin with, then there are no group attitudes to look for.

Recall that the attitude aggregation problem for groups was to determine a function that took the individual group members' attitudes as inputs and yielded a group attitude as an output.

[51] Two, if you happen to be the President or Danny DeVito. One if you're Danny DeVito and were elected President.

We also want this function to satisfy certain constraints, and the impossibility results covered in chapter two showed that several plausible combinations of these constraints cannot be jointly satisfied. I will wait until chapter four to argue for reasonable constraints on the function. For now, I will just attempt to show that the steward approach generates a function that takes the proper inputs to be a candidate solution to the attitude aggregation problem.

The steward approach is as follows. Take some group G, agenda P, and set of attitude types A. Then, suppose there is some neutral agent s, who I call the Steward, who is rational and has no evidence whatsoever with respect to the propositions in P. Suppose s receives G's profile—that is, the set of all of G's members' attitudes towards all of the propositions in the agenda—as evidence. Then, going attitude by attitude and proposition by proposition, we ask "what does this evidence rationally compel s to do?" For instance, should the steward believe, disbelieve, or withhold belief about p? How about q? And so on for each proposition in the agenda. Then we run those questions for the other attitude types in A. For each individual question about attitude type a* to proposition p, there are three cases to consider:

> ***Case 1: s has no rational response of type a* to p***. This is not a possibility that I take very seriously. I deny the possibility of rational dilemmas. However, if case 1 ever holds, then G has no attitude of type a* to p.
>
> ***Case 2: s has exactly one rational response of type a* to p***. This case is entailed by the Uniqueness Thesis.[52] While there are many subtle flavors of uniqueness, the general thrust is that an agent has a single permitted rational stance towards any proposition given

[52] See Feldman (2007). "This is the idea that a body of evidence jsutifies at most one proposition out of a competing set of propositions...and that it justifies at most one attitude toward any particular proposition" (p. 205).

some body of evidence. In this case, the group attitude will also match the steward's attitude of type a* towards p.

Case 3: s has more than one rational response of type a* towards p. If one denies uniqueness, one allows for situations where more than one response is rationally permitted. Perhaps more than one credence is deemed permissible,[53] or perhaps the steward could permissibly either believe or withhold belief about a given proposition. In this case, the group's attitude is either a broad attitude that encompasses all of the permissible attitudes for the steward, or the group attitude is neutral between all of the permissible steward attitudes.

The Steward Approach identifies the group attitudes as the rationally permissible attitudes for s. I deny the existence of rational dilemmas, and so I'm not worried about case 1. But for completeness, I say that the group also has no attitude towards some proposition if s has no rationally permitted attitude to take towards that proposition. If there is exactly one rationally permitted attitude for s, then I identify that as the group attitude. If there is more than one rationally permitted attitude for s, then I say that either the group is neutral between that set of attitudes or that the group takes a broad attitude, a single attitude that encompasses all of the permissible ones, towards the proposition. I can now clarify my steward approach for attitude aggregation:

Steward Approach: Suppose an agent's, the steward's, evidence consists of all and only the attitudes that members of G take towards the propositions in P. Then, for any proposition **p** and attitude type **a**, the attitude of the group, $\mathbf{a}_G(\mathbf{p})$, is determined by the set of rationally permissible attitudes of s, $\mathbf{A}_s = \{\mathbf{a1}_s(\mathbf{p}),\ldots\}$, as

[53] See Castro and Hart (2017).

follows: i) if $\mathbf{A}_s$ is empty, then $\mathbf{a}_G(\mathbf{p})$ does not exist, ii) if $\mathbf{A}_s$ has exactly one element, then $\mathbf{a}_G(\mathbf{p})$ is identical to that element, iii) if $\mathbf{A}_s$ has more than one element, then $\mathbf{a}_G(\mathbf{p})$ is either neutral between those elements or takes a broad attitude that encompasses all and only those members.

In order to make the Steward Approach more concrete, I will provide a few examples to show how the theory works out.

Consider a small group of two persons, Abe and Ben.[54] The proposition in question is "LeBron is the greatest basketball player of all time." The attitude type in question is belief. Abe believes the proposition in question is true, and Ben believes it is false. What should the steward believe in this case? More to the point in this chapter, how might the Steward Approach interpret this rational obligation in assigning a group attitude to Abe and Ben? Rather than argue for what the steward is obligated to do here, I will just consider three possibilities. After I have articulated and defended the Steward Approach in this chapter, I will spend next chapter arguing for the rational constraints on the steward when encountering bodies of evidence.

The steward's obligation	What the Steward Approach says about the group
s has no rational response	The group has no attitude toward p.
s should withhold belief about p	The group withholds belief about p.
s is permitted to believe, disbelieve, or withhold belief about p	The group attitude is neutral between belief, disbelief, and withholding belief in p.

The above table gives the three most plausible views I can think of about the Steward's rational obligations in combining Abe's belief and Ben's disbelief. I am partial to the second response—that s should withhold belief—but that is less important here. It is more important to note how

[54] If it helps, I imagine Abe and Ben as a talk radio sports duo.

the Steward Approach would view the group attitude conditional on any of these three responses to the evidence being the correct ones.

Consider another example. The proposition in question is "It will rain tomorrow" and the three-person group consists of Carly, Dan, and Emily. Their credences are captured in the following profile table:

Name	Credence in "It will rain tomorrow."
Carly	.6
Dan	.8
Emily	.4

Suppose s is confronted with only the facts in the above table, the three different credences that Carly, Dan, and Emily assign to the proposition that it will rain tomorrow. What should s's credence be? Again, we can split the answers into three categories:

Type of Response	s's required credence	What the Steward Approach says about the group
No rational response	None	The group has no credence toward the proposition.
One Response	(.6+.8+.4)/3 = .6	The group credence is .6
	.8 (see footnote[55])	The group credence is .8.
	…	The group credence is …

[55] Obtained by (.4*.6*.8)/(.4*.6*.8+.6*.4*.2).

More than One Response	Anything in [.4, .8]	The group credence is neutral between all of [.4, .8]
	Any range centered on .6; e.g. (.55, .65)	The group credence is neutral between all ranges centered on .6
	…	The group credence is neutral between …

Here it is less clear what view one might adopt as to what s's credence should be, and as such there are more options to consider than in the Abe and Ben case. Again, I find it implausible that s will have no rational response. I will argue in the next chapter for the formula that justifies .8. But for now I just want to be clear on what the Steward Approach says about the group attitude conditional on a particular theory of the right response to the evidence. In each case, the group attitude matches s's prescribed attitude.

Hopefully the Steward Approach is clear at this point. Find what the steward ought believe/crede/prefer/etc. and then that gives us the group attitude. There are three main tasks to accomplish in the coming chapters. First, I need to argue for why the Steward's credences should tell us what the group attitudes are. I do this in the remainder of the chapter. Second, I need to detail the Steward's rational constraints: for instance, which row in the previous tables do I think actually captures s's rational obligations? This is the task in chapter four. Third, and related to the first point, I need to explore whether the results of the Steward Approach are plausible. I do this in chapter 5 by applying the work of this chapter and the next to particular cases. The results given to these particular cases by using the Steward Approach are plausible both descriptively and normatively: my approach coheres with a common-sense view of what groups are like.

As just mentioned, the rest of this chapter will be dedicated to advancing arguments in favor of the Steward Approach. There are four.

Argument 1: The benevolent Dictator

A common theme in fairy tales, at least the versions that appear in shows my daughter watches, is that good Kings and Queens make decisions that take into account all of their subjects' views. This is seen whenever the 'good monarch' character makes a tough decision; their people will all admit, some begrudgingly, that the choice was the right one. I want to leverage this sort of benevolent dictator to argue for the Steward Approach:

1. A benevolent dictator's attitudes should match the group attitude.
2. A benevolent dictator should form their attitudes by combining the group members' testimony.

---

3. Hence, the steward approach is identical to the group attitude.[56]

Often I find myself doing something that chiefly affects another person. For instance, I might pack a lunch for my wife or a friend. When I do this, I have to largely remove my personal beliefs and tastes and replace them with someone else's. I don't care for tomatoes on my salad, but my wife does, and so I throw some tomatoes into the mix. This process of deferring to another's attitudes gives a model for determining what those attitudes are. Of course, I won't entirely eschew my own beliefs: I might make her a recipe that only I know. Or I might prepare

[56] You might note that this argument is not deductively valid. I've included the double underline to suggest that it's an inductive argument. In various iterations I have considered deductively valid forms of the argument, but I think the above construction gets at the heart of the matter. The dictator should set his attitudes to the group attitudes, but he should also form his attitudes based on the testimony he receives. This gives us good reason to think these approaches should yield the same answer, which is the what the Steward Approach claims.

her a healthier meal than she otherwise would have, as healthy eating is more easily done for someone other than yourself.[57] But, by and large, we are used to the act of internalizing another person's attitudes in order to act on their behalf (you might say as their...steward).

I want to generalize this approach to groups of people. Suppose you are called upon to represent the interests of some group. You have been appointed by a charity to spend a large sum of money to benefit a neighborhood called Dunn's Marsh. There are five potential plans for improvement: A, B, C, D, and E. You do not live in Dunn's Marsh, nor do you plan on spending any time there once the project is complete. The residents of the neighborhood have all honestly revealed their preferences on the ordering of the five plans. Let us suppose that you have no evidence, other than the resident's preferences, that any of the plans are better or worse than the others. How should you go about determining which plan to follow? While I think the particular procedure for making this decision may be complicated (see Arrow (1956) for a reminder that you can't just use majority voting to generate the Dunn's Marsh preference), I think it is clear that your choice should depend only on the preferences of the Dunn's Marsh residents. For instance, if all you knew was that everyone who lived there preferred plan C over all of the others, then you should clearly implement plan C.[58]

The argument listed at the start of this subsection is in terms of a "benevolent dictator". Before I defend my two premises, I should clarify what I mean by that term. Consider the following story:

---

[57] Anyone who knows me and my wife will realize the humor in this: she is a far healthier eater than myself. But suspend belief for the sake of the example.

[58] It is tempting to create counter-examples, but I cannot think of any here. In every case where it seems to me that you should not implement plan C despite everyone in the neighborhood wanting you to, you will need to have access to extra information about why C would be a bad idea.

> **The Benevolent Dictator**: Ben finds himself in an odd situation. He is the dictator of Aggreba, your traditional Disney-style Arabian kingdom. Unfortunately, he suffered a head injury that wiped his memory of all of his evidence. He is still able to use language perfectly, and he has perfect insight into all logical truths. There are two further fairy tale twists. One, Ben is struck by beneficence; whereas he used to only seek to satisfy his own desires, he now genuinely cares about everyone else. Two, Ben is not allowed to leave the palace to obtain any information about Aggreba for himself, but he can summon whoever he likes and receive their honest opinions about all of the relevant goings on.

Ben is our Benevolent Dictator. I want to argue for two claims about Ben that correspond to the claims in the argument made at the start of this section.

First, Ben should respect the group attitude. Ben has no evidence of his own, due to his accident. So, in order to determine what is best for Aggreba, he should defer to the group view. Suppose that he only knew for a fact that Aggrebians preferred a six-day work week over a five-day work week. In acting as his capacity as Dictator, Ben should treat the six-day work week as preferable over the five-day work week. Similarly, if it was clear that the group believed any proposition P, then Ben should, too.[59]

Second, Ben should form his beliefs by acquiring the testimony of the people of Aggreba. This is the uncontroversial thesis. As Ben calls people in to the palace and asks them questions, he should update his beliefs and credences—and any other attitudes—accordingly. Taken to the

[59] If the Aggrebians have inconsistent beliefs, then Ben may need to choose between conflicting beliefs or withhold belief. I will consider these sorts of cases in the next chapter.

limit, if Ben wants to form the most accurate attitudes that he can, he should acquire the total testimony of all of his subjects. Of course, there would be practical constraints on how much time he should take, but it's my story, so we can posit whatever magic is necessary for Ben to get all of the testimony.

These two claims give us reason to think that the Steward Approach is on track. Take any group of people that you like. Then, suppose that we assigned them a Benevolent Dictator (or Steward) with no antecedent evidence over some agenda. That person should simultaneously adopt the group attitude and form his attitudes by updating on the total testimony of the group. But this tells us that these two procedures should have the same result. The attitude of the group should just be that attitude that our steward arrives at upon processing all of the testimony.

One area where this line of thinking is prevalent is in union representation. The heads of the National Football League Players Association (NFLPA) and the National Basketball Players Association (NBPA), DeMaurice Smith and Michele Roberts are tasked with representing the interests of the players from their respective sports. These leadership positions are often subject to public criticism if they act in ways that favor, say, the best players making the richest contracts over the journeyman player making the veteran’s minimum salary. To do their jobs well, Smith and Roberts should not be satisfying their own personal desires for what the NFL and NBA should look like. Rather, they are charged with representing the democratically structured groups of players’ desires, wishes, etc.

Argument 2: Deferring to groups

Cariani (2011) notes two broad uses for the formal tools of judgment aggregation. The first is to understand the nature of group structures and decision making. The second is in order for us to defer to collections of individuals, such as a collection of experts. Notice that this second reason doesn't depend on the collection of individuals being anything above and beyond individuals: eliminativists about groups may still want to defer to experts in some particular field. I want to develop this second reason of Cariani's as a motivation for adopting the Steward Approach. The argument:

1. Deferring to each member of a collection of individuals should yield the same answer as deferring to that group of individuals.
2. The Steward Approach applied to a group yields the same result as deference over individuals.

---

3. So, the Steward Approach gets the right answer in group attitude aggregation.

This argument doesn't require the same setup as the benevolent dictator argument. We defer to groups all the time:

- When drafting his fantasy football team, Tyler defers to the rankings on ESPN's Top 300 list.
- After hearing competing views on a talking heads "news" show, Bill has no idea whether climate change is real and/or man-made. Bill says, "I'll just go with whatever the experts think."
- When designing my first Symbolic Logic course and picking a textbook, I went with the received view by consulting a variety of syllabi from logic courses at reputable universities.

Given the plausibility of these scenarios, hopefully you agree that deference to experts is not only a thing that we do in practice but something that we should do. I can now defend the two premises of my argument.

Premise one states that deferring to each member of the collection of individuals is the same as deferring to the group composed of those individuals. Let's consider both directions of this identity. First, if you defer to a bunch of individual experts, then you should be already committed to agreeing with that group of experts. Second, the converse: if you defer to a group of experts, you should be committed to updating on the views of each individual expert in that group. Consider the following scenario:

> Climate Change Panel: There are five major experts on climate change. These experts come together as part of a global initiative on climate change and publish their group findings. None of the experts change their minds in forming this group view, instead they merely combine the previously held individual views to form a single 'group view' on climate change.

Let us consider the two conditionals. First, suppose you had read and showed appropriate deference to each of the five individuals on the panel, but that you did not agree with the group's published group findings. This signals that something is amiss. Either you do not hold them each to be equally experts, or you do not think that they properly aggregated their individual expert opinions into a group view.[60] Second, suppose you first read and defer to the group findings, but then go on to read each expert's individual opinions on the matters at hand. It would be very strange if subsequently learning of the individual positions shifted your view away from the

[60] Take the fantasy football example. Suppose I read five equally good expert opinions on who to start at QB this week. I then look at the rankings that are supposed to combine these five expert opinions. If my view upon deferring individually doesn't match the aggregated view, then there are two plausible responses. First, I could say that I don't think they are all really experts: maybe my views are much more heavily influenced by Eric Karabell than Matthew Berry. Second, I could say that the algorithm for combining the views is incorrect: they are all experts, but when you combine their views you should end up with something other than what the site displayed as their "aggregated view".

agreed upon group view. Again, if you change your position upon reading the initial views, then you are either committed to saying that not every person in the group is equally an expert or that the purported "group findings" are not actually indicative of the group's attitudes.

Premise two states that the Steward Approach yields the same results as deferring to a collection of individuals. This premise is analytic. The Steward Approach is just to pull together the large set of testimony from all of the group members into a (perhaps singleton) set of rational responses. Thus, the Steward approach is a form of deference, though not necessarily to experts, at least it is deference to all of the members of some group.

Our commonplace practice of deferring to experts then gives us strong reason to accept the Steward Approach. Experts often form groups, and we often want to defer to the established expert group opinion. However, this act of deference shouldn't be any different from just deferring to each individual expert that comprises the group. But this points us to the Steward Approach: the attitude of the group is just the attitude that one should arrive at upon aggregating each of the group members' attitudes.

Argument 3: Pressing the analogy with evidence

The third argument in favor of the Steward Approach is to continue the analogy drawn in the previous chapter between aggregating individual attitudes and aggregating evidence. The argument from analogy is as follows:

1. Aggregating individual attitudes is relevantly similar to aggregating evidence.
2. The Steward Approach is appropriate for aggregating evidence.

---

3. So, the Steward Approach is appropriate for aggregating individual attitudes.

The second premise is uncontroversial. The Steward Approach is explicitly tying together the aggregating of evidence and the aggregating of individual attitudes into group ones. The real crux of this argument is whether premise one is true: are these two types of aggregating relevantly similar? In chapter 4, I will talk about the specific rational constraints and procedures for aggregating testimony. When you think about these constraints, you can keep in mind whether the analogy between evidence and group attitudes is apt. For instance, I will claim that when just considering a single proposition and whether to judge that it is true, one should adopt a "majority rules" procedure. As I argue for this claim, you can consider whether the question of processing the evidence lines up with your intuition about characterizing a group with a similar sort of profile. In what follows here in this chapter, I will not get into the finer details of the rationality constraints. Instead, I will focus on more general similarities between aggregating evidence and aggregating attitudes.

Combining evidence and individual attitudes should be order invariant. In chapter two, you might recall that several of the impossibility theorems used a constraint about order not mattering. This is why groups of individuals can be thought of as sets; two sets are equal when they contain the same elements in any order. Something similar holds for processing evidence. Suppose I draw balls from an urn with replacement in order to form beliefs and credences about the contents of the urn. Drawing five white and five black balls from the urn should have the same evidential value regardless of what order I draw them in. There are some instances when order matters: perhaps the colors of the balls changed with my behavior, or maybe the contents of the urn change over time. But, if we hold fixed the contents of the urn, then drawing alternating colors (WBWBWBWBWB) or separated colors (WWWWWBBBBBB) should result in the same credences with respect to the overall distribution of colors in the urn.

Both combining evidence and individual attitudes force the agent doing the combining to satisfy rationality constraints. While I will argue that these constraints are in fact the same in the next chapter, it's worth pointing out here that there are constraints at all. For instance, suppose someone told you that they had successfully aggregated a group's members to yield the result that the group both believed and disbelieved that tomatoes are fruits. This can't be right. It's controversial whether tomatoes are fruits,[61] but the group attitudes should be consistent on this matter. Similarly, I have made a mistake if, upon receiving a good deal of evidence about tomatoes, I arrive at the conclusion that they both are and aren't fruits. Tricky though the question may be, I need to come to a consistent position on the matter.

Combining evidence and individual attitudes are difficult. Of course, many things are difficult, but the difficulties between these two things seem to line up. For instance, both threaten to put us in genuine dilemmas. As the impossibility theorems in chapter two showed, we can generate mirrored impossibility theorems. To put a less formal example on the paradoxes, consider the following two cases:

> **Aggregation Problems:** You are trying to determine the group attitude for a collection of libertarians. They identify as a group, have group aims to further libertarianism, and are free and honest with you about their opinions. But that's where the trouble starts: their opinions often strongly conflict. Some want to abolish all taxes. Others think that some minimum tax is necessary. To make things worse, these people have other commitments that might entail the falsity of

[61] If you find that to be uncontroversial, consider whether hot dogs are sandwhiches, or which of the four seasons is best.

their other opinions. You are stuck trying to combine a collection of conflicting opinions into a roughly coherent "libertarian view".

Evidence Problems: You are a doctor with an ill patient. You run several tests to diagnose the illness, but they put you in a bind. You ran the same blood sample twice, but the test results conflicted (one was positive for condition X while the other was negative for X). The patient likely has one of two illnesses, but her symptoms do not clearly pick out which one. Some of her symptoms match well, but others are rare for the respective conditions. If you knew what the condition was, then you could easily build a post hoc justification for that diagnosis, but without that knowledge, you're pulled equally in two conflicting directions.

These cases are quite similar. In both situations, there's a threat of paradox: you might end up with problematic collections of individuals or selections of evidence that make it difficult for you to meet all of your commitments. Given these similarities, there's reason to think that epistemological solutions to aggregating evidence will also provide solutions for identifying group attitudes. This favors my Steward Approach.

Argument 4: A Promissory Note

The last argument will be developed in the next two chapters. It is inductive argument that is strengthened whenever the Steward Approach yields an intuitive answer to a particular aggregation problem. If the Steward Approach yields a counter-intuitive aggregated attitude for some particular case, that counts against my view. The previous three arguments provided motivations to think that treating attitude aggregation like aggregating testimony from an

evidence-free perspective would do the trick. However, what really matters in solving the attitude aggregation problem is to find a tool that will do the formal job, where that job is specified by arguing for a set of conditions that the aggregation function must meet. Sample conditions were seen in the previous chapter: systematicity, proposition invariance, minimal agenda, etc. In the chapters to come, I will offer my own arguments for the conditions that the aggregation function must meet. Once those lines are drawn, I will be in a position to show that the Steward Approach meets those conditions. For now, I will leave you with this as a promissory note to be revisited later.

There are a few more general benefits of the Steward Approach worth discussing now. First, the Steward Approach places the aggregation problem in a positive light. Whereas the literature has often focused on impossibility theorems, reasons that aggregating will be difficult and/or impossible, the Steward Approach should make us as optimistic about finding a group attitude as we are about being able to process evidence. Of course there are cases when it is difficult to know how to deal with a certain body of evidence[62], but it's fair to say there's a general optimism that there is at least one rational response to the vast majority of epistemic situations that we find ourselves in. Similarly, I'm optimistic that there's an adequate characterization of a group's attitudes in the vast majority of cases.

The second virtue of my approach is that it gives us a better understanding of what the failures to aggregate represent. Strictly speaking, the Steward Approach will generate a set of rational attitudes for any profile, whether or not the agents listed in that profile form a group. However, some of these representations will be extremely broad when the group attitudes are not

[62] See Sleeping Beauty Problems, or cases where different epistemic values conflict, or when you are dealing with a complicated proposition that turns out to be a tautology.

complimentary. For example, suppose there is a group composed of two people and those two people disagree on every proposition in the agenda. When person A believes P, person B disbelieves it. When A is certain that Q is true, B is certain that it is false. And so on. In this case the Steward will not have many rational constraints imposed by the testimonies. In the face of two conflicting pieces of testimony about, say, P, there seem to be two leading candidates for what the Steward is required to do:

> Option 1: The Steward can do whatever she likes; believe, disbelieve, or withhold belief about P.
>
> Option 2: The Steward has no preponderance of evidence for or against P, so the Steward should withhold belief with respect to P.

So, expand this problem to a large group with a large agenda. Suppose that the Steward either has a great deal of freedom (a la Option 1) or that the Steward is stuck holding neutral attitudes toward everything in the agenda (a la Option 2). In this case, what shall we say about the group? I say that either this isn't a group at all, or at least it's a very weak one. As we saw in chapter one, groups are characterized by their having shared goals and the capacity to act. But if it turns out that a group widely disagrees on the agenda, then it's hard to see how they could have any shared goals at all. Moreover, how are they supposed to act if they do not have a reasonably large basis of agreement on which to found action? This is how the Steward Approach allows us to better understand problem cases for aggregation: the problem cases are not problems for attitude aggregation, but rather they are cases that are near the boundaries of counting as groups in the first place.

Hopefully these four arguments have convinced you that the Steward Approach is at least a promising place to look for the solution to the aggregation problem. The beneficent dictator would both adopt the group's attitude as well as form his by updating on their testimony. We defer to groups of experts, but doing so is also accomplished by deferring to the individual experts themselves. Aggregating individual attitudes into a group attitude is similar to aggregating evidence in our ordinary epistemic lives. And, as I will show in the next two chapters, the Steward Approach gets the right results. In particular, the Steward Approach generates a positive approach to the aggregation problem that also helpfully characterizes edge cases, cases where groups are composed of such disparate members that they are barely groups at all.

Chapter 4: Aggregating Evidence and Aggregating Testimony

In this chapter, I will present necessary conditions for a theory of aggregating testimony as that relates to the Steward Approach. Strictly speaking, the fully solved Steward Approach would require having a theory for aggregating any possible body of evidence. I will not be able to accomplish that enormous task in this chapter. Instead, I will flesh out some salient features of evidence aggregation in order to demonstrate that the Steward Approach solves the attitude aggregation problem. The chapter proceeds in two parts. First, I will run through a series of constraints on aggregating both evidence and group attitudes. For each constraint I will evaluate whether it is indeed a necessary condition on rational aggregation. Second, I will build up a theory of aggregation by considering cases of increasing complexity. After developing these tools in this chapter, I will be in position to show that the Steward Approach satisfactorily solves the aggregation problem in Chapter 5.

But before I get to the rules themselves, I want to more carefully lay out what I mean by "evidence". On my view, one's evidence is the (collection of) proposition(s) that one can use to justify doxastic attitudes. This is not the dominant view. Indeed, it is hard to determine what the dominant view is. However, all views on evidence that I am aware of hold that evidence is used to justify attitudes like belief. And, the view of propositions as evidence can be translated into other views. For instance, a Russellian view is that evidence is sense data. These sense data mental objects will directly correspond to propositions, and thus is translatable from my view. Or, on a knowledge-first Williamson view, again there will be an easy one-to-one correspondence between evidence and proposition.

Given that I treat evidence as a set of propositions, one might be puzzled why aggregation of evidence is an issue at all. Aggregating propositions into a single proposition is a simple affair: just conjoin the propositions. Indeed, this is a *feature* of the proposition view of evidence. When I see a red jacket on my daughter, what evidence do I receive?[63] Is it that she's wearing red? That she's wearing a jacket? Do I have to break the evidence down into 'atomic' pieces of evidence? The answer is no. To try to count evidence in this way is a fool's errand in just the same way that counting propositions is a fool's errand. Once you have a collection of propositions, individuated however you like, you can conjoin them into a single piece of evidence.

When I talk of the problem of aggregating evidence, I do not mean the problem of turning multiple pieces of evidence into one. That is trivial. Instead, the problem of evidence aggregation is determining the rational response to a body of evidence: combining them into one's rational doxastic attitudes. This is not trivial, as evidence can pull in different directions. If two pieces of evidence favor a proposition and another piece of evidence disfavors that proposition, more information is needed to determine whether the conjunction of these three pieces of evidence favors the proposition. In this chapter, and the dissertation more broadly, this is the sort of aggregation that I will be comparing to aggregating attitudes of group members.

§4.1 Some Accepted and Rejected Rules for Aggregation

There are several constraints that one might apply to the aggregation function, either for aggregating individual attitudes into group attitudes or for aggregating evidence into a single

[63] It's not critical here, but evidence is usually thought of as evidence *for* something, let's suppose I'm wondering whether Elizabeth will be warm enough on her walk this afternoon.

rational doxastic perspective. Some of these constraints are *limiting* in that they rule out certain approaches to aggregating, while others are *constructive* in that they tell us what the aggregated view should be in certain cases.[64] Later in this chapter, in §4.2, I will build up some general constructive principles by considering simple cases and moving on to more complicated ones. First, though, I want to go through a variety of plausible sounding aggregation rules that merit further discussion.

Each rule will have both an evidential and an attitude aggregation version. I'll signify this by adding a '-A' to the attitude aggregation and '-E' to the evidential aggregation version. I want to show that the -A version supervenes on the -E version. To do this, I will follow a standard format for each sort of rule. First, I will introduce the type of rule. Then, I will formulate and offer a prima facie defense of the rule as a constraint on attitude aggregation. Next, I will formulate and offer a prima facie defense of the rule as a constraint on evidence aggregation, particularly with regards to testimony. I then evaluate whether the evidential rule has any merit, usually drawing attention to counter-examples. Once I have the refined verdict of the rule for evidence aggregation, I will draw the parallel lesson for attitude aggregation.

There are two main upshots of this section. First, I am offering an inductive argument for the relationship between attitude and evidence aggregation. I will show that they are closely connected in this collection of rules, and that should give us reason to think that they are more generally connected as per the Steward Approach. Second, this discussion will frame the constructive approach that I advocate for in the second part of this chapter.

---

[64] Really, constructive constraints are also limiting. If you should believe such-and-such, then that rules out any other requirements on which you aren't required to believe such-and-such. Still, there's an important difference in flavor between constructive and constraining rules. The constructive constraints ensure that there's at least one solution, whereas limiting constraints need not offer any such hope.

Rule 1: Unanimity

In this section I will argue against rules of unanimity—roughly, rules that say that aggregating a unanimously agreed group or body of evidence should just set the aggregate view to match. While this rule might look attractive, and indeed will give the right answer in many cases, it runs into some serious counter-examples.

> **Unanimity-A:** for any attitude type and proposition, if every member of the group has the same attitude toward that proposition, then the group must take that attitude toward the proposition.

This seems sensible. Suppose everyone in my group judges that the Bruins will win the playoffs. Then, this unanimous agreement seems to favor treating the group as taking that judgment as well. Similarly, if everyone in the group believes that a die is fair, then the group can be attributed this belief as well. As a rule for evidence,

> **Unanimity-E:** for any attitude type and proposition, if every member of a body of evidence[65] when taken alone requires adopting a certain attitude towards the proposition, then that whole body of evidence requires taking that attitude towards the proposition.

The unanimity rule looks a little less plausible as a rule of evidence or testimony, but it is still easy to build a prima facie case. If all your evidence points in favor of some proposition, then you should believe that proposition. Similarly, if all of your evidence tells you a coin is fair, then

[65] It might be difficult to unpack this phrase precisely. Evidence need not be easily individuated. I had alternatively included "if there is a partition of one's evidence such that every member of the partition, when taken alone…". Perhaps that is more satisfactory to some readers. In any case, the argument that follows doesn't depend on a particular understanding of individuating evidence members.

your credence should be .5 that it will come up heads on its next flip. It is at least tempting to generalize these two simple cases into the general rule expressed by Unanimity-E.

This generalization turns out to be mistaken. To see this, let us return to the example of Tuberculosis testing. Suppose you run five tests, and all of them are positive for Tuberculosis. Let's further suppose that a single positive test favors adopting a credence of .7 that the patient has TB. According to Unanimity-E, we should set our credences to .7 in the proposition that the patient has TB. But this is a mistake. Instead, we should be much more confident than .7 in this proposition. Supposing that each test is independent, the combined body of evidence favors a credence of $1 - .3^5 = .99757$. Even if the tests are not independent, the combined body of five tests should cause us to be more than .7 confident in the TB hypothesis unless the tests are perfectly correlated.[66] Having unified evidence does *not* mean that you should adopt the same view that any single piece of evidence would favor. Considering the body of evidence as a whole might suggest a different course of action in cases where the evidence is synergistic.[67]

But synergy cases are not the only ones that run counter to Unanimity-E. Suppose you visit a foreign country and ask people about the government, and it turns out that everyone, without exception, has only positive things to say: "The government is perfect. There's virtually no corruption, and our leaders only have our best interests at heart." After hearing the first two, three, maybe ten pieces of testimony to this effect, it seems pretty reasonable believe that this is true. You ought to believe that the government is good, and you should be confident in that

[66] To use Wittgenstein's example in the *Philosophical Investigations*, one shouldn't become more confident about a story in the newspaper by buying and reading four more copies of the same paper. See Sober's *Independent Evidence About a Common Cause* for a more in-depth treatment of the importance of independence.

[67] See Easwaran et al (2016) and Hart (under review).

belief. However, if *everyone* echoes this sentiment, you might begin to grow skeptical. If one thousand pieces of univocal testimony pile up, you should probably be more confident in a competing hypothesis: your evidence is being biased or tainted in some way. Perhaps dissent is punished harshly, and as such anyone who would dissent is already in prison, or worse. Or maybe you are gathering your evidence in a non-random way; it could turn out you were surveying people in an affluent area where very few had any qualms with the government.[68]

We might end up with similar biases in non-testimony evidence cases. Suppose I am searching through journal articles on the deterrent effects of the death penalty and I find a consensus of opinions in favor of its deterrent value. On the one hand, this would speak in favor of being more confident in the death penalty having a deterrent effect. But on the other, I should be surprised to find a consensus on such a controversial topic. Perhaps this should cause me to lower my confidence a touch and to consider whether my search has led me to a biased evidence set. Maybe my search terms were such that positive deterrent effect articles were more likely to come to the surface.

These two reasons given for rejecting Unanimity-E are also reasons to reject Unanimity-A. Again, suppose that my group is composed of individuals, perhaps a panel of doctors at a hospital, who all assign credence .7 to the proposition that some patient has Tuberculosis. If it turns out that each of these doctors formed those opinions by running their own independent

[68] We can characterize this in Bayesian terms. Let your credence function be a probability function Pr(.). We want to determine Pr(government is great|everyone testifies that the government is great). By Bayes' Theorem, Pr(government is great|everyone says it is) = Pr(everyone says it is|government is great)*Pr(everyone says it is)/Pr(gov't is great). This quantity will be very small because Pr(everyone says it is) is very small: we are virtually certain that there will be some dissenters in any political system. You can tell a similar story for why your credence will increase that the government is bad given this body of unanimous evidence.

tests, all of which turned out positive, then the group credence should actually be much higher than .7.

This rejection of Unanimity mirrors our experiences in groups. Suppose you are .6 confident that Hillary will defeat Trump in the Presidential Election of 2016. Now, suppose you find yourself in an entire room full of your colleagues, all of whom assign a similar credence—one that's over .5, but not by a large margin. The group consisting of you and your colleagues will likely act in ways that reflect a confidence higher than any of your individual confidences. In this case, we want the method for attitude aggregation to assign a higher than .6 credence to the group consisting of you and your colleagues. How much higher? That depends on the reasons that you and your colleagues have for assigning .6 individually. I will offer a more thorough solution later in this chapter.

Rule 2: Non-Dictatorial

A dictatorial rule is one in which we identify the group attitude with some individual's attitudes. That individual is the *dictator*. In such a group, if you hold the dictator's attitudes fixed, you can vary the other group members' attitudes however you like without changing the group attitude.[69] Conversely, if you hold every other group member's attitudes fixed, then you can still change the group attitude however you like by varying the dictator's attitudes. A non-dictatorial rule is one that prohibits the aggregation rule from making dictators.

[69] Arrow (1965) allows for an exception to this claim. If the dictator is indifferent between two states of affairs, then perhaps the group preferences are offloaded to other group members. This only arises in practical cases where the group agent must act, however. Until then you can simply treat the group preference as indifference as well.

> **Non-Dictatorial-A:** Consider a group G with members g_1, g_2, …, g_n. The aggregation function F is dictatorial if there is some g_i such that $F(G) = F(g_i)$ regardless of the attitudes of the rest of the members of G; F is non-dictatorial otherwise. The solution to the problem of attitude aggregation is non-dictatorial.

This rule is very plausible for the reasons laid out in chapter 1. There are certainly groups that function dictatorially, but this is not a particularly interesting target for aggregating attitudes. For one, aggregating a dictatorial group's attitude is easy: just figure out what the dictator's attitudes are. For another reason, even dictatorially structured groups can be treated as having democratically structured attitudes: what would the attitude be if the decision-making and acting process were in fact structured democratically? If an aggregation function is dictatorial, it fails to treat each member's views as equally valuable in determining the group attitudes.

> **Non-Dictatorial-E:** Consider a body of evidence E consisting of e_1, e_2, …, e_n. The aggregation rule is dictatorial if there is some e_i such that $F(E) = F(e_i)$ regardless what E/e_i contains; F is non-dictatorial otherwise. The solution to aggregating evidence is non-dictatorial.

There should be no single piece of evidence that, merely in virtue of existing, dominates all other pieces of evidence. For instance, suppose a scientist ran a series of experiments, but the data collected from the third experiment, whatever those results may turn out to be, would fully fix her views on the subject. This would be a horrible approach. It might turn out that her views end up aligning with those results, but it is irrational to determine that prior to considering all the evidence.

As written, however, Non-Dictatorial-E is false. Suppose we are considering our evidence for some proposition P, and it turns out that P is in our evidence. Then F(your

evidence) will require that you believe P, regardless of what the rest of your evidence consists in. Similarly, suppose a piece of evidence entails P. Again, you will be compelled to believe P regardless of what else is in your body of evidence. Similarly, if your evidence contains or entails ~P, that piece of evidence will be dictatorial. Similarly, evidence about the objective chance of P might be dictatorial in the above sense. For instance, suppose a perfectly reliable oracle tells me that P is the result of some chance process and that P will happen with likelihood .2. Then this chancy evidence will fix my doxastic attitudes towards P. It's clearly rational to believe P based on having P, something that entails P or ~P, or chance evidence about P in your evidence, so for Non-Dictatorial-E to work, we'll need to revise it:

> **Non-Dictatorial-E*:** Consider a body of evidence E consisting of e_1, e_2, …, e_n. The aggregation rule is dictatorial if there is some e_i that does not entail the truth or falsity—or objective chance information—of any of the propositions under consideration such that $F(E) = F(e_i)$ regardless what E/e_i contains. F is non-dictatorial otherwise. The solution to aggregating evidence is non-dictatorial.

This is better, as it brackets off the problem just discussed. While I think that Non-Dictatorial-E* is a good one for evidence, I want to consider and respond to an apparent counterexample. Suppose you have a body of evidence, but one of the pieces of evidence is especially good:

> **Dominant Evidence:** You are trying to figure out what your spouse is preparing for your anniversary dinner. It's supposed to be a surprise, so she will not reveal it to you when asked. Your evidence consists of the following. First, you know your and your wife's preferences for foods; call this P. Second, you know she purchased some extra white wine from the store this week; call this W. Third, and most importantly, you snuck into the kitchen and it looked like she was making

chicken cacciatore; call this C. So, you form the belief that chicken cacciatore is for dinner.

In Dominant Evidence, it turns out that the third piece of evidence, C, dominates P and W. You would believe cacciatore was for dinner regardless of your and your wife's preferences and regardless whether you had extra white wine. Dominant Evidence poses a problem for Non-Dictatorial-E* because it looks like a counter-example. C appears to be a dictatorial piece of evidence: F(E = {C, E, W}) = F(E = {C}).[70] [71] This sort of objection dissolves on closer inspection. For Non-Dictatorial-E* to be false, we need a case where it is rational to have a single piece of evidence dominate your others *no matter what those other pieces are*. But, in this case you could suppose that E was instead your wife's testimony that she was going to make something with artichoke that smells just like, but is different than, chicken cacciatore. Supposing she is being honest, this means that C no longer dominates the other pieces of evidence. If C does not entail the truth or falsity of the proposition in question, it won't be dictatorial.

This lesson translates to Non-Dictatorial-A. There are potential counter-examples: sometimes groups contain experts such that their attitudes will presumably carry more weight. For instance, a philosophy department hiring a new faculty member might mostly adopt the views of the resident applied ethicist when it comes to determining which hiring practices are fair. Let me spell out the example more carefully:

---

[70] An English translation of the symbols: The rational response to evidence consisting of C, E, and W is the same as your rational response to evidence consisting only of C.

[71] Something even stronger can be said here: F({~C, E, W}) = F({C, ~E, W}) = F({~C, ~E, W}) = F({C, E, W)} = F({C}).

> **Hiring**: The University of Mendota philosophy department is hiring a new philosopher. There are six faculty members, including Justus the applied ethicist. Each individual philosopher has views on what hiring practices are ethical, but each philosopher also would willingly defer to Justus' views on the subject, since he is the resident expert on such matters. As such, when the group attitude is aggregated on issues of hiring practices, it turns out to match Justus' attitudes.

You might be tempted to think that F(.) is dictatorial here, since it just returns a single member's attitudes. However, this would only be a problem if F(.) were to return Justus' views *no matter what the other group members thought*. But that is not the case here; it is crucial that the other group members judge Justus to have the correct views. So, we do not have a counterexample, and Non-Dictatorial-A is safe from this sort of counter-example.

The reasons for reformulating Non-Dictatorial-E to Non-Dictatorial-E* do require us to change Non-Dictatorial-A, though. For there are some cases when a group's attitudes will be dictatorially determined in a non-problematic way. Let us return to the University of Mendota case. Suppose the proposition in question is J, Justus believes Q, where Q is some other proposition in our language. In this case, Justus' attitude towards Q will dictatorially determine the group's attitude toward J because Justus' attitude towards Q entails J (or not J). In attitude aggregation, it is possible for one of the individual's testimony to entail a proposition if that proposition is about some individual's attitudes. These seem like fringe examples, but a good formulation of the rule should avoid them:

> **Non-Dictatorial-A*:** Consider a group G with members g_1, g_2, …, g_n. The aggregation function F is dictatorial if there is some g_i such that $F(G) = F(g_i)$ regardless of the attitudes of the rest of the members of G for all propositions that are not entailed by the attitudes of the members of G; F is non-dictatorial otherwise. The solution to the problem of attitude aggregation is non-dictatorial.

Here I just added a restriction to the propositions in question. It is fine to aggregate over any set of propositions, but there may be a subset of group-referential attitude questions on which it makes sense to have a dictatorial solution.

Rule 3: Member Invariance

Whereas the non-dictatorial rule was not singling out a single group member or piece of evidence as dominant, this rule is about treating the order of each group member or the order of some body of evidence as irrelevant in determining the group attitude or the rational response to the body of evidence.

> **Member Invariance-A (MI-A):** Consider a group **G** with members g_1, g_2, …, g_n. Any group with those same members will have the same group attitude if we hold the individual attitudes fixed.

This rule is almost unassailable. If a group is defined as a set of members, then a group is not an ordered object. So, permuting the members of the group gives us the same set, and hence the same group.

> **Member Invariance-E (MI-E):** Consider a body of evidence E consisting of e_1, e_2, …, e_n. Any body of evidence consisting of those same elements will yield the same rational response.

This rule can be defended in the same way. A body of evidence is a set of propositions that an agent takes to be true. So permuting those propositions does nothing to alter that set, and therefore our agent should adopt the same attitudes when presented with a permuted body of evidence.

You might think we can build a counter-example to the evidential version. It is well established that order matters. For instance, in determining how long it will take for you to repeat a delicious meal, it matters more whether the beginning or the end of the meal had the most gustatory pleasure. We prefer meals that *end* well to those that *begin* well.[72] This perhaps poses a counter-example: two meals that have the same total pleasure will not induce the same response in an agent. Does this tell against Invariance-E? No. Invariance-E is a norm that order of evidence doesn't matter in determining the rational response to the evidence. Let us consider the example more carefully. Call *early* the meal that starts well early, but ends less well, and call *late* the meal that starts less well, but ends excellently. Suppose that our agent gets the same amount of gustatory pleasure from *early* and *late*; the pleasure is just distributed differently. Is it rational for an agent to prefer a repeat of a *late* meal over a repeat of an *early* meal? I see two plausible answers, and either way this does not turn out to be a counter-example to Invariance-E. Suppose it is rational to repeat *late* sooner than *early*. Presumably this is because we humans are wired to enjoy and dwell on endings more than on beginnings when it comes to meals. But, if this is so, then our agent does not have the same set of evidence in *early* as in *late*. The different orders of pleasure constitute a relevant difference in evidence such that the bodies of evidence are not the same. Thus, these meals do not generate a case where the evidence is the same but the rational responses are different. On the other hand, suppose we should treat the evidential situations in

[72] https://www.ncbi.nlm.nih.gov/pubmed/24894582

early and *late* as equal. Then it is irrational for us to prefer a repeat of *late* to a repeat of *early*. This is a common result of our learning about our own psychological biases. Discovering that I have an implicit bias against those with darker skin does not justify my subconscious favoring of whites over non-whites. Rather, we are shocked by such findings because it reveals irrational bases of our behaviors. Similarly, one can argue that we should be indifferent to *early* and *late*, and it is a strangely irrational disposition to focus on the endings of two equivalent situations. Again, the comparison of *early* and *late* does not generate a counter-example to Invariance-E: the evidence is identical, but the rational response should also be identical.

Rule 4: Propositional Invariance

The previous rules were about the group members and evidence. This rule is about the propositions that are either in the agenda or are relevant to the evidence. Propositional invariance requires that we treat each proposition equally.

> **Propositional Invariance-A (PI-A):** Take some group G's profile P, with an agenda containing at least p and p*. Then take the A-subset of P and call this profile R.[73] If R's p-subset and p*-subset are equal,[74] then $A_G(p) = A_G(p^*)$. PI is a neutrality constraint among propositions: if two propositions have the same set of attitudes held toward them, then the group attitudes of those propositions should match.

[73] I do this because P could contain multiple attitudes, in which case the p- and p*-subsets of R might match even though the p- and p*-subsets of P do not.

[74] Two sets A and B are equal just in case every element in A is in B and vice-versa. That is, A = B iff $(\forall a)(a \in A \rightarrow a \in B) \& (\forall b)(b \in B \rightarrow b \in A)$.

The best prima facie defense of PI-A is to consider an example. Consider a philosophy department as a group again, and suppose we are interested in an agenda that includes these two propositions: (1) The department should hire a new ethicist, rather than an epistemologist next year, and (2) the AV equipment in the department seminar room needs improvement. Further, suppose that exactly the same number of philosophers in the group judge that (1) is true as judge that (2) is true. PI-A is the intuitive constraint that the group judgment of (1) must match the group judgment of (2).

> **Propositional Invariance-E[75] (PI-E):** Any two propositions with the same level of support must have the same aggregated doxastic attitudes.

There is a general sentiment that it is rational to "follow the evidence where it leads us". Notice how this sentiment is neutral with respect to the proposition at hand, as it would undermine the quote to restrict the propositions to which we are led.[76] Another metaphor also favors PI-E: weighing the evidence. If I have evidence in favor of P and evidence in favor of Q, and this evidence is of equal weight, then I should adopt the same attitude toward P and Q.

As intuitive as PI-A is, it is problematic. The impossibility theorems in chapter 2 exploit this rule. We can build cases where the agenda contains the premises and conclusion of a valid argument where as many people judge the premises to be true as judge the conclusion to be false. This should make us skeptical of PI-A.

In chapter 2 I presented the structure of an evidential discursive dilemma. Here the structure for a discursive dilemma:

---

[75] I call this "Support Constancy" in Chapter 2.

[76] To compare: "follow the evidence where it leads, unless it leads us to believe a complex-sounding proposition". Hopefully the reader will agree that this "unless" clause defeats the motivation for the heuristic.

	p_1	p_2	~(p_1&p_2)
e_1	Favors	favors	disfavors
e_2	Favors	disfavors	favors
e_3	Disfavors	favors	favors

Presenting the structure may not convince the skeptical reader that such a case is possible. So, consider the following scenario, call it *Forecast*. I am a weather person trying to determine whether the Fall weather for tomorrow will be rainy and cool (below 70° F, say). I have three pieces of evidence. First, there is a low-pressure system blowing in that will cool the air, and therefore likely condense water in the clouds to produce rain. Second, this Fall has been warm and drizzly. Third, there's a cool, dry wind blowing down from the northeast. Let us represent this evidence in an evidence profile table:

Forecast	Rain tomorrow	Be below 70° F	~(Rain and be below 70°)
Low pressure system	favors	favors	disfavors
Warm, drizzly Fall	favors	disfavors	favors
Cool, dry winds	disfavors	favors	favors

As you can see, we have the same level of evidential support (two favors, one disfavors) for each proposition. They cannot all be true, and they cannot all be false. So, PI-E must demand that we withhold judgment with respect to each of these; otherwise PI-E would demand us to be inconsistent. But this is not a good solution either. Imagine we augment the above weather scenario by adding another unrelated proposition that has two pieces of evidence in favor of it and one piece of evidence disfavoring it:

Forecast+	Rain tomorrow	Be below 70° F	~(Rain and be below 70°)	Unrelated P
Low pressure system	favors	favors	disfavors	neutral
Warm, drizzly Fall	favors	disfavors	favors	neutral
Cool, dry winds	disfavors	favors	favors	neutral
Evidence 1	neutral	neutral	neutral	favor

Evidence 2	neutral	neutral	neutral	disfavor
Evidence 3	neutral	neutral	neutral	favor

Supposing Evidence 1-3 and P are irrelevant to the weather-related propositions above, it is plausible that I should believe P in light of the two to one evidence votes in favor of it.[77] PI-E is now in a serious bind:

1. If PI-E requires us to believe "Rain tomorrow", then it requires us to believe contradictory things in *Forecast*.
2. If PI-E requires us to disbelieve "Rain tomorrow", then it requires us to believe contradictory things in *Forecast*.
3. If PI-E requires us to withhold belief on "Rain tomorrow", then it requires us to withhold belief on other propositions independent of "Rain tomorrow" that we have good reason to believe in *Forecast+*.

None of these are acceptable, and so I reject PI-E.

Still, this begs the question: what should we believe in cases like *Forecast* and *Forecast+*? Evidential discursive dilemmas draw out conflicts between premises and conclusions. Here, we may want to say our premises are true, but these premises entail the falsity of a conclusion with the same level of evidential support. Premises are more basic than conclusions. Consider the following:

	P	**Q**	**P v Q**
e_1	favor	disfavor	?
e_2	disfavor	disfavor	?

[77] There are certainly cases where one piece of evidence can outweigh another two (see my discussion on this in §4.2). However, all I need in this example is to show that there are cases where the two outweigh the one.

How would you replace the '?'s in the above table? Intuitively, they should be replaced with 'favor' and 'disfavor' moving from top to bottom. This is what I mean by the premises being 'more basic'. The level of evidential support for conclusions like PvQ is derived from the level of support for the premises P and Q. On the other hand, that e_1 favors P is taken as basic in the representation above.

With this discussion in mind, I propose the following amendment to PI-E:

> **PI-E*:** Any two propositions with the same level of *non-derived* support must have the same aggregated doxastic attitudes.

The corresponding amendment for PI-A is:

> **PI-A*:** Take some group G's profile P, with an agenda containing at least p and p*. Then take the A-subset of P and call this profile R. If R's p-subset and p*-subset are equal, *and if every $A_g(p)$ and $A_g(p^*)$ is non-derived*, then $A_G(p) = A_G(p^*)$. PI is a neutrality constraint among propositions: if two propositions have the same set of attitudes held toward them, then the group attitudes of those propositions should match.

While the technical notion of some $A_g(p)$ being non-derived is more complicated, the basic idea is simple. An agent's attitude is derived just in case they form that attitude as the result of holding some other doxastic attitude. For instance, if I believe that the Seahawks will win the Super Bowl (partly) because I believe that they will win 11 games this season, and that any team that wins 11 games this season is very likely to win the Super Bowl, then my belief that the Seahawks will win the Super Bowl is (partly) derived.

The non-derived components of the definitions of PI-E* and PI-A* are not formal. It is not sufficient that a proposition looks like a conclusion to determine that its support is derived.

Rather, the derived nature of a proposition's support—be it evidential or in terms of attitude aggregation—depends on how the support was actually built. To see this, let's revisit a variant of the Seahawks example from the previous paragraph. Suppose that I *first* form the belief that the Seahawks will win. And, since I also believe that any team that wins 11 games this season is very likely to win the Super Bowl, I form the belief that the Seahawks will win 11 games.[78] In this case, the belief that the Seahawks will 11 games is derived from my other doxastic attitudes.

There is another objection that applies equally well to both PI-A and PI-A*: there are democratically structured groups that treat propositions differently. As an example, suppose that a western democracy required a supermajority (greater than 67% of the vote) for judicial appointments but only a simple majority (greater than 50% of the vote) to pass resolutions. In this case, a resolution might have the same level of support as a judicial appointment but be treated differently, if that level of support was greater than 50% but less than 67%. Prima facie, this looks like a case where a democratically structured group violates PI-A*. I do not think this objection succeeds in defeating PI-A*, but the objection is complicated. There are two main responses I can give.

1. *The group is democratically structured, but is not treating propositions differently.* In the above scenario, let us identify the two propositions in question: J: Judge Judy should be appointed to the Supreme Court, and R: The resolution should pass. If the only atomic propositions in the agenda were J, and R, then the group does violate PI-A*. However, this is likely an impoverished snapshot of the attitudes of the group members. These group members also will have attitudes about whether J should

[78] It doesn't matter if my inference here is fallacious. But, I could tweak the example to include some statistical information that would prevent me from affirming the consequent.

require a supermajority in order for it to be enacted. Suppose every group member was certain that J is true iff at least two thirds of the population believed J. Similalry, suppose each person is certain that R should only require a simple majority. In that case, there is a relevant difference between the group support for J and R.[79]

2. *The group is not actually democratically structured.* Some of the strength of the objection may stem from the fact it is a *democratic* government, but this does not mean that group is *democratically* structured. It may just be unfortunate that the words used are the same. Even so, the western democracy as described is not a pure democracy, as that would require simple majority votes by the entire population on every action. Instead, representative democracies that add supermajority requirements on judicial appointments, among other such structural features in the decision-making process, are no longer democratically structured in the sense of groups as discussed in this dissertation. These groups will still be close enough to democratically structured so that the Steward Approach will yield a good approximation of their attitudes, but the Steward Approach may not quite reflect the group's actions insofar as these structural features violate the constraints discussed in this chapter.

This is a difficult objection, but I am optimistic that some combination of the above responses will resolve any fully explicated scenario where voting procedures may differ between two different (classes of) propositions. If not, PI-A* would need to be weakened to accommodate the other cases where a democratically structured group may treat two propositions differently despite their having the same level of support.

[79] It is possible that the difference in support is not captured by the 'derived' condition added in PI-A*. This may suggest that further refined principle, say PI-A**, is needed.

Rule 5: Representative Evidence

Rules five and six are different from the first four rules. For both five and six, I will be focused on an evidential rule and then conclude the section by discussing the implications for attitude aggregation. Rule five is that one should treat their total body of evidence as representative of the state of the world. How representative should we take the evidence to be? There are two major contenders for the rule:

> **Perfectly Representative Evidence (PRE):** At any time, an agent should judge that their total body of evidence is *perfectly* representative of the way the world is.
>
> **Roughly Representative Evidence (RRE):** At any time, an agent should judge that their total body of evidence is *roughly* representative of the way the world is.

Let us consider how PRE and RRE function in an example. Suppose there is an urn that contains 10 balls. You know nothing else about the urn. You observe 10 draws with replacement, of which 3 were white and 7 were black. Here are two verdicts you might take from Rule 5:

> ***PRE:*** You should believe[80] there are 3 white and 7 black balls in the urn.
>
> ***RRE:*** You should believe there are *roughly* 3 white and 7 black balls in the urn.

One violates PRE if they take any other attitude than believing there are 3 white and 7 black balls in the urn. On the other hand, RRE gives some wiggle room. I will not present an analysis for "roughly" here. But, the notion is tied up with the notion of misleading evidence. RRE is the requirement that you not think your total body of evidence is seriously misleading. So, if you believed that there were 9 white balls and 1 black ball, you would hold that the chances of

[80] I'm just focusing on belief for simplicity. PRE and RRE would have implications for each doxastic attitude, but belief is the easiest case to unpack and hence is the most appropriate for describing the thrust of the rule.

observing the 3 and 7 as you did were quite low, and so the evidence was misleading. Hence, RRE would indict the belief that there were 9 white balls as irrational. PRE entails RRE, but not vice-versa. I will defend the weaker RRE here.

The most compelling defense of RRE is a practical one. If you were not constrained to treat your evidence as representative, then you risk being crippled by all of the possible states of the world for which your evidence would be misleading. A problem with conspiracy theorists is that they often fail to change their minds in the face of overwhelming evidence. The same problem can plague someone who violates RRE: if you are willing to entertain that your evidence is radically misleading, very few pieces of evidence will actually cause you to revise your doxastic state. This problem is particularly debilitating for certain clusters of views on epistemology, namely ranged credence views while maintaining an impermissive epistemology (see Castro and Hart 2017).

Unfortunately, it is very hard to provide a strictly epistemic defense of RRE (or PRE). This is in part because our evidence is not always representative. In hindsight, all of us can recall many times where our evidence has not been representative of the state of the world. A child might think they are a great singer when their only evidence has been compliments from their parents and hearing themselves sing in the shower. From the outside, we can see how this evidence is biased in favor of their being a good singer. But, from the inside, one can't tell when their evidence is misleading.[81]

This leads to my defense of RRE. Epistemic norms, like any other sort of obligation, are such that 'ought implies can'. But, one cannot know when one's evidence is misleading, and thus one cannot be obligated to treat evidence as misleading. Still, RRE is stronger than this; it

[81] See my *Dissolving Harman's Dogmatism Paradox* (manusript).

requires that an agent treat the evidence as roughly *not* misleading. As much as I would like to offer further justification for this claim, it is just part of what rationality is to respond to one's evidence as if it were not misleading.

Considering apparent counter-examples will show that we are disposed to treat evidence as representative. Suppose there is an NBA player who makes eight of the first nine three pointers that he attempts. Should we conclude that he is perhaps the best NBA three-point shooter in history? After all, no one has been a career 90% three-point shooter.[82] Of course not: the sample size is too small. This may seem to be a counter-example for RRE (and hence PRE): this is a case where our evidence favors a certain assessment of his shooting accuracy, yet we are unwilling to believe that he is indeed that good of a shooter. Despite appearances, this is not a counter-example to RRE. RRE requires that we treat our *entire* body of evidence as non-misleading, but here we are only considering our evidence from the first nine shots the rookie took in his NBA career. We have much more evidence about three-point shooting than this: we know, as was already mentioned, that no one has had that high of a career shooting percentage. We also know that shooting three-point shots is a relatively high variance activity, and that high variance activities are subject to producing misleading results when only small sample sizes are considered. Hence, we are treating our full body of evidence as representative, as prescribed by RRE.

Let me now pivot to the attitude aggregation version of RRE:

Roughly Representative Attitudes (RRA): At any time, a group's attitude is roughly represented by the attitudes of those individuals who comprise the group.

[82] There may be some counter-example out there of a big (person who plays center or forward on a basketball team) who only took and made a single three point shot in his career, but otherwise I am confident in this assertion.

This approximates the constraint that we are interested in democratically structured groups, and the support for this underlies our rejection of dictatorial rules. RRA gives the group a little wiggle room, whereas a Perfectly Representative Attitudes (PRA) rule would demand that the group attitude perfectly matches those of the group members. Both RRA, which I advocate, and PRA provide a general condition for adequacy of a solution to the attitude aggregation problem: if a solution ends up with group attitudes that are very different from the group member attitudes, then the solution is a bad one.

Rule 6: Uniqueness

Evidentialism is the claim that one is justified in their doxastic attitudes towards a proposition at a time just in case their evidence supports those doxastic attitudes. Conee and Feldman (2004) sum up the view as follows:

> "**EJ:** Doxastic attitude D toward proposition p is epistemically justified for S at t if and only if having D toward p fits the evidence S has at t."

Other characterizations of evidentialism are characterized by how they answer three core questions: 1) what is evidence? 2) what does it mean for an agent to 'have evidence'? and 3) what does it mean for a doxastic attitude to fit the evidence? I am sympathetic to evidentialism, broadly construed, and in fact endorse an evidentialist principle in Rule 5: RRE (and derivatively RRA). According to RRE, attitudes fit the evidence just in case one takes their total body of evidence to be representative of the way the world is.

However, I am less sympathetic to a stronger claim that is supported by evidentialists such as Richard Feldman and Roger White. Feldman and Cullison (2014):

> *Same Evidence Principle: If A and B have the exact same total body of evidence, then A and B are justified in believing the same propositions to the same extent.*

Elsewhere, White (2005) advocates for the uniqueness thesis:

> *Given one's total evidence, there is a unique rational doxastic attitude that one can take to any proposition.*

Both of these claims stem from a tightening of the connection between justification and evidence. Whereas EJ posits an uncontroversial 'fit', the same evidence principle and the uniqueness thesis demand a one-to-one correspondence between bodies of evidence and doxastic states.[83, 84] There is something appealing about these claims. If true, they would mean that rational agents should never disagree when they pool their evidence. And, they would mean that there is some objective relation between evidence and propositions that we epistemologists could attempt to ferret out.

As much as I may wish it were true, the uniqueness thesis is false. There are many cases where there is too little evidence to favor a single doxastic attitude over others.[85] There are multiple permissible starting points or standards of reasoning that one can apply to the evidence.[86]

Rejecting uniqueness has implications for the Steward Approach. While I maintain via RRA that a group's attitude must fit the attitudes of the individuals that comprise the group, denying uniqueness means that the Steward may have more than one rationally permitted

---

[83] As Titelbaum and Kopec (forthcoming) point out, there are several notions of uniqueness to be teased apart here. I will glibly proceed with a less-refined treatment, since the details turn out not to matter in my rejection.

[84] Strictly speaking, the text of the same evidence principle only says that two individuals are justified in having the same attitude, and it does not rule out them being justified in having different attitudes. Given that Feldman accepts uniqueness elsewhere, my reading the stronger claim into the same evidence principle is warranted.

[85] See Castro and Hart (2017).

[86] See Titelbaum and Kopec (2017).

response to a given group's attitude profile. The following are the most plausible aggregation rules I can think of in cases of permissive situations for the steward:

> **Non-Uniqueness Broad (NUB)**: Whenever there are multiple permissible responses the Steward could take in response to a group's profile, the attitude of the group is the broad attitude that encompasses all of the steward's rational constraints.
>
> **Non-Uniqueness Indeterminable (NUI)**: Whenever there are multiple permissible responses the Steward could take in response to a group's profile, the attitude of the group is indeterminable among the permissible responses.
>
> **Non-Uniqueness Privileged (NUP)**: Whenever there are multiple permissible responses the Steward could take in response to a group's profile, the group attitude is identical to one of those rational stances permitted to the steward, the privileged rational stance.

NUB gives a unique solution to group attitudes, and for this reason I prefer it over NUI and NUP. However, in practice, these three methods aren't all that divergent. Let's consider an example. Suppose the NAACP is unified in attitudes that African Americans have been historically oppressed and that actions should be taken to address those oppressions. But, we can also suppose that the NAACP members are divided on many issues, which would make it such that the Steward is permitted to take a range of doxastic attitudes toward. For instance, members of the NAACP may diverge on whether abortion is permissible: evangelical African Americans may be Pro-Life while more progressive members are Pro-Choice. Let us suppose that, given all and only the testimony of NAACP members as evidence, the Steward is permitted to have credence anywhere from .2 to .8 that abortion is morally permissible. What do NUB, NUI, and NUP say about what the NAACP's credence is as to whether abortion is permissible?

> NUB: The NAACP has the ranged credence (.2, .8).
>
> NUI: The credence is not outside of (.2, .8), but it's not fixed.
>
> NUP: The credence is somewhere between .2 and .8.

One might supplement NUP with a procedure for determining which attitude is privileged. I am skeptical that such a method exists: if it did, then we could use that argument to show that the Steward wasn't rationally permitted to choose any value in (.2, .8), contra the assumption. Notice that, whichever rule one adopts, the upshot of the divergence is that the NAACP doesn't have an easily identifiable precise attitude about the moral status of abortion. I think this is the right result.

Bodies of evidence, and attitudes of groups of people, may sometimes permit multiple rational responses. This is not a weakness for my solution to the attitude aggregation problem. In cases where there is a broad range of permissions, the group can be thought of as having a weaker, less unified stance on the matter. Analogously, when one has conflicting pieces of evidence relevant to some proposition, one should have less strongly held opinions about that proposition.

§4.2 Aggregation by Cases

This section will contain my constructive argument for how to aggregate evidence, and hence group attitudes using the Steward Approach. There are many sorts of things to iterate through:

- Attitudes: judgment/belief and credences
- Agenda: from simple, single proposition and its negation sets to large, integrated agendas
- Evidence: from small, single pieces of evidence to larger, multi-modal bodies of evidence

Because the complexity can ramp up in each of these three dimensions, there is no single obvious way to develop my account from simplest case to most complex case. Since the ternary belief attitude is simpler than the infinite credence values, I will start with beliefs. I will then

walk up through increasingly complex bodies of evidence and agendas. After, I will do the same for credences.

§4.2.1 Beliefs

In this section I will discuss the procedure for aggregating evidence for the ternary belief attitude. For each proposition an agent considers, the agent can believe (yes), disbelieve (no), or withhold belief (w). There are many related variants of beliefs and judgments. You might treat judgment or belief as binary, or perhaps one is stronger than the other. These different belief-type attitudes will be similar enough to the ternary beliefs discussed here that one can easily translate my rationality conditions to them.

In each of the following cases I will consider some number of propositions (p) and pieces of evidence (e). When there is more than one proposition or piece of evidence I will use subscripts to distinguish them. So, the default in the one proposition, one piece of evidence case is to use p and e. But, when there are two propositions, I will call them p_1 and p_2, and so on.

§4.2.1a One proposition, one piece of evidence

Suppose you are considering a single atomic proposition and you have exactly one piece of evidence relevant to p. A piece of evidence is *relevant* to some proposition just in case that piece of evidence rationally requires or permits a change in one's doxastic stance toward that proposition. In the context of belief, irrelevant evidence neither favors nor disfavors a proposition. Intuitively, the fact that it is raining right now is not relevant to the proposition that I own a cow. On the other hand, that it is raining right now is relevant to whether traffic will be slow on the highway this evening. In the coming pages I will restrict myself to pieces of

evidence that are relevant to at least one of the propositions in the agenda. Evidence that is irrelevant to every proposition in the agenda is easy to aggregate over: the evidence is inert with respect to one's doxastic state.

In a one proposition, one piece of evidence case, there are two possibilities: either the evidence favors p, or it disfavors p. I will argue for two types of rational constraints here, one weak and one strong. Here is the weak claim:

Scenario	**Rational Prohibition**
e favors p	Disbelieve p
e disfavors p	Believe p

The weak claim is that a single piece of evidence in favor of p need not compel one to believe p, but it does prohibit the belief in not-p. RRE favors both this weak claim as well as the strong claim that follows. RRE requires agents to treat their total body of evidence as non-misleading. Here is a stronger claim:

Scenario	**Rational Prohibition**
e favors p	Believe p
e disfavors p	Disbelieve p

This stronger claim is supported by RRE as well. Believing p when your entire body of evidence (albeit just a single piece of evidence) favors a proposition is to treat the evidence as representative of the world. Moreover, the discussion on Unanimity supports this strong claim. Recall that I reject Unanimity as a rule because there are two types of counter-examples: 1) cases where the agreements stack synergistically, and 2) cases where the agreements are unexpected and thus constitute evidence for an alternative hypothesis. In this case, neither counter-example is applicable: there is only one piece of evidence, so there is no synergy to worry about. And, the single piece of evidence is your entire body of evidence, so there is nothing else on which to base any alternative hypotheses.

It is very difficult to imagine a case where one's total body of relevant evidence is just a single piece of evidence. In most real-world cases, even when I only have one piece of evidence that is clearly about some proposition, I still have a much larger body of evidence that bears on how I process that single piece of evidence. Imagine you are teleported to an alien planet. You observe some creature-like thing seeming to drink from a colored liquid. You are wondering whether the colored liquid is potable. My weak claim is that you shouldn't disbelieve that it is potable. My strong claim is that you should believe it is potable. I think this is the right verdict. Of course, it would be foolish to act on such beliefs until you had gathered more evidence, if possible. But, given the one piece of evidence you have favors the liquid being safe to drink, you should take that evidence to accurately represent the world.[87]

§4.2.1b One proposition, two pieces of evidence

In the case where one's total body of evidence contains two pieces of evidence relevant to a single atomic proposition, we can split the cases into two main groups. First, both pieces of evidence could be aligned insofar as they both favor or both disfavor p. In this case, the verdict will be the same as in the one proposition, one piece of evidence case. If anything, the case for the strong claim in that section is even stronger. The interesting case in this section is when the pieces of evidence disagree:

	p	~p
e_1	Favors	Disfavors

[87] As already noted in this paragraph, it is a stretch to pretend one has no other evidence. We have seen other creatures drink things that are unsafe for us to drink, and other creatures drink potable water as well. And you will likely not receive just a single piece of evidence from your observation: our visual field is ripe with information.

e_2	Disfavors	Favors

Here, one must compare the relative strengths of the favoring/disfavoring relation. This is the process of weighing the evidence, after which the problem will reduce to the one proposition, one piece of evidence case. I am unaware of any fully general account of what makes a piece of evidence weightier than another with respect to some proposition and body of evidence. I trust you will agree that any two pieces of evidence will have weights, and thus there are three scenarios:

3. e_1 is weightier: one should believe p.
4. e_2 is weightier: one should believe ~p.
5. e_1 and e_2 are of equal or incommensurable weight: one should withhold belief on p.[88]

My defense in §4.2.1a is the same as for these three claims. For, in the even that, say, e_1 is weightier, you can reformulate the profile table as follows:

	p	~p
e_1 & e_2	Favors	Disfavors

But, this is just a case of one piece of evidence that favors p over ~p. Similarly, if e_2 is weightier, that will reduce to case where ~p is favored by a single piece of evidence. In the case where the two atomic pieces of evidence are equally weighty or incommensurate, your evidence neither favors nor disfavors p. For example, if p is "the next draw from the urn will be a white marble",

[88] If you think this trio is too strong, you could opt for the weaker: 1. not believe ~p, 2. not believe that p, and 3. withhold belief on p.

then two draws with replacement where the first is white and the second is black will neither favor nor disfavor p.

§4.2.1c One proposition, three or more pieces of evidence

Because any two pieces of evidence can be combined into a single piece of (conjunctive) evidence, cases with more than two pieces of evidence relevant to a given proposition will reduce to one of the two cases already considered. Either all of the evidence will favor (disfavor) p, in which case one should believe (disbelieve) p as per the first case, or the evidence is split, with some favoring p and others disfavoring p. In the second case, we can conjoin the favoring evidence and conjoin the disfavoring evidence, and then we are in the case discussed in §4.2.1b.

§4.2.1d Two or more unconnected propositions, One or more pieces of evidence

Two propositions are *unconnected* just in case they are independent and remain independent in light of a body of evidence. This has the consequence that you can split up the evidence profile tables without changing any rational requirements for the agent. This case is trivial, given §4.2.1a-c. If we can separate unconnected propositions into smaller (that is, fewer-columned) evidence profiles, then each evidence profile will be one of the cases already considered. The trickier cases come next, where the propositions are connected.

§4.2.1e Two or more connected propositions, one or more pieces of evidence

There will be simple cases here as well, such as cases where the evidence univocally favors consistent propositions. However, the difficult cases are the interesting ones. And, the difficult

cases are the ones that give rise to discursive dilemmas. Recall those are cases of the following sort:

DD	p_1	p_2	$\sim(p_1 \& p_2)$
e_1	favors	favors	disfavors
e_2	favors	disfavors	favors
e_3	disfavors	favors	favors

You might also recall that above I suggested PI-E* as the proper response. We can believe p_1 and disbelieve $\sim(p_1 \& p_2)$ if the support for the latter is derived while the support for the former is direct. The more problematic case would be if there were contradictory atomic propositions that had equal support. A case like the following, if p_1 and p_2 were contradictories:

DD*	p_1	p_2
e_1	favors	disfavors
e_2	favors	favors
e_3	disfavors	favors

This sort of case, if it were possible, would pose a rational dilemma. I reject rational dilemmas, and so I must find something amiss with DD*.

The problem with DD* is that e_2 cannot support both p_1 and p_2. For a piece of evidence to favor a proposition, it is necessary that the evidence makes that proposition more likely than if the evidence were false. But this means that, if DD* is accurate, $Pr(p_1|e_2) > Pr(p_1)$ and $Pr(p_2|e_2) > Pr(p_2)$. But by assumption, $Pr(p_1|e_2) = 1 - Pr(p_2|e_2)$. This yields a contradiction:

1) $Pr(p_1|e_2) > Pr(p_1)$ Def'n of favoring
2) $Pr(p_2|e_2) > Pr(p_2)$ Def'n of favoring
3) $Pr(p_1) + Pr(p_2) = 1$ p_1 and p_2 are contradictories
4) $Pr(p_1|e_2) + Pr(p_2|e_2) > 1$ combining 1-3
5) $Pr(p_1|e_2) + Pr(p_2|e_2) = 1$ p_1 and p_2 are contradictories

But, 4 and 5 give us a contradiction. So, we should reject that e_2 favors both propositions.

Despite some bodies of evidence being very difficult to distill into a rational set of beliefs, there are no rational dilemmas. For any set of possible evidence and set of atomic sentences, we can split them up into evidence profiles that conform to one of the above cases. And, in each of those cases, one can weigh the evidence to yield a rational set of beliefs.

§4.2.2 Credences

In §4.2.1, I often referenced an evidence's weight. Unfortunately, I do not think there is a sufficiently general account of weight to offer more precise rules for aggregation. In the case of credences, though, we can utilize probabilities to do the job. This will allow me to be more precise for strategies for combining pieces of evidence. To this end, it will be helpful to place a meaningful value between 0 and 1 in each cell of our evidence profile when aggregating evidence into credences. This will enable us to identify which pieces of evidence are stronger, and thus they can have more pull in determining our posterior credences.

But, there are several difficulties in determining which numbers to place in these cells. Our evidence, inconveniently enough, does not come stamped with probabilities on them. When I look out my window and see some dark clouds blowing in, God does not whisper "that looks like some .8 evidence in favor of rain, Casey." Rather, we do something like the following: Given what I already know, when I see dark clouds blowing in like this the chance of rain is .8. Formally, it looks like this:

Pr(rain|dark clouds plus my previous evidence) = .8

This looks promising. Perhaps we should just put the likelihoods of each proposition given the piece of evidence in our profile table.

As nice as this would be, this solution is unworkable. The project that I am engaged in is to determine what an agent should do when they acquire their first body of evidence. But, this means that all that our evidence-less Steward must work with is his prior credence function. If someone solves the problem of the priors, then so much the better for my approach here: we can simply use conditionalization to update on any body of evidence. But there is no such agreed upon solution. In the absence of such a solution, I need to find some way to meaningfully quantify the strength of evidence.

To motivate my approach, consider the following example. You are out hiking and very, very thirsty. You have already drained your water bottles, and you are looking greedily at a nearby stream, trying to decide whether it is safe to drink. While you wait, you see another hiker head by and drink from the stream. This emboldens you, and so you go and have a drink. Twelve hours later, you still feel fine. Let's build an evidence profile:

Thirsty	Water is safe	Water is unsafe
Hiker drank	favors	disfavors
Feel fine after drinking	favors	disfavors

If we were interested in combining this evidence into belief, the "favors" and "disfavors" above would be sufficient: you should believe the water is safe. However, we need to augment the above profile to include what should be obvious from the details of the Thirsty story: that the hiker drank is far weaker evidence than that you feel fine after drinking the water. After all, it was possible that the hiker got sick shortly after drinking, for all you know.

I propose the following method for filling in evidence profiles for credence aggregation. If a piece of evidence favors a proposition, it receives a number greater than .5. The stronger the evidence, the larger the number. If a piece of evidence entails the proposition in question, it receives a "1". Similarly, evidence that disfavors a proposition gets a value of less than .5, and

the number approaches zero the stronger the evidence is against the proposition. This estimation by strength method is a good approximation of the likelihood of the proposition given the piece of evidence. In the case where these likelihoods are available, then by all means use them. But, in the case where an agent doesn't have such prior credences available to update on, this method of mapping likelihoods onto strengths is a plausible and fruitful approach.

In §4.1.5, I argued for RRE: your evidence is roughly representative of the world. This means that, the stronger the evidence, the more likely the propositions supported by the evidence are to be true. One way to make this explicit is to say that the probability that evidence will have strength m in favor of p given p is proportional to m. While the details are tricky, the idea is simple: we are just cashing out the RRE principle in a mathematical way that will prove useful.

One upshot of cashing out RRE this way is that our assumptions about evidence now match perfectly with a mechanism for combining credences articulated by Easwaran et al (2016). They call their rule UPCO. I have defended a similar rule called Bico, and I present a justification for my version in chapter 6. Since I think the Easwaran et al justification for the rule is most appropriate here, I will use their name in this chapter. The rule is as follows:

> **UPCO**: When combining a set of credences m, n, o, … towards some proposition p, the updated credence will be Pr(p) =
> (m*n*o*…)/[m*n*o*… + (1-m)*(1-n)*(1-o)*…].

This rule has many desirable features. It is *synergistic*: many high (or low) credences can have a synergistic confirmation (or disconfirmation) of each other, sending one's posterior credences higher than each individual one. UPCO is also commutative: it does not matter which order you receive the evidence in (see §4.1.4 and MI-E). Lastly, the rule preserves relations between different propositions, so applying UPCO to a number of propositions and pieces of evidence will not result in your violating any of the rationality constraints of, say, the probability axioms.

With our rule in hand, we can easily apply it to some sample evidence profiles:

	p	~p
e_1	.7	.3

you should assign credence .7 to p. The singleton case is boring mathematically. And, the justification is the same as in §4.2.1a. Let us consider a more complex case:

	p	~p
e_1	.7	.3
e_2	.6	.4
UPCO	.42/(.42+.12) = .78	.12/.54 = .22

This is a sensible result. If you have two pieces of evidence in favor of p, and both pieces are mildly strong, then you should be even more confident in p than if you had only learned either piece of evidence in isolation.

While this synergistic effect is desirable in an updating rule, it is not always warranted. Suppose it is Christmas morning, and Juliette is trying to figure out what present is waiting inside of the colorfully wrapped box. She knows that her parents either got her a heavy chemistry set or a light box of art supplies, which is much lighter than the chemistry set. She picks up the box and feels that it is very heavy. She sets it down, and then picks it up again. We can build the following table:

Present	Present is chem set	Present is art supp.
First lift	.9	.1
Second lift	.9	.1
UPCO	.81/(.81+.01) = .99	.01

Notice that UPCO allows the two strong pieces of evidence to synergize, resulting in a Juliette who is very confident that she will be getting a chemistry set. But, this is problematic, since it

means that Juliette can bootstrap her way to near-certainty merely by lifting the box several times.[89]

Why is it that multiple lifts do not constitute strong additional evidence in favor of Juliette's present being a chemistry set? Getting the epistemological details might be more complex, but it is clear that Juliette does not gain much additional information with her second lift. She already knew the box was heavy, and while the second try may confirm the reliability of her first trial, it should not add much to her confidence that the present is a chemistry set. Formally, there are a few ways to implement this reasoning into our profile table. First, we could merge the two rows into a single row, denying that there is more than one piece of evidence in "Present". This approach would work, but I am leery of getting into the business of counting evidence or propositions. The second option is to replace the '.9' next to the second lift with a number closer to .5. This for the simple reasons that the second lift is not nearly as strong a piece of evidence as the first relative to "The present is a chemistry set." The two pieces of evidence appear equally strong since they are such similar actions; in both Juliette picks up a box. However, the second one contains far less information about the contents of the box for Juliette, since she has already acquired most of that in her first lift. Thus, we get the following evidence profile:

Present	Present is chem set	Present is art supp.
First lift	.9	.1
Second lift	.55	.45
UPCO	.495/(.495+.045) = .912	.088

[89] This resembles the problem of old evidence. Bayesians have traditionally thought of the confirmatory value of evidence for some proposition as the increase that the evidence adds to one's credence in the proposition. But if this is so, then old evidence, which supposedly is fully integrated into one's doxastic state, shouldn't be able to offer any confirmation. In the above case of excessive synergy, the worry is that UPCO might confer too much weight to redundant evidence.

These results are much better. The second lift does make Juliette more confident that she got a chemistry set, but only slightly so. If she lifted it a third time and it felt heavy, the number would be very nearly .5 under "Present is chem set". You might be surprised to see '.45' in the second far right column. At first sight, this struck me as odd, too. But remember that the .45 does not give confirmation that the present is art supplies; the .45 indicates that the second lift only constitutes weak evidence *against* this claim. And that is exactly right.

While I needed to go through a number of cases for belief, the general nature of UPCO makes such an argument by cases unnecessary for credences. It has all of the formal properties needed to ensure that rational inputs will generate rational outputs. The real test for this rule, and all of the rest of the rules accepted and rejected throughout this chapter, is to see whether they yield plausible results for aggregating group attitudes.

Chapter 5: Applying the Steward Approach

In chapter 3, I argued in theory for the Steward Approach, and in chapter 4 I fleshed out some of the details for how that will work in practice. In this chapter, I demonstrate that the Steward Approach yields plausible results for concrete cases of attitude aggregation. I will start with a few "easy" cases, ones where there is not a great deal of conflict among the attitudes of the group members. I will then consider more difficult cases, especially ones that give rise to Condorcet-style paradoxes.

To standardize the form, each of the following sections will feature an example of a group composed of agents g_1, g_2, … g_n. These agents are considering the agenda of propositions P composed of atomic p_1, p_2, … p_n, along with their conjunctions and negations.[90] For brevity, I will only explicitly list the atomic sentences on the profile table unless some molecular (non-atomic) formula is relevant to the example. Also, for each aggregation I will be considering two attitudes, belief and credence. So, each cell in the profile table will be an ordered pair of the form <belief-status, credence> (e.g. <disbelives, .2> indicates the agent disbelieves a proposition and assigns it credence .2). Every example will be an augmented profile table, where the last row is the aggregated attitudes of the group members achieved by the Steward Approach. The task in this chapter is to show that this final row is a viable candidate for the group attitude, thus vindicating the Steward Approach.

[90] Hence also their disjunctions, material implications, and so on.

§5.1 The Single Person Group

The first test for the Steward Approach is to get the single-person group right—a group with only one person in it, not a group of unmarried individuals. Of course, it is not clear that one person can even be thought of as a group, but it will still be useful to show that my theory works even at the extreme edge (non-)cases. Here is my solution to the general case for aggregating a single-person group:

Single-Person	$\mathbf{p_1}$	$\mathbf{p_2}$	**...**	$\mathbf{p_n}$
$\mathbf{g_1}$	$\langle b_{1,1}, P_{g1}(p_1)\rangle$	$\langle b_{1,2}, P_{g1}(p_2)\rangle$	...	$\langle b_{1,n}, P_{g1}(p_n)\rangle$
Steward	$\langle b_{1,1}, P_{g1}(p_1)\rangle$	$\langle b_{1,2}, P_{g1}(p_2)\rangle$	...	$\langle b_{1,n}, P_{g1}(p_n)\rangle$

The Steward's beliefs and credences should match g_1's, and hence the group attitudes are identical to g_1's. I need to defend two things:

1. The Steward's doxastic attitudes should match those of another person if the Steward's only evidence is that set of attitudes.
2. The attitude of a single person group matches the attitudes of the person who solely comprises the group.

I take it that (2) is intuitive. I can see no other plausible candidate for the attitudes of a single-person group than those of its only member. It is the only case where a dictatorial rule successfully captures a democratically structured group.

On the other hand, (1) is more controversial. When we receive testimony in our day to day lives, we rarely feel compelled to match those attitudes exactly. And, even when matching is warranted, this usually is supported by having a body of evidence that supports the reliability of the person whose testimony we are taking. But these cases are not directly analogous to the Steward's case. We are imagining a Steward whose *only* evidence is the testimony of the group members. In chapter 4, I argued that one should believe whatever one's only piece of evidence favors. Since testimony is a specification of evidence, those arguments apply equally well here.

The one piece that requires further argument is to show that credences that the Steward obtains meet the same constraints as the "strength of evidence" numbers that I used in the evidence profiles for aggregating into credences. The assumption necessary for the application of UPCO was that the strength of evidence given p is proportional to its strength. This meant that stronger evidence made p more likely. In the testimony case, the analogous assumption is that g_i's credence in p_j given p_j is directly proportional to that credence. This means that the Steward must assume that group members are accurate indicators of the truth. Is this assumption really a rational requirement? I think so, for the same reasons I used in defending RRE. While there might not be anything inconsistent or incoherent about assuming, say, that every piece of evidence was an anti-indicator of truth, I have no problem with calling such a person irrational. Since testimony is evidence, we should take that evidence to be representative of the way that the world is. This is a defeasible attitude: we should stop believing liars, and we should be skeptical of folks we have reason to be skeptical of. But the Steward's only evidence is the testimony of the group members, and so unless that very testimony gives the Steward reason to doubt its reliability, the Steward should take it as a strong indicator of the truth. Thus, the Steward Approach gets the single-person group aggregation right. Next, we will see how it handles two-person cases.

§5.2 Sports Radio Duo, United

A popular medium for sports talk shows is to have two personalities that have dueling opinions. In this section, I will show how the Steward Approach aggregates a toy case where sports personalities are largely in agreement. In the next section, I will shift gears and see how the Steward Approach aggregates divided hosts.

In the 2016-2017 NBA season, Kevin Durant left the Oklahoma City Thunder to join the Golden State Warriors. This was an unprecedented move, since Durant was a top five player in the league, and the Warriors had already been a historically great team without Durant. Durant's move drew a great deal of ire from sports fans. Some criticized him for his lack of loyalty, while others criticized him for his lack of competitive spirit: you should want to beat the best, not join them. Consider the following profile table and Steward Approach aggregation for a pair of sports talk show hosts:

United Duo	disloyal	uncompetitive	Durant erred
Mike	<believe, .97>	<believe, .6>	<believe, .99>
Mad Dog	<neutral, .5>	<believe, .9>	<believe, .85>
Steward	<believe, .97>	<believe, .93>	<believe, .998>

Mike and the Mad Dog both think that Durant erred in going to Golden State, and the Steward Approach yields a group attitude that is stronger than either host's individual attitudes. Also notice that Mad Dog is indifferent as to whether Durant was disloyal, and as such the aggregated attitude matches Mike's with respect to that proposition. This does not violate our non-dictatorial rules: Mad Dog's views are being factored into the group attitude just as Mike's would be if their views were switched.

§5.3 Sports Radio Duo, Divided

The previous example featured our sports duo with synergizing beliefs, but these talk shows do not get their viewers by generating agreement. Far more often the attitudes of hosts, or at least the attitudes that the hosts pretend to have for ratings, are in stark opposition.

In the fall of 2016, and continuing into 2017, there was a controversy in the NFL as some players, notably Colin Kaepernick, knelt during the national anthem. Many thought kneeling during the anthem was wrong because it insulted American armed forces. Others thought the protests were at least permissible: freedom of speech is an American cornerstone. Further, many of those thought the protests were a good thing, since these protests raised awareness of serious social inequality. We can imagine our radio hosts being split along these lines:

Divided Duo	unpatriotic	Free speech is sacred	Raises awareness	Kneeling is bad
Mike	<believe, .9>	<neutral, .5>	<believe, .55>	<believe, .88>
Mad Dog	<disbelieve, .2>	<believe, .9>	<believe, .8>	<disbelieve, .01>
Steward	<believe, .69>	<believe, .9>	<believe, .83>	***<disbelieve, .07>***

Mike strongly believes that kneeling is unpatriotic. He does not think that free speech is relevant to the kneeling, though he does think the kneeling has slightly raised awareness about inequality. Mad Dog, on the other hand, strongly believes that it is not unpatriotic, that free speech protects the kneeling, and that it has raised awareness of an important social issue. I think the Steward Approach gets these aggregations spot on.

I have bolded the bottom righthand cell because the method for arriving at this value does not seem quite right. In aggregating the attitudes, I adopted a column-centered approach, one in accordance with a strict Propositional Invariance rule. But, when I built the values for Mike and Mad Dog's attitudes, I followed a rule so that their attitude for "Kneeling is bad" was a function of their other three attitude pairs. If Mike and Mad Dog did indeed form their beliefs in that way, then our above profile table is incomplete. It should also provide the Steward with the following column:

Addendum	F(col1-col3) = col4[91]
Mike	<believe, 1>
Mad Dog	<believe, 1>
Steward	<believe, 1>

Once the Steward is supplied with this additional information, the Steward Approach yields <disbelieve, .05> under "Kneeling is bad". The initial answer was not far off, but this second one better represents the aggregate attitude of the duo. This shows that toy examples can often be misleading about what the group attitude should be. Strictly speaking, my first aggregation was correct: if the Steward's only evidence was the profile contained in Divided Duo, then Steward should take <disbelieve, .07> attitudes toward "Kneeling is bad". But, if we add the information contained in addendum, the Steward has different rational requirements.

You might be wondering why I was able to aggregate the "believe" and "disbelieve" in the last column of Divided Duo into a Steward attitude of "disbelieve". I did this because of the constraints that credences make on beliefs (and/or vice-versa). This will prove very important in solving an instance of the discursive dilemma in the next section.

§5.4 Tenure Committee

Cases that appear problematic can sometimes be resolved by appealing to the way in which attitudes interact. Generally, I am a pluralist about attitudes. I am open to the possibility that some attitudes can be reduced to others, but my tendency is to treat attitudes as distinct. This does not mean that attitudes are fully independent. For instance, suppose an agent judges that p,

[91] I applied a version of UPCO to combine the first three credences into the fourth: col4 = col1*col2*col3/[col1*col2*col3 + (1-col1)((1-col2)(1-col3)].

withholds belief in p, and assigns p credence 0: this would clearly be irrational. Even if you think that different attitudes are irreducible, you might still think an agent is subject to coherence requirements that hold between these distinct attitudes. While I am generally not swayed by the literature on, say, reducing beliefs to credences, I find such considerations helpful in articulating components of such coherence requirements. In what follows, I will use a few candidate requirements in hopes that they can free the steward from some binds posed by impossibility scenarios.

Importantly, these between attitude coherence requirements allow us to solve the discursive dilemma with the Steward Approach. Consider the following version of the discursive dilemma, Tenure:

Tenure	Research	Lecture	1&2 iff Tenured	Tenured
A	Believe	Believe	Believe	Believe
B	Disbelieve	Believe	Believe	Disbelieve
C	Believe	Disbelieve	Believe	Disbelieve
Steward	?	?	Believe	?

A, B, and C are three members on a committee determining whether some faculty member will make tenure this year. They all agree that the candidate should receive tenure just in case the candidate is a good researcher and a good lecturer. Traditionally, cases like Tenure are taken to show that aggregation is doomed: one cannot maintain Propositional Invariance and maintain consistency. I will show that adding a credence profile to Tenure enables a plausible aggregation.

Let's extend this case to include credences of A, B, and C on the relevant propositions. Let us assume that each member takes R and L to be independent. That is, their respective

credences in the conjunction R&L is equal to the product of their individual credence that R and that L. Again, we can apply the UPCO method of credence aggregation for the Steward:

Tenure*	Research	Lecture	1&2 iff Tenured	Tenured
A	Believe, .8	Believe, .9	Believe, 1	Believe, .72
B	Disbelieve, .1	Believe, .96	Believe, 1	Disbelieve, .096
C	Believe, .7	Disbelieve, .3	Believe, 1	Disbelieve, .21
Steward	Believe, .51	Believe, .99	Believe, 1	Believe, .94

There are a few salient coherence requirements that enable me to fill in the creedal details for this case:

> **Conjunction**: If someone believes that P and judges that Q, then they cannot judge not-(P-and-Q).
>
> **Entailment**: If P logically entails Q, then one must have at least as high a credence in Q as in P.
>
> **Confident Judgments:** If one's credence is over .5 in a proposition, then one cannot judge the proposition to be false. And, if one's credence is under .5 in a proposition, then one cannot judge a proposition to be true.[92]

I want to be careful to qualify the conclusions one might draw from this application of the Steward Approach. I have not shown that one can make an end-run around the impossibility results of the Discursive Dilemma. You can, without too much difficulty, construct credence profiles for A, B, and C that preclude rationally committing the steward to a single attitude profile. Or, you could leave the Steward with only the profile shown in Tenure (not Tenure*). In

[92] This formulation might suggest that credences are ontologically prior to beliefs. This may be right, but I intend no such implicature.

that case, the impoverished nature of the evidence will leave the Steward with multiple rational responses. I will consider how the Steward Approach handles these cases in the next section.

What my approach does show is that, in many situations, we can "tighten" the evidential and coherence requirements in a way that yields the right results. In Tenure*, I take it that we would like to treat the committee as a rational agent, and my approach allows us to recover that verdict.

§5.5 Impoverished Profiles

The previous section showed how fleshing out a profile by adding in credences can resolve otherwise difficult cases. But, what does the Steward Approach say about 'impoverished' profiles where this is not possible? Let's revisit Tenure, where there are no credences associated with the propositions:

Tenure	Research	Lecture	1&2 iff Tenured	Tenured
A	Believe	Believe	Believe	Believe
B	Disbelieve	Believe	Believe	Disbelieve
C	Believe	Disbelieve	Believe	Disbelieve
Steward-N	Neutral	Neutral	Believe	Neutral
Steward-P	Believe	Believe	Believe	Believe
Steward-C	Neutral	Neutral	Believe	Disbelieve

I see three plausible aggregations of this. Steward-N sees that a majority-rule would result in an inconsistency, and responds by taking the neutral attitude to all of the offending propositions. Steward-P favors the premises, believing "Research" and "Lecture", whereas Steward-C favors the conclusion "Tenured".[93] This might appear to be a problem: the Steward Approach does not yield a unique rational attitude.

[93] You could also create variants of Steward-C where the Steward believed one premise or the other, but not both, rather than taking neutral attitudes toward both of them.

The failure to return a unique attitude in this case is a strength, not a weakness. Groups with sparse profile tables, or with strongly conflicting attitudes are less unified. If Tenure (as opposed to Tenure*) fully captured the attitudes relevant to the group's opinions, then the group would have very weak opinions indeed. In a case where a committee was divided like this, it would be a mistake to treat the group as having a stronger opinion than was warranted. Cases where the Steward Approach yields very weak attitudes, or many attitudes, are cases where the group is not unified.

§5.6 Further Work: Higher Order Credences

One worry with aggregation schemes is that individuals who comprise groups can game the system to make their voices heard. For instance, now that you know how UPCO works, you might artificially push your credences farther from .5, since this results in the rule interpreting these pieces of evidence as stronger. I can slightly sidestep this problem by fiat: the Steward gets free and honest insights into the attitudes held by group members. However, we can still build problematic cases:

Super Bowl	Texans Win	Eagles Win
A	Believe, .7	Disbelieve, .3
B	Believe, .7	Disbelieve, .3
C	Neutral, .5	Neutral, .5
Steward	Believe, .84	Believe, .16

Suppose that the above profile depicts the attitudes of three agents watching the Super Bowl between the Texans and the Eagles. As is, the Steward's aggregation seems appropriate. But, we can add more details to the story. Suppose that A and B know little about football, and C is a football connoisseur; were A and B to know about C's attitudes, then they would gladly defer to them. We could thus modify the profile:

Super Bowl	Texans Win	Eagles Win	C is expert
A	Believe, .7	Disbelieve, .3	Believe, .8
B	Believe, .7	Disbelieve, .3	Believe, .8
C	Neutral, .5	Neutral, .5	Believe, .8
Steward	Neutral, .5	Neutral, .5	Believe, .98

In this case, the Steward utilizes the "C is expert" evidence to defer on A and B's behalf in aggregating the attitudes. Essentially, A and B's strong beliefs that C is the expert about football serves as a defeater for their attitudes on the outcome of the Super Bowl, justifying the Steward in matching her attitudes to C's about the winner.

This approach works for discrediting a member's credibility as well as propping one up as an expert. Imagine a church with a single person who always complains about the sound system being too low. If other members of the group identify this person as a 'complainer', and these attitudes show up on the profile table, then the Steward Approach will use this evidence to tone down the value of the complaining person's attitudes.

I think determining the general mechanisms for processing higher order testimony, and higher order evidence, will be extremely difficult to pin down. However, it is clear that we do and should take into account evidence about the reliability of our evidence. Since the Steward Approach just applies these principles to the testimony of a group, the Steward Approach should also perform such corrections on group member attitudes.

§5.7 Concluding Thoughts

This chapter demonstrated that the Steward Approach can aggregate a variety of profiles into plausible group attitudes. The next question is: what do we do with this information? In the next two chapters I will look at some repercussions for being able to identify the attitude of a group. In chapter 6, I argue that aggregating attitudes is nicely suited to addressing the problem of peer

disagreement, and so I apply tools developed in this dissertation to that problem. In chapter 7, I argue that democratically structured attitudes ground norms for group decision making, and this means that arguments made for rational constraints on the Steward are constraints on how groups are permitted to make group decisions. In particular, I use this to argue against winner-takes-all voting systems.

Chapter 6: An Application to Peer Disagreement

A core theme of this dissertation is that we should look for intersections between combining agent's attitudes and combining evidence. One area where this intersection is particularly salient is peer disagreement. In this chapter, I will show how the Bico rule, which is extensionally equivalent to the UPCO rule discussed in chapter 4, can generate a quasi-solution to the problem of peer disagreement.

§6.1: Introduction

You and your coworker disagree about whether it will rain tomorrow. You have no reason to suppose she is any better or worse than you are at predicting the weather, and you have no reason to think she has any better or worse evidence than you have regarding whether it will rain. What should you do? There is a simple answer: you should ask for her evidence and for an explanation about how she came to her conclusion about the weather. Then, you can evaluate her reasoning and process this evidence as usual. But before we even get to that point, you might have wondered how your attitude should change the instant you discover that you and your co-worker disagree. So, we can bracket off practical considerations (who really cares if it will rain tomorrow?) and the obvious follow up questions (why do you think it will rain, Sue?). What is the epistemic import of the disagreement itself? In this chapter, I will offer a formal answer for how your credences should change upon learning that a peer disagrees with you. Here is the position I will defend:

> ***My Central Thesis:*** *In cases like the one given above, discovering that your credence* ***y*** *differs from your peer's credence* ***p****, you should adopt the new credence given by Bico, where Bico(y,p) =* $\boldsymbol{py}/[\boldsymbol{py} + (1-\boldsymbol{p})(1-\boldsymbol{y})]$.

The rule is so named because it is derived by learning a biconditional. Of course, Bico will require some caveats (thus the "restricted" in my title). So, my thesis will be conditional: *If you want such-and-such from your solution*, take Bico.

Here is a roadmap for the chapter. First, I will more carefully lay out the problem of peer disagreement so as to make plain what a successful solution will do (§6.2). Then, I will present the formal solution, Bico, highlighting some of its properties (§6.3). I will offer two stories in favor of Bico being the solution to the problem of peer disagreement (§6.4). First, it has the formal properties that we want. And second, we can rephrase problems of peer disagreement to see why updating on a biconditional would be appropriate. With Bico defended, I express some of its virtues and limitations in sections 5-6. Bico is practical, fits nicely on the continuum with deference, and it is appropriate in many real-world situations.

§6.2 The Problem of Peer Disagreement

The problem of peer disagreement involves you discovering that you disagree with a peer about some proposition. Both "disagreement" and "peer" require some clarification. After clarifying these terms, I'll discuss some solutions on offer in the literature on route to determining what the right solution should do.

A peer, more precisely an epistemic peer, is someone who you have no special reason to think is any better or worse than you at forming a doxastic attitude toward the proposition in

question.[94] It is often easier to identify when someone is *not* an epistemic peer than when they are one. My intoxicated friend is not a peer when determining what portion of the bill I should pay: I have reason to think his level of intoxication makes me much better at remembering who ordered what and performing the simple arithmetic required. Similarly, Stephen Hawking is not my peer when it comes to evaluating the details of String Theory: I have good reason to think he is much better than I am in that domain. In both cases we see an asymmetry between me and the other party. This asymmetry comes to bear on how I react to discovering their views. I quickly discount my inebriated friend, while I also quickly diminish my own views in light of Stephen Hawking's on scientific matters.[95]

It is important to relativize the concept of peer to some proposition or class of propositions in question. While Hawking's scientific views are importantly more valuable than mine, I reasonably do not place much weight on his philosophical assertions. It is rare that two people are epistemic peers simpliciter. For clarity, when we say that two individuals are epistemic peers, we should be able to specify the subject or set of propositions relative to which they are peers.

The previous paragraphs can be distilled into the following definition:

*An **epistemic peer** with respect to proposition p is an agent who we have no antecedent*

---

[94] Kelly (2005) claims he owes "epistemic peer" to Gutting (1982), who defines epistemic peers as those who share your level of "intelligence, perspicacity, honesty, thoroughness, and other relevant epistemic virtues" (p. 83).

[95] There's an additional distinction worth mentioning here, borrowing from Elga's footnote 14 (2007). Here are two ways to phrase the problem: 1) What should you do when you have reason to believe you disagree with a peer? And 2) What should you do when you disagree with a peer? These two questions come apart when your evidence does not accurately represent the situation. You could be misled about whether the other party is in fact a peer, or you could be misled about whether there is a disagreement. For the purposes of this paper, I will focus on cases where there is in fact a peer disagreement and both parties are reasonably aware of this disagreement.

> *special reason to think is i) more or less likely to have true or otherwise epstemically valuable properties in their attitudes towards p or ii) likely to have better or worse evidence than we do towards the proposition.*

In the above example with my friend and the bill, I have special reason to think that (i) applies, since my friend's level of intoxication gives me good reasons to expect his attitudes will be less good than mine. Between Stephen Hawking and I, I have special reasons to think that both (i) and (ii) apply. Hawking will likely have better scientific views than me generally (i), and I also expect his evidence to be much better than mine (ii). But notice that both of these cases are relativized to some proposition, either about the bill or String Theory.

In setting up plausible problems of peer disagreement, it can be a little tricky to ensure the "peer" part holds. We can easily enough avoid the (ii) concern by supposing that the parties have little or no evidence, or by supposing that they share all of their evidence. But here is a worry: what if the disagreement itself constitutes a violation of (i)? Imagine a coworker and I attempt a math problem just given to us. We get different answers. I don't have any reason antecedently to think she's better or worse than I am at math, and let's suppose there's no evidential disparity. Still, I might take the very disagreement in question as evidence that my coworker isn't a peer. Since I must have thought initially that I got the problem right, I am at least somewhat committed to thinking that my disagreeing coworker got it wrong.

There are broadly two ways to deal with this sort of move. First, we could claim that there aren't going to be any peer disagreement cases at all, since the disagreement itself will undermine the peer status of the relationship. This is unsatisfying. Clearly there is some phenomenon that we have in mind when we talk about peer disagreements.[96] Second, we could

[96] Elga (2007) offers another argument against this view. He argues that iterating disagreements will result in the absurd consequence that you think you are much better than your once-thought peer. But, Elga thinks that you should not be able to conclude that you are much better than your "peer" merely by discovering many disagreements. I am not persuaded by Elga's argument, but it has the same

still call these cases peer disagreements, supposing that the peer judgment is made with respect to all other epistemic positions prior to the current disagreement. I'm opting for the latter characterization, so in my definition I include the word "antecedently".[97]

This brings us to clarifying "disagreement". Suppose you believe that it will rain tomorrow and I do not. This clearly constitutes a disagreement. But what if you think the chance of rain tomorrow is .89 while I think it is .90? I will treat both cases as disagreements:

> *One agent **disagrees** with another agent about some proposition p just in case they do not have the same set of propositional attitudes toward that proposition.*

This is a very broad definition of disagreement since it includes all propositional attitudes. On this definition, if I hope the Seahawks win the Super Bowl and you do not, then we disagree. Of course, this is not our target peer disagreement.[98] We are interested in doxastic attitudes, generally either beliefs or credences. In this paper, I will restrict myself to credal disagreements:

> *One agent **credally disagrees** with another agent about some proposition p just in case they do not have the same credence toward that proposition.*[99]

Notice this means that two agents disagree even if they have very similar credences. If you think there is a 91% chance of rain and your coworker thinks there is a 91.5% chance, then you disagree. This disagreement isn't as dramatic as one being certain it will rain while the other is

---

upshot: adding "antecedently", or "prior to" in Elga's case, to our framing of peer disagreement allows us to get at the heart of what makes peer disagreement so puzzling.

[97] Kelly (2005) takes a third route. He uses the disagreement to break the symmetry, but still regards the two disagreeing party as peers: "Two chess players of equal skill do not always play to a draw".

[98] Upon reflection, I am less sure of this claim than when I first wrote it. Our hopes are affected by rational norms insofar as we shouldn't hope for two propositions to be true when they are inconsistent. Moreover, discovering that a reasonable person has different hopes than us might cause us to reevaluate our own. Still, hopes, desires, and other non-doxastic attitudes aren't typically the focus of peer disagreement discussions in epistemology, so I will set them aside for now.

[99] Both agents must assign some credence to the proposition in question. If one or both parties fail to assign any credence to the proposition, it is improper to ascribe a credal disagreement.

certain that it won't, but both are credal disagreements nonetheless.

A successful solution to the problem of peer disagreement must deal with the symmetry of such cases. Part of what compels us to think that we should respond to the other person's views is that, in virtue of being our peer, we deem their epistemic position to be indistinguishable from ours "from the inside." If we were to switch doxastic perspectives with them, we would find ourselves in the same boat. A solution can handle this symmetry concern in one of two ways. First, the solution could offer the same advice to both parties. In this case, the symmetric problem gets met with a symmetric solution. Second, the solution could break the symmetry. If we could point out a respect in which one peer's views were better than the other's views, then this could ground favoring one attitude. With the problem clarified, we can now explore some of the proposed solutions. Solutions mainly fall into two camps: steadfast and conciliatory. Steadfast approaches tell one to stick to one's guns, while conciliatory approaches advise one to compromise to some degree.

The steadfast approach is best captured by Thomas Kelly. Kelly (2005) claims that "disagreement does not provide a good reason for skepticism or to change one's original view." One of his central arguments to this end focuses on symmetry. Kelly accepts that the setup of a peer disagreement between you and me is symmetric: suppose we are equally smart, equally well-informed, and so on. However, the asymmetry comes up when the disagreement is discovered because "of course, this means you have misjudged the probative value of the evidence" by my lights. With the asymmetry thus broken, Kelly opts for endorsing a steadfast solution: the disagreement does not constitute a reason for you to conciliate; standing your ground is permitted, if not required, in cases of peer disagreement.

Kelly's view is *not* that any and all peer disagree-ers should stubbornly refuse to change

their minds. His view, and others like it, have been characterized as "right reasons" views.[100] Suppose that two peers share a common body of evidence, yet they come to disagree about some proposition P: one believes P while the other believes ~P. Further suppose that the evidence objectively favors P. The right reasons view is that the peer who believes P should hold her ground, vindicated by the objective fact that the evidence favors her view. The person who believes ~P should change her belief, but it is not the disagreement that compels her to do so. She should change because that is what the evidence directs her to do in the first place.

The conciliatory approach is nicely summarized by Christensen (2007): "The *fact* of disagreement is old, but bad news…But adjusting our beliefs in the direction of those peers with whom we disagree should be welcomed as a valuable strategy for coping with our known infirmities." Christensen advocates that we "split the difference". Elga advocates an Equal Weight View to a similar end (2007). Splitting the difference, or giving equal weight can be expressed with the following sorts of rules:

> **Belief:** If I believe p and you disbelieve it, then splitting the difference entails that we both withhold belief with respect to p.
>
> **Estimations:** If I estimate that "…. is N…" while you estimate that "…. is M…" where M, N are numbers, then splitting the difference entails that we adjust our estimate to (M+N)/2.[101]
>
> **Credence:** If I assign a proposition credence p and you assign it credence y, then splitting the difference entails that we assign the proposition credence (p+y)/2.[102]

Thus, splitting the difference represents a maximally conciliatory approach, juxtaposed with

---

[100] See Titelbaum 2013 and Elga 2007.

[101] For example, if I estimate that there are 200 jellybeans in the jar and you estimate 500, then splitting the difference should lead us to both estimate 350.

[102] For instance, if I have credence .7 that it will rain tomorrow and you have credence .9, then splitting the difference lands us both at .8.

Kelly's maximally steadfast approach.

Kelly and Christensen give us polar opposite views: hold your ground, or meet in the middle. But there is plenty of conceptual space in between. One could be a "thirder" and move a third of the way towards their interlocutor's credence. The problem that such positions face is a failure to address the symmetry concern. For instance, if I have credence .05 and you have credence .95, the thirder solution sends me to .35 and you to .65. But this doesn't resolve the disagreement! This isn't to say that such proposals are doomed, but they are much less attractive initially.

Both Kelly-like steadfast approaches and Christensen-like conciliatory approaches have initial appeal in different circumstances. There are times when I feel that I should hold my ground, whereas there are others where learning of a disagreement pulls me toward conciliation. A solution to peer disagreement should either honor these feelings, or give an error theory to explain why at least some of them are unfounded. In the next section, I will outline Bico, my proposed solution. Sometimes Bico looks steadfast, other times conciliatory. In this way, Bico might bring in the best of both sorts of approaches.

§6.3: Bico and Its Properties

I came across Bico in some other unrelated work in Bayesian Epistemology on biconditional updating (thus "Bico"). It became apparent to me that its properties had useful application to the problem of peer disagreement. In this section, I will briefly present the formal features of biconditional updating. Then, in §4, I show how these formal properties make Bico a good fit as a solution to peer disagreement.

Sometimes we learn that a biconditional holds between two propositions that were formerly thought to be independent. Our initial credences that both propositions are true is the

product of their respective credences.[103] Then, we discover that one is true just in case the other is. Here's a real-world case that has this feature:

> **Matching Shirts**: Your two friends Abe and Ben have very restricted wardrobes. Both have only white and black shirts. As far as you can tell, each randomly selects a shirt from his closet each morning. Abe wears white 80% of the time, whereas Ben wears white 30% of the time. You are talking to Abe on the phone, and he reveals that he and Ben are wearing the same color shirt.

Let your credence function be the probability function cr(.). You have enough frequency data to yield cr(Abe wears a white shirt) = .8 and cr(Ben wears a white shirt) = .3. But, upon discovering that they are wearing the same shirt, you need to make these two credences match; they are either both wearing white or both wearing black. So, we will update by conditionalization:

> cr(Abe wears white| Abe wears white iff Ben does) =
>
> cr(Abe and Ben wear white)/[cr(both wear white) + cr(neither wear white)] =
>
> .8*.3/[.8*.3 + .2*.7] ≈ .63.

In Matching Shirts, you should be approximately 63% confident that both are wearing white. The rule Bico just applies the above conditionalization to peer disagreement. Formally:

> ***Bico(y, p)*** *=* ***py****/[****py*** *+ (1-****p****)(1-****y****)], provided* $|y-p|<1$,[104] *where y is your credence in some proposition, and p is your peer's credence in that proposition.*

[103] P and Q are independent iff Pr(P&Q) = Pr(P)*Pr(Q). Equivalently, P and Q are independent iff Pr(P|Q) = Pr(P).

[104] This condition prevents the case that y = 0 and p = 1 or vice-versa. If you are certain a proposition is false and certain another is true, bad things happen if you learn that a biconditional holds between them.

I should note that Bico is a particular instance of a general rule of Easwaran et al.'s UPCO[105] (2016). Since our rules are formally the same, many of the benefits of the approach that they tout are also benefits of my proposal, and vice-versa. However, I justify the rule in a completely independent way. It is a happy and interesting finding that we arrived at the same place.

Bico has six notable (for our purposes) properties:

- **Bico(p, q) = Bico(q, p)**. It doesn't matter which order the entries are in.
- **If p, q are > .5, then Bico(p, q) > p, q**. If p and q are both above.5, then the Bico rule will end up even higher.
- **If p, q are < .5, then Bico(p, q) < p, q**. If p and q are both below .5, then the Bico rule will end up even lower.
- **If p > .5 and q < .5, then q < Bico(p, q) < p**. If p and q are on different sides of .5, then the Bico rule will end up between them.
- **Bico(.5, q) = q.** The .5 value collapses to any other value upon using Bico.
- **Bico(1, q) = 1**. The 1 value pulls any other value to itself upon using Bico.
- **Bico(0, q) = 0**. The 0 value pulls any other value to itself upon using Bico.

I won't elaborate on these findings here. Some are surprising (see Hart and Titelbaum (2015)), but the math is straightforward. I will show why these features are desirable for a solution to peer disagreement in the next section.

§6.4: Bico as a Restricted Solution to Peer Disagreement

In this section I will offer my two arguments in favor of Bico as a solution to peer disagreement. The first is that the formalism gives the right answers to proposed cases. The second is that we

[105] Updating on the Credences of Others.

can re-describe cases of peer disagreement as cases of learning biconditionals. If you accept such a re-description, then it follows that biconditional updating is appropriate.

§6.4.1: Curve Fitting

The careful reader may have noticed that Bico has something in common with the steadfast approach. There are situations when Bico recommends that you not change your credence at all, or at least that you change it only slightly. Conversely, there are also times when Bico recommends a conciliatory verdict. However, the fact that Bico is sometimes steadfast and sometimes conciliatory is not a defense: it must give these verdicts in the right cases (§6.4.1) for the right reasons (§6.4.2). Here I will focus on particular cases.

To illustrate that Bico gives intuitive answers to cases of peer disagreement, I will consider three examples and show that Bico gives the right result. Here is the first example:

> ***Election*:** Diane and Robyn are colleagues speculating about whether a Democrat will win the presidency in 2020. Suppose they are peers: equally well-informed about politics with equally good track records of election predictions. Diane says "I'm flummoxed. I have enough evidence, but I just can't see how it favors the Democrats over the Republicans, or vice-versa. My credence is .5." Robyn, on the other hand, is confident that a Democrat will win in 2020; her credence is .85.

Suppose you are a third person who knows nothing of politics, and you want to adjudicate the dispute by setting your own credence. How confident should you be that a Democrat will win in 2020? Diane's testimony is neutral: the evidence that Diane has gathered indicates that it's as likely as not that Democrats will win. On the other hand, Robyn's testimony is relatively strong in favor of a Democrat winning. It seems reasonable to count Robyn's testimony strongly and to

mostly discount Diane's 'flummoxed' testimony: your credence should be near .85. Bico captures this result. Bico(.5, α) = α *no matter what α is*. So, Bico(.5, .85) = .85.

For our next case, consider the tried and true philosophy of science example: given the results of a tuberculosis test, how confident should we be that the patient has turberculosis?

> ***Tuberculosis*:** Suppose our patient takes two tests. Both tests are negative for TB. The first test has a .1 chance of false negatives or positives, while the second test has a .05 chance of false negatives or positives. What is the chance that the patient has TB?[106] Two doctors form their credences on whether the patient has TB, using Bayes' Theorem and their knowledge that the base rate of TB in the patient's demographic is 1%. Doctor Smith uses the first test to yield a credence of .0011.[107] Doctor Jones uses the second test to yield a credence of .00053.[108] Smith did not know about Jones' test and vice-versa. They then learn of each other's credence that the patient has TB.

Again, we see a disagreement between the two results. Should we adopt one of these as our credence? Should we end up in between? Neither of these moves are appropriate: the chances of getting these *two* false negatives is even less than .00053. In fact, the probability that the patient

---

[106] You might object that this isn't a case of peer disagreement, and you would be right. But, we can easily modify the case: suppose doctor A administers one test and forms a credence as to whether the patient has TB, and then doctor B does the same with the second test. Then, they will end up disagreeing about the chance that the patient has TB. What should they do? I think the answer is the same as in the version that I gave, but adding in the disagreeing scientists would only serve to add unnecessary details to the story.

[107] Pr(TB|Neg test) = Pr(Neg test|TB)Pr(TB)/[Pr(Neg test|TB)Pr(TB)+Pr(Neg test|~TB)Pr(~TB) = .1*.01/(.1*.01+.9*.99) = .0011.

[108] .05*.01/(.05*.01+.95*.99) = .00053.

has TB is about .00006.[109] Bico(.00053, .0011) = .0000006. Bico's result does not match the actual probability because Bico uses less information: when Smith and Jones discover the disagreement, they only learn that they disagree and the corresponding credences. The probability, on the other hand, is achieved by conditioning on all of the available evidence in the scenario. Bico correctly identifies that the two low credences are complementary, what Fenton-Glynn et al. call synergy, and should be driven down.

Let us now consider a case where a more conciliatory approach seems appropriate:

> **Conflicting Experts:** a juror hears testimony from both the prosecution's expert and the defense's expert about whether the DNA on the crime scene belongs to the defendant. Both experts seem equally good from the juror's perspective. However, the experts express conflicting conclusions. The defense's expert says the chance that the DNA belongs to the defendant is .05, whereas the prosecutor's expert says that the chance that the DNA belongs to the defendant is .95.

Intuitively, these experts' testimonies should cancel each other out. Bico gives this result, since Bico(.05, .95) = .5. More generally, Bico (.5 – a, .5 + a) = .5 for any $0 < a < .5$. The moral is that Bico perfectly splits the difference between any disagreement centered on .5.

But sometimes we want a conciliation where the disagreement is not centered on .5.

> **Conflicting Experts 2:** This scenario is just like Conflicting experts, except that the defense expert says the chance that the DNA belongs to the defendant is .4 and the prosecutor's expert says the chance is .7.

Again, a conciliatory response seems appropriate. The juror should end up somewhere between .4 and .7 in their confidence that the DNA belongs to the defendant. Bico respects this, since Bico(.4, .7) ≈ .61.

While Bico captures the intuition that we should conciliate in Conflicting Experts 2, you

[109] Pr(TB|test1&test2) = Pr(test1&test2|TB)*Pr(TB)/Pr(test1&test2) = .00006, assuming that the two tests are independent conditional on the patient having TB.

might object that it doesn't get the precisely correct answer. Proponents of a split the difference approach think that the experts should meet precisely in the middle at .55. I have three responses to this objection. First, it's not clear that the essential element of split the difference views is to achieve the average. Fenton-Glynn et al. (2016) diagnose the essence of split the difference views as symmetry (in my terminology). Since Bico offers the same advice to both experts, and would offer the same advice if they switched roles, Bico honors this component of conciliatory approaches. Second, the demand for a precise arithmetic average is unsupported. We are not given a reason to prefer a linear average over, say, a geometric average.[110] Third, even if we grant that adding the credences and dividing by two fits our intuitions the best, splitting the difference will yield counter-intuitive results in the other two examples.

In this section, I have shown that Bico tracks our intuitions about several cases. Election calls for a steadfast response, Conflicting Experts and Conflicting Experts 2 require conciliatory responses, and Tuberculosis features complimentary credences. In each case, Bico yields an appropriate result.

§6.4.2: Disagreements as Learning Biconditionals

The formula for Bico is derived from learning a biconditional holds between two independent propositions. In this section, I will re-cast the story of peer disagreement as one in which you learn a such biconditional. If you buy this retelling of the story, then Bico emerges straightforwardly as the solution to peer disagreement.

Oftentimes, formal epistemologists try to capture credences by thinking about betting behavior, whether dispositional, actual, counterfactual, etc. So, what it means to say I have

[110] The geometric average of .4 and .7 is sqrt(.4*.7) ≈ .53.

credence .8 in a proposition is that I think that it's four times as likely to be true as false, and so I should be willing to bet $4 to win $1. We might actually be more risk averse, or find gambling immoral, and so there is no direct one-to-one relation between credence and betting behavior. Nonetheless, this discussion of betting can help us understand why and when Bico is appropriate in cases of peer disagreement. Call a $1 to $1 bet that a proposition is true an *even bet*. Think of assigning a credence of at least .5 to a proposition as a willingness to accept an even bet on that proposition. Conversely, treat assigning a credence below .5 in a proposition as a willingness to accept an even bet on the negation of the proposition.[111] Let us suppose that rational agents are well calibrated: they will win their even bets at a frequency equal to their credence. For instance, if I assign a credence of .6 to "It will rain tomorrow", I would place an even bet that this proposition is true and will win such bets 60% of the time. If I thought that I would win this bet more than 60% of the time, then I should assign a higher credence.

Let me now tell the story of peer disagreement with the biconditional at the center. You meet Peter at a conference. Peter is a peer in the sense described in §2. There is a somewhat controversial proposition, Q. You both have credences in Q, call Peter's p and yours y, and you have a brief moment to talk with Peter before ducking into another conference talk. Let us make two key assumptions:

- Your and Peter's credences are good approximations of your accuracy at winning even bets that you are willing to take. That is, you and Peter win roughly X% of even bets that you take with credence X/100.
- You and Peter formed your opinions separately, so let us assume that your opinions

[111] This is entailed by the previous sentence if the agent has a credence function that is a probability function.

are independent. This is a strong assumption, and I will consider objections to it in the following section, so withhold judgment on it for now. This assumption means that learning that you would place an even bet on Q doesn't tell you whether Peter would place an even bet on Q. The controversial nature of Q makes this plausible.

You exchange credences. What should you do about the disagreement? There are four cases:

Both y ≥ .5 and p ≥ .5. This means that both you and Peter would place even bets on Q. Let Y be the proposition that you win your bet on Q and P be the proposition that Peter wins his bet on Q. You have learned Y iff P: either you both win your bet or you both lose it.

Y	P	Pr(.)	Pr(.|Y iff P)
T	T	yp	yp/[yp+(1-y)(1-p)]
T	F	y(1-p)	…
F	T	(1-y)p	…
F	F	(1-y)(1-p)	…

So, how likely should you think Q is? Since Q is true just in case Y&P in this scenario, you should end up at the highlighted cell: your credence in Q should be yp/[yp+(1-y)(1-p)], which is what Bico prescribes.

y ≥ .5 and p ≤ .5. You would place an even bet on Q, but Peter would place an even bet on ~Q. In this case you also learn a biconditional: Y iff ~P: either you win the bet and Peter loses, or Peter wins the bet and you lose. Calculating our values for the following state table will be slightly different. Above, it made sense to let the initial credence in P be p, but here the probability of P is (1-p). This is because in the first case Peter bets that Q is true, and we assume that his credences match the chances of a proposition being true. But here Peter bets that Q is

false, so p represents the chance that he will lose the bet.

Y	P	Pr(.)	Pr(.|Y iff ~P)
T	T	y(1-p)	…
T	F	yp	yp/[yp+(1-y)(1-p)]
F	T	(1-y)(1-p)	…
F	F	(1-y)p	…

So, how likely should you think Q is? Since Q is true just in case Y&~P in this scenario, you should end up at the highlighted cell: your credence in Q should be yp/[yp+(1-y)(1-p)].

y ≤ .5 and p ≥ .5. This is just the reverse of the previous case. You still learn Y iff ~P.

Y	P	Pr(.)	Pr(.|Y iff ~P)
T	T	(1-y)p	…
T	F	(1-y)(1-p)	…
F	T	yp	yp/[yp+(1-y)(1-p)]
F	F	y(1-p)	…

Q is true here just in case ~Y&P. Again, Bico gives the right result: yp/[yp+(1-y)(1-p)].

y ≤ .5 and p ≤ .5. Finally, consider if both you and Peter would take an even bet on ~Q. This is like the first case. Since you make the same bet, you learn that Y iff P.

Y	P	Pr(.)	Pr(.|)
T	T	(1-y)(1-p)	…

T	F	(1-y)p	…
F	T	y(1-p)	…
F	F	yp	yp/[yp+(1-y)(1-p)]

Here Q is true just in case we both lose our bets, when ~Y&~P. Thus, we should assign credence yp/[yp+(1-y)(1-p)] to Q.

In each of the four mutually exclusive and exhaustive cases above, conditioning on what we learn about the results of our even bets gives the same result as Bico. Regardless of whether you or your disagree-er assign a credence above or below .5, you should update your credence to yp/[yp+(1-y)(1-p)], if the two assumptions hold. In the next section, I will explore how viable those assumptions are.

§6.5: Restricting the Solution

So far, I have argued for Bico as the solution to peer disagreement, but you might recall that my goal is to argue for Bico as a *restricted* solution to peer disagreement. Here I will highlight those restrictions.

The first restriction is that you might accept a theoretical solution to peer disagreement but nonetheless want to find a more practically satisfying solution. You might think that the problem of peer disagreement is much more easily solved than the prior discussion would suggest. When two peers, who share a common body of evidence, reach different conclusions, at least one of them made a mistake.[112] The onus is then on the one who made the mistake to revise

[112] This is too quick. If there are multiple permissible ur-priors, or multiple conflicting sets of epistemic values, etc., then you might think that two parties could reasonably disagree even if they share the same body of evidence. However, the cases that populate the literature usually seem to admit only one solution (e.g. solving a math problem).

his credence; the other party is in the clear. Call this the Right Reasons (RR) solution.[113] In the event where *both* parties fail to respond rationally to the evidence, then both should revise their credences. To make this solution plain, consider the following case:

> ***Algebra***: Elizabeth and Nicole are peers solving an algebra problem. They both correctly copy the problem from the board and solve for x. Elizabeth gets x = 4 and Nicole gets x = 5 as their solutions. How should they respond?

The RR response is that Elizabeth and Nicole's disagreement is not epistemically significant. While it tells us that at least one of them must have made a mistake, it does not give any new rational requirements. Instead, both Elizabeth and Nicole have the same requirements they had when they initially encountered the problem: correctly solve the algebra problem.

I think there is something importantly right about the RR solution. I accept a RR view as my theoretical approach to peer disagreement, but the RR has two noteworthy shortcomings. First, the real world is messy, and there are many peer disagreements where we don't have identical bodies of evidence. So, even if both parties don't make any mistakes, they might still disagree and want to respond to that disagreement. Second, RR is a particularly unhelpful piece of advice, since agents in cases like Algebra can't tell who—if not both of them!—made the rational mistake in the first place. This leads us to a conditional problem of peer disagreement: given that we don't know if—or which—one of us made a rational mistake, what are we rationally required to do in the face of disagreement?

Titelbaum (2013) responds to these sorts of objections to RR. He notes that RR is a conditional: if you draw the conclusion required by your evidence, then you should not budge in the face of peer disagreement. It is not an objection to RR, Titelbaum argues, that it might be

[113] Following Titelbaum 2013, and Elga 2007 (page 11), who references a passage from Kelly 2005.

difficult for you to identify whether the antecedent is satisfied. I think this is right: the failure to give good advice is not a reason to think that RR is false. However, it still might be a worthy—though different—project to seek out good advice. There are many ways to frame the problem of peer disagreement. One is: what rationally all-things-considered should I do when I discover that I disagree with a peer? Another is: supposing that I might have made a rational mistake, and I will be unable to determine whether I made such a rational mistake, what should I do upon learning that I disagree with a peer? RR gives an answer to the first question: treat your evidence as you were supposed to when you first got it. RR is silent about the second question.

This second version of the problem of peer disagreement may seem strange, but it is similar to a very natural sort of question that we ask in ethics. I might ask: I know it is wrong for me to lie on my taxes. But if I do, should I buy my spouse a gift or donate the money to my church? Reaffirming that I shouldn't lie on my taxes is not to answer this question, but to dodge it. I am asking what morality demands of me if we suspend the requirement that I not lie on my taxes. The problem of peer disagreement parallels this ethical question: I know we should both be rational and end up with the same attitude, but in the case that we don't and end up disagreeing, what should we do? This is why RR falls short: it just reasserts that I shouldn't have made the mistake in the first place.

Bico supplements the two weaknesses of the RR solution. You can apply Bico even when you don't know which of you, if not both, committed a mistake. Further, you can apply Bico in cases where you have separate bodies of evidence and *neither* of you made a mistake. Consider the following case:

> **Sampling Bias**: Nate and Claire are pollsters trying to figure out whether the HR 172 will pass in the Senate. Both gather relatively simple random samples,

> but both get unlucky: Nate gets a biased sample towards "pass" and Claire gets a biased sample towards "fail". They correctly interpret their respective sampling data: Nate assigns "pass" a .7 credence while Claire assigns it .35. They discover they disagree. What should they do?

If we assume that Nate and Claire followed their polling procedure correctly, and did their math right, then RR doesn't offer any advice here. RR only directs one to change their credence in the face of disagreement when a mistake has been made. Bico, on the other hand, allows Nate and Claire to combine their poll-driven credences.[114]

I have shown that Bico nicely fills the gaps in RR, so you should turn to Bico in cases where RR is not feasible. But there are two additional constraints on Bico that were noted in §4.2. Bico is appropriate when 1) you can treat your and your peer's credence as good estimates of the expected frequency that you win even bets that you are willing to take, and 2) you can treat your and your peer's even betting behavior as independent. In what follows I will unpack these constraints and show that they are reasonably met in real world peer disagreements.

As to the first constraint, why should we expect our peers' credences to generate reasonable expected frequencies of successful even bets? There are many ways to express what is meant by having a particular credence. One already discussed is that your credences reveal your betting behavior. Relatedly, it is natural to treat your credences as your confidence that a proposition is true. Most accept Lewis' (1980) Principal Principle, which roughly states that we

[114] Bico(.35, .7) ≈ .56. You might worry that Bico might not reflect the precise statistical answer that would be arrived at by pooling the samples. My response here is the same as with the TB example in §4.1: Nate and Claire don't learn the statistical data that underpins their peer's credence. So, Bico will be operating on less than full evidence. Again, it would be best for Claire and Nate to exchange their data and condition on all of it, but Bico is about updating on the disagreement alone.

should set our credences in P to the objective chance of P whenever we know the chance. My constraint here is different, but it is rooted in the same understanding of relating one's credences to chances. It is a charity principal. If you assign credence *n* to each of 100 propositions, then I should expect approximately 100*n* of them to be true. This constraint is widely met in cases of peer disagreement. If I encounter a person who violates this condition, then I likely don't consider them to be a peer in the first place.

The second condition is less widely met. When and why is it appropriate to treat two peer's bets as independent? To get a grip on this, it will prove helpful to think about when this condition is violated. If you strongly expect your peer to disagree with you, or you strongly expect them to agree with you, then it would be surprising to find out that your attitudes don't match.[115] Here's one such example:

> **Expecting Agreement:** You and your peer have the same academic advisor, and this advisor has strongly influenced your views on epistemology. You both form credences about some epistemological claim E. Then, you learn that you ended up at credences .7 and .71, respectively. What should you do?

In this case, your credence should not move very much. Bico(.7, .71), on the other hand, would send you much higher to a credence of about .85. Bico increases the credence because it treats the two credences as independent corroborators of the likely truth of E. However, here it's reasonable to believe beforehand that you and your peer will have pretty similar credences, and so learning that you have different, but similar, credences should not alter your credence very

[115] See Vavova (2014) for a helpful discussion on how the surprising-ness factors in to how much one should budge in cases of peer disagreement.

much. In this case, something like Christensen's splitting the difference approach seems much more plausible. Consider a case where the disagreement is expected:

> **Expecting Disagreement:** You are the defense attorney's expert from Conflicting Experts (in §4.1). You assign credence .05 that the DNA belongs to the defendant. You see that Joe has been called as the prosecution's expert in this case. You know Joe well, and he is your peer when it comes to DNA evidence. But you also know that on these sorts of cases he will be nearly certain that the DNA is the defendant's. Sure enough, Joe claims his credence is .95.

This is similar to Expecting Agreement. You come into the situation expecting your peer's credence to be approximately what you find it to be, and so it would be a mistake for you to overreact to learning this predictable piece of information. Perhaps you should move a little closer to .5, but not all the way there, as Bico would prescribe.

The last two examples show that if you come into the situation already expecting a certain answer from your peer, you should not change your credence much in the face of disagreement. Bico is plausible and effective on the assumption that you are learning that there is a relationship between you and your peer's even bets. But this leads to a further question: is *that* condition ever satisfied? Shouldn't we always have at least *some expectation* whether our peer's will agree with us? One can mount a reasonable objection against it being met. The condition was not met in Expected Agreement because the two peers had a powerful common influence. Broadly, we all hope that our credences are, directly or indirectly, influenced by the truthmakers of the true proposition in question. For instance, when I assign a high credence to the fact that it will rain tomorrow, I am hoping that the causal factors that will influence it raining tomorrow are rightly related to my credence forming process. Similarly, my peers on weather related issues

should form their opinions by being related to those causal processes as well. So even if it's not as obvious an overlap as sharing an influential advisor, perhaps the independence condition is never met. And this would be bad for Bico.

Fortunately, this objection can be dealt with. There are strong counterexamples to the claim that we should always have an expectation that our bets are dependent with our peers' bets. Consider your views on controversial moral issues, say, abortion. Or consider other difficult questions: what is the answer to a hard math problem? What is the best surgical procedure for a complex medical issue? Should I be a one- or a two-boxer?[116] In these sorts of cases, I may not have any expectation about whether a particular peer's views will align with mine.

The second, and more practical, defense of Bico from the objection is that independence is not a costly assumption even when we expect it's not true. Take calculating the tip after dinner, for example. You will likely have an expectation that you and your peer will agree. You assign credence .8 that a tip of $10 is adequate, and your peer assigns a credence of .9 that $10 is adequate. What is the cost of the independence assumption here? Bico(.8, .9) = .97. This number is too high. You are right in ending up higher than .9, but perhaps it shouldn't be as high as .97. How much should we subtract from .97 to end up at the right spot? The answer is messy, and it depends on exactly how independent you took your credences to be at the outset. I think this is an interesting area for further work, but Bico seems to do the job well enough and satisfies the practical constraint that it is very difficult to quantify the expected level of independence between you and a peer. You might want to adopt a modified account of Bico, call it Bico*, whereby you shave off a little bit from credences that have complimentary effects.

[116] I actually do not think this is a hard question or that it should be controversial. But, that's another matter.

§6.6: Continuity and Peer Agreement

Bico is continuous with intersecting epistemological issues. This virtue means that Bico is flexible. Moreover, it counts in favor of Bico that rules related to Bico seem appropriate in epistemological issues related to peer disagreement; Bico seems to be on the right track. The problem of peer disagreement is closely related to issues of deference.[117] And learning that you disagree with a peer is not all that different from more generally learning what a peer thinks, whether they agree or disagree.

In §2 I noted that Stephen Hawking is not my peer when it comes to String Theory. Indeed, he is an expert, and I would defer to him on such matters. If I have no expertise whatsoever in a subject, then it seems reasonable for me to set my credence to whatever the expert credence is on such matters. Conversely, if I am the expert, then I will not defer at all to those who have no expertise whatsoever. But there are many degrees in between; we often find ourselves interacting with people who are a little bit better than us. It would seem wrong to defer wholesale to them, but perhaps we should defer a little bit. The special case is when we find that we are exactly as much of an expert as the other party. This is the case of peer disagreement. How should we defer, if at all, when we are exactly as good epistemically speaking as the other person? I have already suggested Bico as a restricted solution. Given the continuity between peer disagreement and deference more generally, it would be especially good if Bico could be modified to handle other deference cases.

There is such a modification. Imagine a case where you and someone else disagree about

[117] Elga (2007) starts his discussion on peer disagreement by clarifying issues of deference to experts and gurus.

a proposition. Let's say both of you end up assigning a credence of over .5 to the proposition, and thus you both would place an even bet on Q. Should you defer?

You bet Q is true	He bets Q is true	Pr(.)
T	T	Both win
T	F	You win
F	T	He wins
F	F	Neither win

Bico sets the chance of the "both win" cell to the product of your credences, because it approximates your frequency of winning even bets with your initial credences. But when thinking about deference more broadly, we should treat the more expert party as having more sway. So instead of multiplying the credences, we should first move the relatively non-expert party's credence closer to .5. Here is an example to demonstrate such a modified rule:

> **Hoops:** Alex and Ben are talking basketball. Ben is an avid NBA expert, and Alex knows nothing about the sport. Ben assigns credence .9 that the Warriors will win the NBA Championship in 2017 while Alex assigns credence .6 to this proposition. Alex defers to Ben using a modified version of Bico, instead of computing Bico(.6, .9), he computes Bico(.6-.1, .9) = .9. He simply modifies his credence by bringing it all the way back to .5, signifying that his credence should have no weight compared to Ben's expert credence.

Alex could have subtracted a smaller number than .1 if he had a little more relative expertise. For instance, if Alex is a casual fan, he might compute Bico(.6-.02, .9) = .92. Hopefully this gives the intuitive result: when Alex has no expertise, he and Ben end up collapsing to Ben's

credence. But when Alex has a little bit, then his slight contribution bumps Ben up to .92.

This brings me to a broader point: Bico isn't just for disagreement! The focus on peer disagreement is understandable. It is frustrating when you disagree with a colleague, and you feel some pressure to respond such that you end up agreeing. And there doesn't seem to be a "problem of peer agreement": when you agree, why change your mind at all? I think the right way to frame the issue is to think about learning a peer's doxastic state, whether he or she agrees or disagrees. And it turns out that the assumption that you shouldn't do anything when you learn a peer agrees with you is mistaken. Take a version of the Tuberculosis example from §4.2.

Suppose you and a peer form your credences of whether a patient has TB based off on running independent tests. Suppose it turns out that you have matching credences of .95 that the patient does not have TB. You agree about the issue in the fullest sense. However, you should not stay put. Instead, both of you should be even more confident in your diagnosis.

Broadening the discussion from peer disagreement to learning a peer's attitudes in general strengthens the case for Bico. Bico works just as well for like values, such as Bico(.2, .2) or Bico(.95, .95), as it does for different credences. Steadfast approaches like Kelly's and conciliatory approaches like Christensen's seem to indicate that you should stay put in cases of agreement, but Bico does not make this mistake.

§6.7: Conclusion and Extensions

Disagreeing with a peer puts you in a difficult spot. You experience pressure to conciliate, since you hold your peer's views on a par with your own. But you formed your opinion for a reason, and that stance shouldn't be discounted lightly. Moreover, while it's possible that you or your peer made a mistake, rarely will you be able to quickly and clearly identify that mistake. The epistemically best course of action would be for you and your peer to lay out all your evidence

and go through the attitude-forming process again. But not only is this often practically impossible, it dodges the problem: what should you do the moment you discover the disagreement?

I have argued that Bico is a satisfactory answer in these cases, as long as two additional conditions are met. First, you and your peer's credences are good approximations of your accuracies on even bets. Second, you and your peer's views can be treated as independent. When those conditions are met, Bico is a powerful tool in reacting to disagreements.

However, you might be bothered by the restricted nature of my thesis. To ease some of that discomfort, and in the spirit of continuity discussed in §6.6, I will conclude by gesturing at some extensions of Bico for cases where the above two conditions are not met.

Credences don't match accuracy: It might turn out that you and your peer are less accurate than your credences might indicate. Perhaps you are biased, and this leads to overconfidence. For instance, I might assign a higher credence to the Seahawks winning the Super Bowl than I ought to as a result of my Seahawks fandom. Nonetheless, when I discover that I disagree with my football peer Mike, whose Bay Area roots bias him against Seattle, I might want to use something like Bico to update. Instead of plugging in our credences to Bico, we should remove the bias from the credences before plugging them in. So, if I assign credence .4 to the Seahawks winning, perhaps that should get pulled down to .3. This isn't an exact science; if it were, then we would already know the objective chance in the first place. Instead, it's a practical way to deal with peer disagreement when we have reason to think our credences are not good approximations of our (peer's) accuracy in placing winning bets.

Bets are not independent: When the bets are independent, knowing your own credence doesn't generate an expectation about whether your peers will place a similar bet. But sometimes we expect our peer's views to track more closely with our own. Recall Expecting Agreement from §5: two students share an advisor and as such have similar opinions. In such a case, you already expect your peer to have similar views to your own, so discovering that they do should have relatively little impact on your own credence. There are two ways to alter Bico to get this result. The first is to move your peer's credence closer to .5. The second is to apply Bico and then "back off" the result, moving closer to one's initial credence. The point is to approximate, as best you can, conditionalizing on the combined evidence of you and your peer.

What about non-credences?: Rarely do we actually exchange credences to discover a disagreement. Instead, you might tell a peer in passing that you believe that Trump will win Ohio, but your peer judges otherwise. Since credences aren't conveyed here, Bico does not have its required inputs. In order to extend Bico to these cases, we will need two things. First, we need values for the credences. Second, we need a principle that converts credences to judgments.

These two things require more discussion than I can provide here, but I can give a quick sketch of how this might work. Filling in the credences is not too tricky: you should have access to your own credence—let's say it's .8—and you will have to guess your peer's based on the nonverbal content of their judgment. Perhaps your peer said that Trump would not win, but his voice was shaky and he hemmed and hawed a little before stating that judgment. Then it might be appropriate to attribute a credence below but close to .5 to your peer. Let's call it .45. Now you can perform Bico(.45, .8) to yield .77. So, Bico recommends that in the face of this disagreement that you lower your credence from .8 to .77. What does it say about your judgment? Here we can augment the approach by supposing that you judge that p whenever your

credence is over .6.[118] So, you should still judge that p.

[118] I offer no argument for this. But, it seems something like the following is plausible: if you judge that not-p, your credence is lower than .4. If you judge that p, your credence is greater than .6. If your credence is between .4 and .6, you withhold judgment about p.

Chapter 7: Concluding Remarks

In this concluding chapter, I will tie together some central themes from this dissertation and point toward future developments for the project of attitude aggregation. I began in chapter one by motivating the project. I argued for robust realism, the view that some groups, those picked out in my common-sense view of groups, are agents in a substantial and non-metaphorical way. I presented arguments that groups were ineliminable from adequate theories of our natural language, natural selection, and our attributions of praise and blame to groups. Robust realism requires that groups have attitudes, and that led me to pursue a procedure for identifying those beliefs, credences, desires, etc.

The second chapter laid out the problem of attitude aggregation. As nice as it would be to develop simple mechanisms for aggregating individual attitudes into a group one, the 'easy' answers are inadequate. Moreover, I proved that a plausible set of conditions for the aggregation function cannot possibly be satisfied. But I also gave reason for optimism: the same sorts of impossibility theorems can be used to argue that there is no rational process for an individual to combine (at least some) bodies of evidence. But epistemologists are not, by and large, skeptical about our capacity to respond adequately to even very complex bodies of evidence. So, if we can harness this optimism in individual epistemology, we may also be able to harness this optimism to solve the aggregation problem in social epistemology.

In chapter 3, I proposed the Steward Approach, which capitalizes on the similarity between aggregating attitudes and aggregating evidence. I offered several arguments for the Steward Approach, but the strongest was that we defer to groups of experts and individual

experts alike, but we should end up with the same attitudes by deferring to either the group or the collection of individuals.

Since the impossibility theorems in chapter 2 are valid, a proper account of aggregation cannot keep all of the constraints used in those proofs. In chapter 4, I considered a set of constraints and argued for modifications on them. Then, I followed these modifications with a constructive account for how to aggregate beliefs and credences in a variety of situations. With these on the table, I applied these rules in chapter 5 to show how the Steward Approach handles several two- and three-person example groups. While these were toy cases, they demonstrated that the Steward Approach provides plausible results and is extendable to more complicated examples.

The Steward Approach is powerful not just because it solves a vexing problem, but because it enables us to apply these group attitudes to many other domains where groups are involved. And, as was made clear back in chapter one, groups abound. Chapter six demonstrates that combining individual attitudes into group attitudes gives us a partial solution to peer disagreement. Literally any case where the attitude of a group is a key factor is a case where the Steward Approach may shed some light. To give another example, I will end this dissertation with the sketch of an argument against majority voting in Democracies, particularly when it is a single vote per person, ‘winner-takes-all’ majority voting system.

The democratically structured group attitude grounds norms about decision making for the group, particularly in democratic governments. The strong appeal of democratic governments is that they are more efficient at maintaining the consent of their people insofar as the citizens play direct roles in passing policies and electing those that pass policies. Since the Steward Approach gives us the democratically structured attitude of a group, I can compare the attitudes

reflected by a government with the actual democratically structured attitudes of a group. I can then evaluate whether the structure employed by the government would be consented to by the democratically structured group. Voting structures that lead to a disconnect between the expressed attitude of the group and the democratically structured attitude captured by the Steward Approach are, as far as they violate the consent of the people, unjust.

There are two major problems, at least as analyzed by the Steward Approach, with winner takes all majority (or plurality) voting systems, like those common to broadly western democracies' elections. First, as the discursive dilemma makes clear, such a system is subject to enacting inconsistent polices (or, at least, politicians who will enact inconsistent policies). Second, the ability to only vote for one candidate leads to cases where the enacted policy or elected official does not reflect the preferences of the democratically structured group. Consider the following example:

> Plurality: There are four candidates on the ballot: A, B, C, and D. A, B, and C are all very popular, but they split the vote of 60% of the country, each receiving 20% of the votes. Candidate D is beloved by 40% of the country but hated by the other 60%. He wins the plurality of the vote.

I think it is uncontroversial that Plurality is undesirable. The Steward Approach enables us to say why. If the Steward was to single-handedly elect one of A-D given all and only the testimony of all of the citizens of the country in Plurality, he should choose one of A-C. The Steward would not select D, and certainly not in a landslide (in Plurality, D wins by a margin of 20% over his nearest competitor). Voting systems that make such mismatches between group actions and the democratically structured attitudes of the group are bad.

According to my argument thus far, the required procedure for any group decision is to hire the Steward to make the decisions. That approach would guarantee a perfect match between the expressed attitude of the group and the group's democratically structured attitude. As much as I would enjoy being a Steward Philosopher King, this aggregation method is clearly impractical. There are no perfectly rational beings to make these decisions, nor would such a Steward have access to every group member's attitudes in a timely enough manner to implement group decisions. In practice, voting procedures will allow for these sorts of injustices to occur. The task for building a just voting procedure is to minimize these cases while managing all of the practical constraints that accompany enacting a voting procedure for a given group. There is bound to be a tradeoff between how efficient a voting system is and how accurately that system captures the democratically structured attitude of the group in a range of cases.

The obvious and hopefully fruitful continuation of this project is to consider how we can best design practices for expressing and executing group actions. While the tone of the previous paragraphs focused on plurality voting procedures, groups are varied: there can be a variety of decision procedures for families, sports teams, philosophy departments, and so on. And each of these domains comes with different sets of circumstances that may require different group structures for just and efficient expression of the democratically structured attitude of the group. While the Steward Approach does not give us a pre-packaged answer for the best decision procedure, it does give us the resources to ground evaluative claims about the goodness of certain decision-making procedures of groups.

References

Arrow, K. J. (1963). *Social choice and individual values* (2nd ed.). New York: John Wiley & Sons, Inc.

Austen-Smith, D. and J. S. Banks (1996). 'Information Aggregation, Rationality, and the Condorcet Jury Theorem'. *American Political Science Review* 90(1): 34-45.

Bacharach, M. (2006). *Beyond Individual Choice: Teams and Frames in Game Theory* (edited by N. Gold and R Sugden). Princeton: Princeton University Press.

Barrett, M., Donoghue, M. J., & Sober, E. (1991). Against consensus.*Systematic Zoology,* 486-493.

Barry, B., & Hardin, R. (1982). *Rational man and irrational society?: an introduction and sourcebook*. Beverly Hills: Sage Publications.

Bell, D. E., & Keeney, R. L. (2009). 'Altruistic utility functions for joint decisions'. *The Mathematics of Preference, Choice and Order* (pp. 27-38). Springer Berlin Heidelberg.

Ben-Yashar, R. and S. Nitzan (1997). 'The Optimal Decision Rule for Fixed-Size Committees in Dichotomous Choice Situations: The General Result'. *International Economic Review* 38: 175-86.

Berend, D. and J. Paroush (1998). 'When is Condorcet's Jury Theorem Valid?' *Social Choice and Welfare* 15: 481-8.

Berend, D. and L. Sapir (2007). 'Monotonicity in Condorcet's Jury Theorem with Dependent Voters'. *Social Choice and Welfare* 28(3): 507-28.

Black, D. (1948). 'On the Rationale of Group Decision-Making'. *Journal of Political Economy* 56(1): 23-34.

Boella, G., Damiano, R., & Lesmo, L. (2000). Cooperation and group utility. In*Intelligent Agents*

VI. Agent Theories, Architectures, and Languages (pp. 319-333). Springer Berlin Heidelberg.

Boland, P.J. (1989). 'Majority Systems and the Condorcet Jury Theorem'. *The Statistician* 38: 181-9.

Boston College (2011). *How people assign blame: Cohesive groups hold members less responsible for individual actions*. ScienceDaily. <www.sciencedaily.com/releases/2011/12/111208121400.htm>. Accessed October 1, 2017.

Boyd R, Richerson PJ. 1985. *Culture and the evolutionary process.* Chicago: University of Chicago Press.

Braham, M. and M.J. Holler (2008). 'Distributing Causal Responsibility in Collectivities' *Rational Choice and Normative Philosophy*. T. Boylan and R. Gekker. London: Routledge. 145-63.

Brennan, G. (2001). 'Collective Coherence?' *International Review of Law and Economics*

----- (2003). 'In Praise of Inconsistency?' *Paper presented at Como Conference*.

Briggs, R., Cariani, F., Easwaran, K., & Fitelson, B. (2012). Individual coherence and group coherence. *Essays in Collective Epistemology. Oxford University Press. To appear*, *68*.

Cariani, F. (2011). Judgment aggregation. *Philosophy Compass*, *6*(1), 22-32.

Castro, Clinton and Casey Hart (2017). "The Imprecise impermissivist's dilemma." *Synthese.* 1-18.

Chapman, B. (1998). 'More Easily Done Than Said: Rules, Reason and Rational Social Choice'. *Oxford Journal of Legal Studies* 18(2): 293-329.

Christensen, David. "Disagreement as evidence: the epistemology of controversy." *Philosophy Compass* 4, no. 5.

Citizens United v. Federal Election Commission 558 US_2010. < http://www.supremecourt.gov/opinions/09pdf/08-205.pdf>. Accessed May 7, 2016.

----- (2002). 'Rational Aggregation'. *Politics, Philosophy and Economics* 1(3): 337-54.

Condorcet, M. d. (1785). *Essay sur l'Application de l'Analyse a la Probabilite des Decisions Rendues a la Pluralite des Voix*. Paris.

Conradt, L. and C. List (2009). 'Group Decisions in Humans and Animals: A Survey'. *Philosophical Transactions of the Royal Society B* 364: 719-42.

Conee, Earl, and Richard Feldman (2004). *Evidentialism: Essays in epistemology*. Clarendon Press.

Dietrich, F. (2006). 'Judgment Aggregation: (Im)possibility Theorems'. *Journal of Economic Theory* 126(1): 286-98.

----- (2007). 'A Generalised Model of Judgment Aggregation'. *Social Choice and Welfare* 28(4): 529-65.

----- and C. List (2004). 'A Model of Jury Decisions Where All Jurors Have the Same Evidence'. *Synthese* 142: 175-202.

----- and ----- (2008a). Opinion pooling on general agendas.

----- and ----- (2008b). The aggregation of propositional attitudes: towards a general theory.

----- ----- (2007a). 'Arrow's Theorem in Judgment Aggregation'. *Social Choice and Welfare* 29(1): 19-33.

----- ----- (2007b). 'Strategy-Proof Judgment Aggregation'. *Economics and Philosophy* 23(3): 269-300.

----- ----- (2008a). 'Judgment Aggregation without Full Rationality'. *Social Choice and Welfare* 31: 15-39.

----- ----- (2008b). ‘A Liberal Paradox for Judgment Aggregation’. *Social Choice and Welfare* 31: 59-78.

----- ----- (2010). ‘The Aggregation of Propositional Attitudes: Towards a General Theory’. *Oxford Studies in Epistemology* 3.

Easwaran, K., L. Fenton-Glynn, C. Hitchcock, and J. Velasco (2016). “Updating on the Credences of Others: Disagreement, Agreement, and Synergy.” *Philosopher’s Imprint* 29, no. 1: 146-162.

Eisenberg, E. (1961). Aggregation of utility functions. *Management Science,7*(4), 337-350.

Elga, Adam (2007). “Reflection and Disagreement.” *Nous* 41, no. 3: 478:502.

Erskine, T. (2003) (ed.). *Can Institutions Have Responsibilities? Collective Moral Agency and International Relations*. London: Palgrave.

Feinberg, J. (1968). ‘Collective Responsibility’. *Journal of Philosophy* 65(21): 674-88.

Feldman, R. (2007). Reasonable religious disagreements. In L. M. Antony (Ed.), *Philosophers without Gods: Meditations on Atheism and the Secular Life*.

Feldman, Richard and Andrew Cullison (2014). “Evidentialism”. *The Bloomsbury Companion to Epistemology*.

French, P.A. (1984). *Collective and Corporate Responsibility*. New York: Columbia University Press.

Fuller, S. (1988). *Social epistemology* (Second ed.). Bloomington: Indiana University Press.

Gardenfors, P. (2006). ‘An Arrow-like Theorem for Voting with Logical Consequences’. *Economics and Philosophy* 22(2): 181-90.

Giere, R. (2002). ‘Distributed Cognition in Epistemic Cultures’. *Philosophy of Science* 69: 637-44.

Gilbert, M. (2001). 'Collective Preferences, Obligations, and Rational Choice'. *Economics and Philosophy* 17: 109-120.

Gold, N. and R. Sugden (2007). 'Collective Intentions and Team Agency'. *Journal of Philosophy* 104: 109-37.

Goldman, A. I. (1999). *Knowledge in a social world*. Oxford: Clarendon Press.

Goodman, Nelson (1955). *Fact, Fiction, and Forecast*. Cambridge, Massachusetts: Harvard University Press.

Gorman, W. M. (1959). Separable utility and aggregation. *Econometrica: Journal of the Econometric Society*, 469-481.

Greenfield, Kent (2015). "If Corporations Are People, They Should Act Like It". *The Atlantic*. <https://www.theatlantic.com/politics/archive/2015/02/if-corporations-are-people-they-should-act-like-it/385034/>.

Guilbaud, G.T. (1966). 'Theories of the General Interest, and the Logical Problem of Aggregation'. *Readings in Mathematical Social Science.* P.F. Lazarsfeld and N.W. Henry. Cambridge, MA: MIT Press. 262-307.

Gutting, Gary (1982). *Religious Belief and Religious Skepticism*. (Notre Dame: University of Notre Dame Press, 1982).

Haddock, A. (2010). *Social epistemology*. Oxford: Oxford University Press.

Hardin, R. (1995). *One for All: The Logic of Group Conflict*. Princeton, N.J.: Princeton University Press.

Hausman, D. M. (2007). Group risks, risks to groups, and group engagement in genetics research. *Kennedy Institute of Ethics Journal*, *17*(4), 351-369.

----- . (2008). Protecting groups from genetic research. *Bioethics*, *22*(3), 157-165.

Hayek, F.A. (1945). 'The Use of Knowledge in Society'. *American Economic Review* 35:519-30.

Held, V. (1970). 'Can a Random Collection of Individuals Be Morally Responsible?' *Journal of Philosophy* 67: 471-81.

Helm, B.W. (2008). 'Plural Agents'. *Nous* 42(1): 17-49.

Henrich, Joseph and McElreath, Richard. *The Evolution of Cultural Evolution*. Evolutionary Anthropology: Issues, News, and Reviews 12.3 (2003): 123-135.

---- and R. McElreath (2007). *Dual inheritance theory: the evolution of human cultural capacities and cultural evolution.* Oxford Handbook of Evolutionary Psychology, R. Dunbar and L. Barrett, eds., Ch. 38. Oxford: Oxford University Press.

Hindricks, F. (2008). 'The Status Account of Corporate Agents'. *Concepts of Sharedness – New Essays on Collective Intentionality*. B. Schmid, K. Schulte-Ostermann, and N. Psarros. Frankfurt: Ontos Verlag: 119-44.

----- (2009). 'Corporate Responsibility and Judgment Aggregation'. *Economics and Philosophy* 25: 161-77.

Hylland, A. and R. Zeckhauser (1979). 'The Impossibility of Bayesian Group Decision Making with Separate Aggregation of Beliefs and Values'. *Econometrica* 47: 1321-1336.

Ijiri, Y. (1971). Fundamental queries in aggregation theory. *Journal of the American Statistical Association*, *66*(336), 766-782.

Jackson, F. (1987). 'Group Morality'. *Metaphysics and Morality: Essays in Honour of J.J.C. Smart*. P. Pettit, R. Sylvan, and J. Norman. Oxford: Blackwell: 91-110.

Kelly, E. (2003). 'The Burdens of Collective Liability'. *Ethics and Foreign Intervention*. D. K. Chatterjee and D. E. Scheid. Cambridge: Cambridge University Press. 118-39.

Kelly, Thomas (2005). "The epistemic significance of disagreement." *Oxford Studies in Epistemology* 1: 167-96.

Knight, J. and J. Johnson (1994). 'Aggregation and Deliberation: On the Possibility of Democratic Legitimacy'. *Political Theory* 22(2): 277-96.

Konieczny, S. and R. Pino-Perez (2002). 'Merging Information under Constraints: A Logical Framework'. *Journal of Logic and Computation* 12: 773-808.

Kornhauser, L. A. and L. G. Sager (2004). 'The Many as One: Integrity and Group Choice in Paradoxical Cases'. *Philosophy and Public Affairs* 32: 249-76.

Ladha, K. (1992). 'The Condorcet Jury Theorem, Free Speech and Correlated Votes'. *American Journal of Political Science* 36: 617-34.

Lakoff, George and Mark Johnson (2008). *Metaphors We Live By.* Chicago University Press.

Lehrer, K. and C. Wagner (1981). *Rational Consensus in Science and Society: A Philosophical and Mathematical Study*. Dordrecht (D. Reidel Publishing).

Lewis, David (1980). "A Subjectivist's Guide to Objective Chance." In *Studies in Inductive Logic and Probability*. Vol. 2, ed. R. Carnap and R.C. Jeffrey, 263-93. Berkeley: University of California Press.

List, C., & Pettit, P. (2011). *Group agency: the possibility, design, and status of corporate agents*. Oxford: Oxford University Press.

List, C. (2001). 'Mission Impossible: The Problem of Democratic Aggregation in the Face of Arrow's Theorem'. DPhil thesis in Politics, Oxford University.

----- (2003). 'A Possibility Theorem on Aggregation of Multiple Interconnected Propositions'. *Mathematical Social Sciences* 45(1): 1-13 (with correction in *Math Soc Sci* 52, 2006: 109-110).

----- (2005). ‘Group Knowledge and Group Rationality: A Judgment Aggregation Perspective’. *Episteme: A Journal of Social Epistemology* 2(1): 25-38.

----- (2006). ‘The Discursive Dilemma and Public Reason’. *Ethics* 116(2): 362-402.

----- and P. Pettit (2004). ‘Aggregating sets of Judgments: two impossibility results compared’. *Synthese*. London: LSE Research 140(1-2): 207-235

----- and C. Puppe (2009). ‘Judgment Aggregation: A Survey’. *Oxford Handbook of Rational and Social Choice*. P. Anand, C. Puppe, and P. Pattanaik. Oxford: Oxford University Press. 457-82.

Ludwig, K. (2007). ‘The Argument from Normative Autonomy for Collective Agents’. *Journal of Social Philosophy* 38(3): 410-27.

Lyle, Denniston (2010). *Analysis: The personhood of corporations*. SCOTUSblog. <http://www.scotusblog.com/2010/01/analysis-the-personhood-of-corporations/>. Accessed May 2016.

May, K. O. (1954). Intransitivity, utility, and the aggregation of preference patterns. *Econometrica: Journal of the Econometric Society*, 1-13.

May, L. (1987). *The Morality of Groups: Collective Responsibility, Group-Based Harm, and Corporate Rights*. Notre Dame: University of Notre Dame Press.

May, L. and S. Hoffman (1991) (eds). *Collective Responsibility: Five Decades of Debate in Theoretical and Applied Ethics*. Lanham, Md: Rowman & Littlefield.

Miller, M. and D. Osherson (2009). ‘Methods for Distance-Based Judgment Aggregation’. *Social Choice and Welfare* 32(4) 575-601.

Miller, N. R. (1996). ‘Information, Individual Errors, and Collective Performance: Empirical Evidence on the Condorcet Jury Theorem’. *Group Decision and Negotiation* 5(3): 211-

28.

Mongin, P. (1995). ‘Consistent Bayesian Aggregation’. *Journal of Economic Theory* 66: 313-51.

----- (2008). ‘Factoring out the Impossibility of Logical Aggregation’. *Journal of Economic Theory* 141(1): 100-13.

Morgan, Molly (2001). “Abolish Corporate Personhood.” Speech transcript: < http://moenvironment.org/environment-blog/wp-content/uploads/2017/02/Personhood_Talk.pdf>. Accessed September 05, 2017.

Moriarty, Jeffrey, "Business Ethics", *The Stanford Encyclopedia of Philosophy* (Fall 2017 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2017/entries/ethics-business/>.

Nehring, K. (2003). ‘Arrow’s Theorem as a Corollary’. *Economics Letters* 80(3): 379-82.

----- and C. Puppe (2008). ‘Consistent Judgment Aggregation: the Truth-Functional Case’. *Social Choice and Welfare* 31:41-57.

----- and ----- (2010). ‘Justifiable Group Choice’. *Journal of Economic Theory* 145(2): 583-602.

Ober, J. (2008). *Democracy and Knowledge: Innovation and Learning in Classical Athens*. Princeton: Princeton University Press.

Okasha, S. (2011). Optimal Choice in the Face of Risk: Decision Theory Meets Evolution*. *Philosophy of Science*, *78*(1), 83-104.

Pauly, M. and M. van Hees (2006). ‘Logical Constraints on Judgment Aggregation’. *Journal of Philosophical Logic* 35: 569-85.

Pettigrew, R. and M. Titelbaum (draft). Deference Done Right.

Pettit, P. (2001). ‘Deliberative Democracy and the Discursive Dilemma’. *Philosophical Issues (supp. To Nous)* 11: 268-99.

----- (2002). *Rules, Reasons, and Norms: Selected Essays*. Oxford: Oxford University Press.

----- (2007). 'Rationality, Reasoning and Group Agency'. *Dialectica* 61(4): 495-519.

----- and D. Schweikard (2006). 'Joint Action and Group Agency'. *Philosophy of the Social Sciences* 36: 18-39.

Pigozi, G. (2006). 'Belief Merging and the Discursive Dilemma: An Argument-Based Account to Paradoxes of Judgment Aggregation'. *Synthese* 152: 285-98.

Quine, W. V., 1948, "On What There Is", *The Review of Metaphysics*, 2(1): 21–38. Reprinted in Quine 1953: 1–19.

Richerson PJ, Boyd R. 1976. *A simple dual inheritance model of the conflict between social and biological evolution*. Zygon 11:254-262.

Richter, R. (1985). Rationality, group choice and expected utility. *Synthese*,*63*(2), 203-232.

Rothstein, P. (1990). 'Order-Restricted Preferences and Majority Rule'. *Social Choice and Welfare* 7: 331-42.

Rovane, C. (1997). *The Bounds of Agency: An Essay in Revisionary Metaphysics*. Princeton, NJ: Princeton University Press.

Rubinstein, A. and P. Fishburn (1986). 'Algebraic Aggregation Theory'. *Journal of Economic Theory* 38: 63-77.

Satterthwaite, M. (1975). 'Strategy-proofness and Arrow's Conditions: Existence and Correspondences for Voting Procedures and Social Welfare Functions'. *Journal of Economic Theory* 10(April): 187-217.

Searl, J. (1990). 'Collective Intentions and Actions'. *Intentions in Communication*. P.R. Cohen, J. Morgan, and M. E. Pollack. Cambridge, MA: MIT Press. 401-15.

Shapley, L. and B. Grofman (1981). 'Optimizing Group Judgmental Accuracy'. University of

California, Irvine.

Sheehy, P. (2006). *The Reality of Social Groups*. Aldershot, Ashgate.

Simpson, Edward H. "The interpretation of interaction in contingency tables." *Journal of the Royal Statistical Society. Series B (Methodological)* (1951): 238-241.

Sober, E., & Wilson, D. S. (Eds.). (1999). *Unto others: The evolution and psychology of unselfish behavior* (No. 218). Harvard University Press.

Sober, E (2015). *Ockham's Razors*. Cambridge University Press.

Špirková, J., & Kráľ, P. (2010). Aggregation functions and personal utility functions in general insurance. *Acta Polytechnica Hungarica, 7*(4), 7-23.

Stegenga, J. (2013). An impossibility theorem for amalgamating evidence.*Synthese*, *190*(12), 2391-2411.

Surowiecki, J. (2004). *The Wisdom of Crowds: Why the Many Are Smarter than the Few*. London: Abacus.

Tielbaum, Michael (2013). "Rationality's Fixed Point (Or: In Defense of Right Reason)". *Oxford Studies in Epistemology* 5: 253-94.

Titelbaum, Michael and Matthew Kopec (forthcoming). "When Rational Reasoners Reason Differently". *Essays on Theoretical and Practical Thinking*. Oxford University Press.

Vavova, Katia (2014). "Confidence, Evidence, and Disagreement." *Erkenntnis* 79: 173-183.

West, M. (1984). Bayesian aggregation. *Journal of the Royal Statistical Society. Series A (General),* 600-607.

White, Roger (2005). Epistemic Permissiveness. *Philosophical Perspectives* 19, 445-459.

Wilson, R. (1975). 'On the Theory of Aggregation'. *Journal of Economic Theory* 10: 89-99.