8-18-2017

Thinking Together: Joint Commitment and Social Epistemology

Nathan Sheff

University of Connecticut - Storrs, nathan.sheff@uconn.edu

Thinking Together: Joint Commitment and Social Epistemology

Nathan Sheff, PhD

University of Connecticut, **2017**

Abstract: We often credit groups with reasoning well. Juries can be diligent, or committees negligent. We even credit the groups independently of crediting the members. But what are groups? How can they form beliefs, and how are those beliefs epistemically evaluated? This dissertation aims to answer these questions.

The first chapter addresses a question about shared action. What does it take for some people to do something together, as a group? I defend a view based on Margaret Gilbert's work, where joint commitment is the basic building block of togetherness. Joint commitments are directive-like representations which people direct towards themselves and others, and which they take up from themselves and others. This view avoids the problems typically attributed to Gilbert's account while capturing its virtues.

In the second chapter, I address metaphysical worries about plural subjects, or jointly-committed people. One concern is that groups are entities "over and above" their members (a view which some have endorsed). I develop a broadly reductionist theory about group ontology, where facts about group persistence are replaced with facts about the history of joint commitments. This captures the sense in which groups are nothing but their members, even though there is more to groups than facts about their members.

The third chapter explains group belief, and why it matters. The view developed here largely follows Gilbert's own theory, that group belief is a joint commitment phenomenon. However, the view that groups establish their beliefs by jointly committing to beliefs *via* negotiation has drawn serious criticisms, which I answer here.

Finally, the fourth chapter develops a virtue-theoretic account of epistemic evaluation for group beliefs. Since we often credit groups for their epistemic responsibility, I apply the theories which take epistemic responsibly seriously – namely, virtue responsibilist theories – to group belief. What results is an account of group virtue, where well-formed beliefs flow from joint attitudes of the sort constitutive of virtue. The overall view takes joint commitment to be central at every stage of analysis: central to groups, central to their beliefs, and central to their epistemic agency.

Thinking Together: Joint Commitment and Social Epistemology

Nathan Sheff

B.A., University of Wisconsin-Madison, **2010**

M.A., University of Connecticut-Storrs, **2014**

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

at the

University of Connecticut

2017

Doctor of Philosophy Dissertation

Thinking Together: Joint Commitment and Social Epistemology

Presented by
Nathan Sheff, B.A., M.A.

Major Advisor _____

Michael P. Lynch

Associate Advisor _____

Donald L.M. Baxter

Associate Advisor _____

Austen Clark

University of Connecticut
2017
iii

Acknowledgements:

So many people have made the present work possible that producing a proper thank-you is a daunting task, but I'll do my best.

To the philosophy department at UConn, I give my sincerest gratitude. They accepted me into the graduate program, and into an intellectual community that has felt like home for the last six years. I hope to pass on that friendly spirit wherever I land. To my fellow graduate students, thank you for being great friends and thinkers. I say this over and over, but it bears repeating: I learned as much about how to be a philosopher from talking to all of you as I did from taking classes. In particular, I say thanks to Casey Johnson, for being an excellent role model; to Nathan Kellen, for being a friend, and for fostering community among the graduate students; to Mayank Bora and David Pruitt, for stimulating conversations; to Hanna Gunn, for her generosity with her time and intellectual resources. I want to give a special thanks to Andrew Parisi, from whom I've learned the most. We spent dozens of hours together commuting to Hartford, and his feedback and friendship have been invaluable. Our conversations have shaped the present work to an extent I cannot yet fully appreciate. I feel truly lucky to count him as a friend.

Thanks to my committee: Michael Lynch, Don Baxter, and Austen Clark. Michael was one of the first people I met at UConn; his seminar was the very first I attended; and his mega lecture was the first I TA'ed for. Oh, and he's my main advisor and spent who knows how many hours reading material of varying quality. Don taught me how to grapple with metaphysics, and Austen did likewise for philosophy of mind. It's no accident that I took two seminars with both of them. I hope I've done okay at

following their leads. Every member of my committee gave me vital encouragement, and were the best mentors and exemplars I could have asked for.

I also wish to thank two philosophers who have had a more circuitous influence on this work: Margaret Gilbert and Ruth Millikan. Gilbert's influence, of course, will be plain from the first pages; I would not have written this dissertation if I hadn't read her work. Millikan's influence is less obvious, but readers familiar with her work on semantic content will recognize her footprint on the directive model I develop in the first chapter. From reading their work, I've learned the value of systematic model-building, and seeing how far one gets with it.

Finally, I want to thank my family, who I love dearly. Kathy and Jerry Sheff never second-guessed my decision to major in philosophy, or to pursue a graduate education. They supported me completely and without reservation. I hope I've made them proud. My brother Jake often asked for updates on my dissertation, and here's the last one for now: it's done! I also thank Jourdan White, my best friend and partner in everything. I've learned a lot about joint commitment from our adventures together, and she kept me committed to my work when my confidence was flagging. I should also thank my cat Maggie, the first friend I made in Connecticut, for her companionship.

Introduction

Suppose we credit Watson and Crick with the discovery of the structure of DNA. It is something *they* did, not just something done by Watson and by Crick. We recognize that they deserve credit for their epistemic labor, and this follows a more general practice of attributing knowledge or ignorance to collectives. So in principle, we're comfortable with making claims of the following very broad form: "There is a group, G, that believes a proposition p, and they so believe in the right way." But take inventory here of what we put forward as true by accepting claims like that. We're saying (1) that there are groups, (2) these groups can have mental states, and (3) they can have those mental states in a way that we can appreciate as praiseworthy (or blameworthy, as the case may be). By saying that groups of people can know or otherwise believe well, we assent to several substantive propositions.

In this dissertation, I develop a theory for making sense of those propositions. I largely build my case on the work of Margaret Gilbert, who posits *joint commitment* as the central concept understanding for social activity, and the dissertation proceeds by addressing each substantive proposition in turn. But while Gilbert takes joint commitment to have far-reaching consequences and applications, my use of it will be

more limited in scope. For instance, she takes political obligation, patriotism, love, and more to be joint commitment phenomena. I'm not as confident as Gilbert that joint commitment forms the core of all social phenomena; I do, however, take it to be the heart of things like group belief. My concern is to use it to develop a view of how groups reason well.

The first two chapters address the first substantive thesis, concerning the existence of groups. In the first chapter, the theory of joint commitment, based on Gilbert's work, is developed. Joint commitment explains how groups come to be. The second chapter addresses metaphysical puzzles about groups, and where they fit into our broader theories about what exists. The third chapter is devoted to the second major commitment, and builds a theory of group belief. Finally, the fourth chapter shows how groups can believe in an epistemically good way.

The reader might wonder what is so difficult or puzzling about these apparently substantive commitments. What's wrong with saying that there are only individuals, not groups; that a group believes just what the right individuals in the group believe; that a group belief's epistemic goodness is determined entirely by that of the members' beliefs? This would allow us to talk about nothing other than what we're already familiar with: individual people who believe in ways that sometimes amount to epistemic achievements.

As tempting as these proposals are, they are mistaken. To explain what goes on in groups, we often have to go beyond paltry conceptual resources afforded by notions like individual belief, personal commitments, and the like. As we will see, group life is too complicated to capture otherwise. But as soon as we admit this,

those of a certain philosophical persuasion (present author included) are apt to get heartburn. If we say that the standard conceptual toolkit for understanding individuals somehow fails to capture what happens in groups, we get uncomfortably close to phrases like "irreducibility" or "something over and above," signals that our ontology is expanding to include social substances or group minds. Perhaps we can't get by without a somewhat inflationary ontology, but I take it to be a guiding methodological principle that we should try to do more with less until we can't. Thus, the view developed here is meant to adopt a lean ontology that can still account for the data, and can be used to build a novel approach within social epistemology.

The first chapter puts forth the most effort to that end. There, I take Gilbert's view of joint commitment as a starting point for saying what it is that explains the jointness of joint activity. To see why this needs explanation, consider the following example of joint activity: play. When toddlers play in the same general location, they play near each other, but not necessarily *with* one another, which developmental psychologists call "parallel play." Later on, once typically-developing children reach preschool age, they start playing together, marking a major developmental milestone. There's an intuitive distinction between these two kinds of play, but identifying the deep difference that explains the difference in behavior is hard. We could say that kids play together by intending to play together, but then we have to explain the intention to play *together* in a way that doesn't presuppose the kind of social concept that we're trying to explain. To explain the *togetherness* of interactive play, and of shared action in general, Gilbert invokes what she claims is a familiar notion, but one which has gone without a name for a long time. The concept here is joint

commitment. For people to do something together, they must jointly commit to performing that action, which is not the same as each individual forming the requisite personal commitments.

Joint commitment is needed to explain shared action because shared activity involves norms that could not otherwise be explained. Personal commitments alone could not explain why, for instance, one would be violating an obligation to one's partner if they left their shared activity without a word. One concern about this new concept, however, is that it brings in a new primitive (that is, undefined) concept to explain behavior, marking a sharp break between how we explain individual action and how we explain group action. I will address this worry by appealing to the basic structure of commitments in general, which personal and joint commitments are species of. Commitments, whether personal or joint, are directive-like representations where the producer and consumer are within the same person. In the personal case, there is a single producer and consumer, but in the joint case, there are many. In the joint case, each person issues a directive to themselves and to the others, and each consumes the directives from themselves and each other. This overall structure shows how joint commitments generate the right kind of normativity in a way that goes beyond merely individual-focused concepts but without introducing strange new primitives. Nothing unfamiliar is invoked to explain togetherness.

One might still worry that, even though joint commitment is not a mysterious social primitive, it still generates new entities in the world, what Gilbert calls "plural subjects." These plural subjects must be something other than the members that compose them (if composition is the right notion to use here). After all, if groups are

identical to their members in some way, then distinct groups who share all the same members would turn out to be identical. Given that distinct groups can have the exact same members, groups are not identical with their members. So if groups are not *just* their members, shouldn't they be something *over and above* them as well? Unfortunately, Gilbert has not written on the ontology of plural subjects, or their metaphysical status. But the second chapter uses her ideas to defend the view that groups are *not* something over and above their members. There are good reasons to think that they are, since the metaphysics of groups bears strong similarities with the metaphysics of persons and ordinary objects. Given these parallels with persons and other objects, we might think they have the same metaphysical status as them; groups might be constituted by, but not identical with, their members, for instance. But when you track the history of a joint commitment – who made it, how it was passed on – you have all the facts you need to know for tracking a plural subject's existence through time. No new entity like a group or plural subject is needed. This is a kind of reductionism, in a Parfitian spirit, about groups. The facts about people and the joint commitments that bind them together tell the whole story.

The first two chapters address the first of our substantive theses in claiming that groups can believe well or responsibly. The third chapter addresses the thesis concerning group beliefs. Again, the natural impulse here is to say that group beliefs just consist in what the group members believe. This follows standard practice in, say, reporting survey results, as in, "Americans believe that their president is untrustworthy." More importantly, though, it seems to capture a more basic intuition, that *whatever* groups do *must* be determined by what their members do; group

activity supervenes on member activity. Groups cannot act without their members contributing something, and the same goes for group belief. However, the claim that group belief supervenes on member belief has to be distinguished from the claim that group belief *just is*, say, what most members believe. Indeed, as we will see, there are good reasons to reject the latter view. Some cases show that groups can believe something that no member believes, meaning that member belief is not necessary for group belief, and other cases show that two groups can differ in their beliefs even though they have the same members, showing that member belief is not sufficient for group belief. To explain the supervenience of group belief on member activity, Gilbert develops the joint commitment account of group belief. What it is for a group to believe that p is for them to be jointly committed to acting as much as possible like a single body that believes p (in a sense to be explained). Furthermore, groups determine their beliefs through a negotiation-like process in conversation. Members propose claims that they might jointly accept or reject, and then they negotiate those proposals.

This view encounters some difficulties. First, it just doesn't seem like an account of group *belief*. We don't always choose what to believe (or perhaps only rarely choose what to believe), yet the process leading up to a group's adoption of their view looks entirely voluntary. Whatever group "views" or "positions" are, they aren't beliefs proper. Second, genuine group beliefs as understood by Gilbert cannot easily be distinguished from group lies. A group that jointly commits to acting as if p is true could be lying about p, but Gilbert's view would say that they really do believe p. This makes the joint commitment account of group belief incapable of telling

beliefs and lies apart. The details of the responses will wait until chapter 3, but they can be stated in brief here. First, while belief might be involuntary, it doesn't follow that the processes which lead to the formation of a belief must be involuntary as well. Second, when it comes to distinguishing group beliefs from group lies, the devil is in the details of the cases, and as we will see, the cases do not conclusively establish that Gilbertian group beliefs are indistinguishable from lies. The details of the negotiations which establish the beliefs in question will settle questions about whether a group is being deceptive.

With the character of group belief established, chapter 4 addresses what it takes for groups to believe well, or responsibly. At first blush, this might seem like an abstruse philosophical concern, but common practice takes it for granted that groups can be epistemically responsible or irresponsible. Teams of researchers can be diligent or negligent. Chapter 4 provides a rationale for these practices, for holding groups responsible for epistemic successes or failures. While social epistemology has had a lot to say about social aspects of thinking (like the giving and receiving of testimony), or about the epistemic benefits of aggregating diverse, independent judgments about a given proposition, not much has been said about the character of deliberation, and how it might have unique epistemic features. And not much has been done with the robustly social concepts from the shared action literature, with the exception of Frederick Schmitt's argument for a kind of group reliabilism *via* joint commitment. Chapter 4 takes the overall Gilbertian view developed in the prior chapters as a basis for a novel approach in social epistemology to group reasoning.

The overall approach adopts a virtue epistemology framework of a particularly

responsibilist variety. That is, it takes the responsibilist approach to individual epistemic responsibility and extends it to groups. Groups can have virtues which their members lack, and these group virtues play an ineliminable role in saying how well a group has performed at thinking together. I follow recent virtue theorists who claim that virtues are clusters of attitudes (not propositional attitudes, but attitudes as understood by social psychologists), and argue that groups can jointly have attitudes which constitute group virtues, like intellectual perseverance. At the end of the chapter, the developed approach is applied to the epistemological significance of disagreement. Conciliationists in the disagreement literature argue that peer disagreement can give us reason to lose confidence in the views we disagree about. The responsibilist view here shows how disagreement with one's group can lead to a loss of confidence in one's view, as conciliationists claim. If I take credit for responsibly forming my own view, *and* for how I formed the group view with my fellow group members, then I credit myself for reasoning responsibly to contrary views. Since I cannot rationally hold the contrary views, it is rational to downgrade my confidence in one of them. The responsibilist approach to deliberation thus provides a good foundation for an aspect of a prominent view on the significance of disagreement.

The dissertation thus attempts to give a systematic approach to an under-explored area of social epistemology. Inasmuch as we endorse claims about epistemically responsible groups, we commit ourselves to answering a number of questions about groups, questions about ontology, about philosophy of mind, and about epistemology. In developing the overall view, I've done my best to respect the intuition that

there is something *really* different about group phenomena. Margaret Thatcher's claim that there is *no such thing as society* drastically misses the mark. All the same, groups cannot enjoy some sort of distinct existence from their members and what they do, like social Cartesian egos embedded without explanation in a cluster of bodies. While Margaret Gilbert's work does not do much to navigate these waters, all the right resources are there, and the present Gilbertian dissertation uses them to steer between the Scylla and Charybdis of crude reductionism and naive non-reductionism. Or so I hope.

Contents

Marco Polo describes a bridge, stone by stone.

"But which is the stone that supports the bridge?" Kublai Khan asks.

"The bridge is not supported by one stone or another," Marco answers, "but by the line of the arch that they form."

Kublai Khan remains silent, reflecting. Then he adds: "Why do you speak to me of the stones? It is only the arch that matters to me."

Polo answers: "Without stones there is no arch."

–Italo Calvino, *Invisible Cities*

There have always existed in the breasts of philosophers, including our own breasts, two conflicting tempers... The slogan of the first temper is 'Nothing But...'; that of the other 'Something Else as Well...'

–Gilbert Ryle, "Thinking and Saying"

Chapter 1

Joint Commitment

1.1 Introduction

What is it to do something together, as partners or as a team? There seems to be a difference between many people acting together or with one another, and many peoples' behavior happening in concert. We can bring out this contrast by considering the difference between two sentences: "Jones and I are going to the airport," and "Jones and I are going to the airport together." The latter entails the former, but not vice versa. In this chapter, I consolidate the details of Margaret Gilbert's plural subject theory, developed throughout various books and articles. According to Gilbert, people act together by forming a joint commitment with one another, thereby forming what she calls a plural subject. Clearly, an explanation of her plural subject view must start with the core of her theory: joint commitment. I will suggest that Gilbert's notion of joint commitment places it on a par with the commitments

of individuals, so that individual and joint commitments are species of a broader genus. I build on her work by developing a directive model of joint commitment, where the commitments in question are understood as a kind of implicit imperative-like state. By giving the full account of and extending her theory, we can compare and contrast it with other theories of collective intentionality (those of Searle, Bratman, and Tuomela) to fully assess its virtues and vices.

1.2　What are we trying to explain?

Some activities can be carried out alone or with others. Alice and Beatrice can each go for a walk by themselves, but they can also walk together. The following cases adapted from Gilbert (1990) will draw out the distinction.

> *Walk 1*: Alice goes for a walk on Horsebarn Hill Road. She crosses paths with her friend Beatrice, who explains that she is walking down Horsebarn Hill Road to see the horses. They chat amicably as they walk side by side, and when they reach the horse barn in question, they part ways, with Alice continuing her walk alone.

We can describe Walk 1 by noting that Alice is taking a walk, and Beatrice is taking a walk. Their paths happen to coincide at one point, providing a pleasant diversion, but later they separate. Did they take a walk together? Consider this case instead:

> *Walk 2*: Alice asks Beatrice, "Shall we walk together?" Beatrice replies in the affirmative. They walk down Horsebarn Hill Road, but when they

reach the horse barn, Beatrice says to Alice, "Actually, I'd like to go see the horses — bye!" She then leaves Alice's side, walking towards the horses.

There are several salient differences between Walk 1 and Walk 2. First, Alice and Beatrice began by walking alone in Walk 1, and their paths coincide by accident. By contrast, the coincidence of their paths in Walk 2 is by design, since Alice initiated their walk by inviting Beatrice, who then accepted.

Second, these cases differ in how appropriately Beatrice behaves when she separates from Alice. When Beatrice goes her own way in Walk 1, she is not acting out of line in any way. Alice and Beatrice decided to go on their walks separately and for different reasons. When their plans no longer allow them to chat, they part ways, with no sense of impropriety. This is not the case with Walk 2, however, where Beatrice's behavior reads as bizarre, even rude. Beatrice should not have simply left without excusing herself or asking permission. While Beatrice has not necessarily violated any moral obligations, we would sympathize with certain reactive attitudes from Alice here. Alice would rightly be annoyed with Beatrice on this occasion.

So there are two important differences between Walk 1 and Walk 2: the explanations for how Alice and Beatrice's walks coincide, and the appropriateness of Beatrice's departures. The interesting features we find in Walk 2 but not Walk 1 appear elsewhere. Consider Searle's (Searle, 1990) examples:

Park 1: A number of people are enjoying the outdoors in a park when a sudden storm strikes. Everyone scrambles to get out of the rain and huddle under the various shelters in the park.

4

Park 2: A number of people from the local avant garde dance troupe are in the park. All at once, they scramble, as if to get out of the rain, and huddle under various shelters in the park, in accordance with a new routine they've rehearsed.

In Park 1 and Park 2, we imagine the two groups of people performing the same exact bodily movements, while recognizing a major difference between the two groups. Their internal motivations seem to be different. In the first case, each individual wants to stay dry, and so each individual finds and runs to the nearest shelter. In the second case, each individual acts as they do because of a plan made with the others. In the first case, we can understand what everyone does by understanding the motivations of the individuals, but the second case requires more than that. We cannot understand what the individuals are up to without reference to the group.

What precisely is the difference between Walk 1 and Walk 2, or Park 1 and Park 2? For both pairs, the second case looks more like a case of people doing something together, acting as one body. That is the phenomenon to be explained, what it is for two or more people to do something together, rather than simply doing something in parallel. To put it another way, consider Walk 1 and 2 again. We might naturally say of each case, "Alice and Beatrice went for a walk," but there are two interpretations available for that sentence, both of which correspond to the different cases. These interpretations can be seen through the lens of what Frege calls "telescoping":

If we say 'Schiller and Goethe are poets', we are not really connecting the proper names by 'and', but the sentences 'Schiller is a poet' and 'Goethe is a poet', which have been telescoped into one. It is different

with the sentence 'Siemens and Halske have built the first major telegraph network'. Here we don't have a telescoped form of two sentences, but 'Siemens and Halske' designates a compound object about which a statement is made, and the word 'and' is used to help form the sign for this object. ((Frege, 1979), quoted in (McKay, 2006, p. 11))

In some instances, the word "and" in a noun phrase serves to collapse two conjoined sentences into one. We can see this telescoping effect when describing Walk 1 with "Alice and Beatrice went for a walk." This sentence results from telescoping the longer "Alice went for a walk, and Beatrice went for a walk."

On other occasions, the word "and" does not serve to telescope conjoined sentences, as in Frege's "Siemens and Halske" example, since you cannot reverse the telescoping out to conjoined sentences; that is, "Siemens and Halske have built the first major telegraph network" does not entail "Siemens has built the first major telegraph network, and Halske has built the first major telegraph network." This effect comes out when we describe Walk 2 with "Alice and Beatrice went for a walk," where, as in Frege's example, we cannot telescope out the noun phrase. While "Alice went for a walk, and Beatrice went for a walk" sufficiently describes Walk 1, it does not sufficiently describe Walk 2. In explaining togetherness, we will want to explain the intuition that the description of Walk 2 is inapt when we try telescoping out the noun phrase.

According to Gilbert, these intuitions reflect a familiar, though often unnoticed, cluster of conceptual resources we possess. When we describe the behavior of many people at once, we can either describe the behavior of a collective, or collectively

6

describe the behavior of many individuals. I take the question of collective intentionality to concern the behavior of collectives, and how a collective can exhibit intentionality. Having considered examples of the phenomenon in question above, we can now ask more pointedly: What is the difference between the behavior of many individuals, and the behavior of individuals acting together? What does togetherness consist in? Over the course of this chapter, I will consolidate and extend Margaret Gilbert's plural subject theory, a theory about what explains group behavior.

Her theory does so by saying that groups are plural subjects. People are sometimes the sole subjects of their intentions and mental states; in a group setting, people can be plural subjects of shared states. Just as a belief can be mine because I am the one who has that mental state (i.e. I am that mental state's subject; I'm the one who believes), a belief can be ours because we are the subjects (or rather, the plural subject) of that belief. We form a plural subject when we make a joint commitment.[1] In Walk 1, Alice commits to walking, and so does Beatrice, but neither commits to do anything more than what they commit to as individuals. In Walk 2, they commit to do something together. The commitment is *theirs*. So joint commitment sits at the heart of plural subject theory: some people come together

---

[1] This means that making a joint commitment is sufficient for forming a plural subject. However, won't this rule out joint commitments to *dissolve* a plural subject? For example, a married couple might jointly commit to getting a divorce, bringing their plural subject to an end with a joint commitment. There are two ways to understand this case in ways that do not threaten the sufficiency claim. First, we might say that the commitment to a divorce does dissolve the marriage, but it also suffices for forming the plural subject of the commitment to the divorce. One plural subject replaces another, and with a protracted legal ordeal like divorce, which will involve a lot of cooperation, we might expect this. Second, we might say that the commitment to divorce is best understood, not as a separate commitment, but as a stage of planning within the larger commitment to the marriage. A married couple deciding to divorce is not understood as undertaking a whole new commitment, but as deciding, within the marriage, to release one another from the bonds of commitment. Thanks to Donald Baxter for raising this issue.

as a plural subject when they have made a joint commitment. But what is a joint commitment?

1.3   Commitment as genus

Before explaining joint commitments, it will help to understand commitments more generally. Different philosophers and scientists use the word "commitment" in a variety of ways, and do not always mean the same thing by it. According to Sen, commitments impact our decisions in a manner distinct from considerations of our own welfare. He thinks that commitment breaks "the tight link between individual welfare... and the choice of action," so that one could be committed "to help remove some misery even though one personally does not suffer from it" (2005, p. 7). According to Pettit, Sen's commitments involve "recognizing the goals of another and, regardless of whether or not this answers to independent goals of one's own... letting them impact how one behaves" (Pettit, 2005, p. 15). This gives the impression that commitments are, by nature, distinct from concerns over one's own welfare.

This is congruent with how we describe relationships as committed to refer to their special, reason-conferring powers. As Chang (2013) points out, being in a committed relationship with John gives Alice reasons to take John's interests seriously, but if John is just the barista she chats with in the morning, then no special reasons arise. In a committed relationship with John, Alice has reasons to get him a birthday present, but absent such commitments, such reasons are likewise absent. Indeed, psychologists also use "commitment" to denote an action that is necessarily

other-regarding (Nesse, 2001). On these views, commitments are a kind of social strategy for influencing the behavior of others, since commitments (or at least the credible ones) create expectations in other people. Thus, some commitments have a pointedly social purpose. In the Wisconsin dialect, for example, we have the driving terminology "committing to the turn," which refers to driving behavior where one waits to turn left at a green light by pulling out to the middle of the intersection. This tells other drivers that, once the light turns yellow, and then red, you are taking a left turn, no matter what; one is so far out in the middle of the intersection that one would be in the way of cross traffic if they didn't turn. Thus, drivers commit to the turn by putting themselves in a traffic position where it would be unreasonable not to take the turn. This also reveals your intentions to oncoming traffic; by seeing that you have committed to the turn, they will be less likely to try making the red light. Thus, the commitment to the turn does double duty, by influencing both how you and others drive. (Examples of commitments in this sense include: changing lanes without looking, taking religious vows, caring for a spouse with Alzheimer's, threatening arson, etc. (Nesse, 2001, p. 12).)

Philosophers consider commitments to things beyond persons. One notable example would be the notion of ontological commitment. Suppose the physicists at CERN tell us that their current physical laws say that there is a particle, the Higgs boson, which is responsible for the mass of bodies in space-time. When a theory says that a given entity exists, that entity is said to be in the theory's ontology, the stock of things that exist according to the theory. So, for example, the ontology of physics now includes the Higgs boson particle. The theory is ontologically com-

mitted to the Higgs boson, because the physical laws given by the theory can only be true if the Higgs boson really exists. While this might seem a far cry from the normative landscape of promises and expectation, an ontological commitment has its own reason-giving power, since a theory ontologically committed to particular entities (like Higgs bosons) cannot then coherently say anything that implies such entities do not exist. If ontological commitments are endorsements of some kind, then they force people who have so committed to continue endorsing them. Future propositional attitudes will be constrained by the commitment; having an ontological commitment to bosons means that you cannot suspend judgment on their existence. Failing to remain true to these commitments renders you incoherent; if you don't treat the things in your ontology as existing, you're not coherent, rational, or really committed.[2]

There are also commitments to practices and norms. Lynch, for example, discusses what he calls epistemic commitments, commitments to basic, fundamental epistemic principles: "To commit to such a principle is to trust the method it recommends; to employ or being willing to employ it; and to adopt a policy of relying on the principle as true, of using it as a premise in your deliberations over what to think or do" (Lynch, 2013, p. 344). As with Sen, Pettit, and Chang, commitments here are action-guiding and reason-conferring. They make for standing policies that constrain our moment-to-moment decisions; in the case of epistemic commitments, these decisions concern what or how we think. Some notion of commitment appears to be at work across philosophy and psychology, as Gilbert (2013a, p. 899) attests:

---

[2]Thanks to Andrew Parisi for discussions on this point.

Everyday discourse suggests that commitments are a central feature of human lives. Many refuse invitations citing prior commitments. Some are said to fear commitment. Others are praised for taking their commitments seriously.

According to Gilbert, the unifying theme of these different sorts of commitments is that they all involve facts about one's normative situation, "about what one has reason to do" (ibid). Even ontological commitments confer reasons in this sense; they constrain what one can coherently assert or deny. Having a commitment can entail having reasons to believe or assert certain things. In the broadest sense, according to Gilbert:

A person P is committed to performing act A, given consideration C, if and only if, given C, P has sufficient reason to do A... P has sufficient reason to do A, given C, if and only if, given C, P would be wrong not to do A, all else being equal (ibid).

Note that C does not provide conclusive reasons in all cases, due to the "all else being equal" clause. Suppose I make plans to go to the pub with my friends tonight, but I also made plans to go home and play video games. Naturally, I agree to go to the pub. Taken in isolation, both commitments provide me with sufficient reason to do what they commit me to — my plan to go to the pub gives me sufficient reason to go, and my plan to go home gives me sufficient reason to go home. I cannot fulfill both plans, however, and my commitment to my friends provides a strong countervailing consideration against going home. With that in mind, ultimately I

am committed to the pub, all things considered. It would be wrong of me not to go in this case. What explains this wrongness?

Gilbert insists that the wrongness involved here is not moral wrongness. The normativity in question is not moral normativity, and it need not be. After all, there are more kinds of normativity than moral normativity. For example, there is such a thing as epistemically inappropriate cognitive conduct, as in cases of wishful thinking contrary to one's evidence. There's also imprudent behavior, as when I wear shoes without much traction on an icy day. Neither wishful thinking nor poor choice of footwear will earn me blame from the moral perspective, but some sort of negative response to my behaviors would be warranted, like chastising me for my foolishness. So the normativity there would not be moral normativity, but something else, like epistemic or prudential normativity. This is not to say that epistemic or prudential normativity are totally sui generis sources of reasons (see Cuneo (2007a) for arguments that epistemic and moral norms share many interesting features, e.g. their authority); it is simply to point out that there's more to being right (or wrong) than being morally right (or wrong).

To illustrate this point another way, suppose, in order to make up for my shortage of clean forks at home, I make plans to steal silverware from the pub tonight. When the waiter brings our silverware, I will stuff it all in my pocket. Clearly, this is wrong, and I have strong moral reasons not to do it. However, if I go to the pub and forget to steal the silverware, I will have failed to carry out my plans. Failing to carry out my plans due to forgetfulness is morally fortuitous, in that I did not steal after all. However, it would still be reasonable for me to chastise myself once I get

home without my ill-gotten silverware. I failed by the light of practical reason. Now consider the converse, where I have my plans and act on them. If I did carry out my plans, we might say of my conduct that, while it was wrong, it made sense that I acted as I did, in light of the plans I made. In these cases, we see moral reasons coming apart from practical reasons. I can behave rationally without behaving morally, and vice versa.

In addition to a difference between moral reasons and other kinds of reasons, there is a difference between reasons generated voluntarily and reasons generated non-voluntarily. Some of the reasons I have depend on my will, and some do not. Moral reasons provide the easiest cases of reasons incurred involuntarily. On common-sense morality, I often have reasons, say, to give to charity or not to lie, which apply to me regardless of my will, desires, or goals. It is not unusual to think that moral requirements demand action (or inaction) no matter what we think; for example, there is no decision I must make in order to have a reason to donate to charity. Some reasons, on the other hand, do depend on my will. Obligations generated by promises provide one such example. If I promise you that I will make gnocchi for the party, I now have an obligation that I would not have had without making the promise. So some reasons are generated by an act of will (e.g. by making a promise), while others are not.

For Gilbert, it would be wrong of me not to go to the pub not because it would be immoral, but because it would violate a commitment of the will that I have made (ibid, 900). I have a non-moral reason incurred by an act of my will to go. According to Gilbert, commitments of the will are independent sources of reasons. They provide

13

reasons, but those reasons are conditional upon our willing, and upon what we will (the content of what is willed). The reason I have to go to the pub in the above story depends on my willed commitments, unlike the reasons I have to donate blood, for example, which find their source somewhere else. For Gilbert, once I decide to do something, I have changed the normative facts. Having committed myself to doing something, I now have reason to do it. This illustrates a feature of commitments of the will in general, namely, that they arise from the making of a decision. The decision need not involve a psychologically intensive process, with long planning or careful weighing of reasons (Gilbert, 2006, p. 127). By deciding to do something, the decision-maker is thereby committed to doing it; the decision-maker becomes the subject of the commitment (2000, p. 52).

Decisions, like intentions, are teleological; they have ends. Gilbert stresses, however, that decisions and intentions differ in how they bind their subjects, in that decisions have a trans-temporal reach that intentions lack (Gilbert, 2006, p. 130). Compare my decision to make coq au vin tonight with my intention to get a glass of water. In a sense, my intention gives me sufficient reason to get a glass of water, in the same way that my decision to make coq au vin gives me reason to make the delicious meal. Intentions and decisions alike bind me in providing me with reasons, in giving me reason to fulfill their ends. Intentions cannot bind me in another sense, however, because intentions are, in a sense, flimsy. If I stop intending to get a glass of water, the reason I had which derived from that intention no longer has force, or more strongly, it evaporates. "In contrast," says Gilbert, "once I have decided to do something I cannot just 'stop deciding'. I have to repudiate my decision or I continue

to have sufficient reason to conform to it" (ibid). A subject cannot be released from a commitment unless they have rescinded their decisions where such rescission is "a matter of deliberate repudiation, as when one 'change's one's mind' — deciding not to do what one had previously decided upon" (Gilbert, 2007, p. 263). That is, while intentions change the normative facts, their effects only last as long as the intention does. Decisions have more normative inertia than intentions; they must be unmade, not just forgotten.

To see why, suppose I forget my coq au vin plans, and dinnertime rolls around. Thinking about what to eat, I remember my plans, and realize that I have none of the ingredients. I'd feel rightly annoyed at myself — I'd been hoping for culinary glory, but did not so much as go shopping. This kind of self-chiding is underwritten by the answerability that decisions and their commitments create (Gilbert, 2013b, p. 38). My decision to make coq au vin committed me to making it, and in turn, rationally required me to make it, all else being equal. In failing to do so (and without rescinding the decision), I have not met my obligation. Note here that, even if I had forgotten my plans altogether (say, I absent-mindedly have leftovers instead), I would remain accountable for my failure to cook coq au vin, even if nobody had been witness to my plans but me. This is not the case for the fleeting normativity of intentions.

Thus, for Gilbert, commitments of the will have some general, overarching features. When a decision is made, the decision-maker thereby changes their normative situation by becoming the subject of a commitment. By being the author of the commitment, the subject is in position rescind the commitment through a further

act of the will — merely forgetting the commitment will not do. As the subject of the commitment, they can be criticized for failing to fulfill it, even if only by themselves. So being the subject of a commitment carries some significant normative powers, namely, the power to alter or rescind the commitment in virtue of bringing the commitment about.

1.4   The species of commitment

Who can make decisions, and become the subject of a commitment? Clearly, as the forgoing examples show, individual persons are capable of making commitments of the will. Personal commitments, we can say, come about through personal decisions. My commitment to making coq au vin would clearly count as a personal commitment. I decided that I would make coq au vin, and thereby committed myself to doing it. Decisions deliver commitments as their output. By being committed, I make myself answerable for the presence of coq au vin in my home, and have the standing to encourage my own efforts towards that end, and rebuke myself for any failures. Of course, if I realize later that I would rather opt for a simpler meal, I could decide to do something else, and no longer be beholden to making coq au vin. Personal commitments can be made and unmade unilaterally. Since I'm the only one responsible for the commitment, I'm the only one with the relevant normative powers.

But how does a decision confer these powers on me? One manner of representing decisions suggests an answer, which Gilbert mentions in different places (Gilbert,

2000, pp. 55–56; Gilbert, 2013b, p. 174). From Gilbert's suggestions, I will build the directive model of commitment, where making a commitment is understood as producing a directive representation for oneself. To understand these representations, I will focus on what G.E.M. Anscombe calls their direction of fit (Anscombe, 1957). To bring out the intuitive distinction between the directions of fit for descriptive and directive states, consider her detective example. At the grocery store, a man consults his grocery list to find out what to put in his cart; meanwhile, a detective tasked with shadowing the man records every item that goes into the man's cart. By the end of the shopping trip, the man's list and the detective's notepad will contain all the same information, and yet the list and the notes serve different functions. The grocery list is meant to make the world conform to what is on the list; if the list says "milk," then consulting the list makes the man put milk in his cart. The detective's notes, on the other hand, are meant to conform to the world, so that when milk goes into the cart, "milk" goes into the notes. The list and the notes thus have different directions of fit. Likewise, mental states like beliefs and desires have different directions of fit. Beliefs, like the detective's notes, have a mind-to-world direction of fit; a belief is supposed to make the mind conform to what the world is like. They purport to describe the world. Desires and other conative states, on the other hand, function more like the grocery list. They have a world-to-mind direction of fit. Their job is to bring the world into alignment with what is desired. Moreover, utterances can exhibit direction of fit as well. Like beliefs, indicative sentences try to conform to the world, while imperative sentences function like desires by making the world match their content. Indicative sentences try to describe what is the case;

17

imperatives direct their audiences to action, to make it so.

The direction of fit metaphor is not perfect. Some utterances and mental states do not seem to have any direction of fit (for example, greetings do not seem to have any direction of fit, nor does imagining a talking giraffe), while others seem to be hybrids of directive and descriptive representations (like pushmi-pullyu representations (Millikan, 1995)). Regardless of the precise details, the basic idea behind direction of fit points to an important feature that distinguishes descriptive representations from directive ones. I claim, following the suggestions by Gilbert, that the outputs of decisions are states with (at least) a world-to-mind direction of fit. Hence, the view I develop here is the directive model of commitment. To make a commitment is to bring about a kind of directive representation.

Directive representations (whether utterances or mental states), with their world-to-mind direction of fit, come in many varieties, as desires, intentions, or commands. In developing the directive model of commitment, I will talk about directive representations as if they have the structure of commands. The directive representation that has A as its content, for example, can be represented by the sentence "Do A!" So if making a commitment is to bring about a directive representation, it is natural to represent them as commands. When I decide to make coq au vin tonight, that produces a commitment to make coq au vin, and thus a directive representation of myself making coq au vin. As a result, I have directed something like a command to myself to make coq au vin tonight. This is not to say, however, that commitments are literally a kind of command. Talking about them as if they are commands provides convenient notational shorthand that carries the world-to-mind flavor of directive

18

representations.

Commands (taken as a paradigm example of directive representations) always have a two-part structure. There is the producer of the command, which issues a command with a particular content (e.g. "Do A!"), and there is the consumer of the command, the recipient of the command, which carries out whatever the producer commanded.[3] In Anscombe's case, the shopping list is a directive representation, where the producer is the list-maker, and the consumer is the shopper himself, using the list to stock his shopping cart. Paradigm commands have an asymmetrical structure. The producer is supposed to provide input with regards to what the consumer is to do; they set the agenda for the consumer, not vice versa. Meanwhile, the consumer, the addressee of the directive, is supposed to do whatever is on the agenda. While producers can change the agenda (e.g. "Do A! On second thought, do B instead!"), consumers cannot; they are answerable to carrying out what is on the agenda. If they fail to carry out the command, the producer can hold the consumer accountable. Thus, when a foreman tells the worker, "Fetch a slab," he can expect the worker to return with a slab. He can also cancel or alter the command, if he so chooses. If he does not, however, the worker is on the hook for bringing the slab to the foreman. This is generally how commands work, with an asymmetrical relationship between producer and consumer.[4]

---

[3]This way of framing the issue is substantially influenced by Ruth Millikan's work, especially (Millikan, 1984).

[4]Another way to make this point is that commands provide a good model of the directives in question. But as with any model, certain aspects of commands have to be smoothed away for the model to do its job as a simplifying interpretative device. After all, commands have some features which are unique to commands, which are not possessed by directives in general. For example, commands are normally authoritative for a recipient only when the command-giver is already in a position of authority over the recipient. Directives in general do not require an already-established

On the directive model of commitment, we can model my decision to make coq au vin as producing a command to myself to make coq au vin. The paradigm command considered above involves one person, a foreman, issuing a command to another person, the worker. In principle, however, producers and consumers can be located in one and the same organism. Such is the case on the directive model of commitment, where a personal commitment is understood as a person producing and consuming a directive representation, like a command. When it comes to my commitment to coq au vin, insofar as I am the producer of the command, I can hold the consumer accountable (in this case, myself), and insofar as I am the consumer of the command, I have sufficient reason to do what is commanded, lest I be faulted for not doing so. Thus, we understand how commitments generate normative powers by understanding them in terms of commands, where the command is produced and consumed by one and the same person. It explains how I can be held accountable (because I am the consumer of the command), and how the decision gives me a reason to do what I decided to do.[5]

It can also explain how I can unilaterally rescind my decision, and also why I must rescind a decision prior to doing anything other than what I decided upon. That is, the directive model explains why I am allowed to change my mind about my dinner plans, provided that I decide to rescind those plans; it also explains

---

authority structure for them to fulfill their function. A to-do list written for oneself is a directive, and doesn't presuppose that the author of the list has a position of authority over the recipient of the list. Even so, it can be useful to think of the to-do list as a self-directed command, in order to understand how it functions as a directive representation.

[5]It is worth noting again that this is not to imply that commitments *are* commands. Rather, commitments are a kind of directive, and commands are good models for directives, especially with regards to the powers of the producing and consuming devices.

why the decision to rescind is required, lest I rebuke myself for falling short of my gustatory destiny. Insofar as I am the producer and originator of a command to myself, I can alter or rescind the command to myself to make coq au vin. Such are the powers of producers of commands. Yet insofar as I am the recipient of a command, I have to act on the commands given to me. If I am out of line with whatever commands I've been given (e.g. to make coq au vin), then I can be blamed for it. So if I'm to do anything other than what I was originally commanded to do, the original command has to be rescinded or altered first. The directive model tells the right story about how commitments give us reasons to act, why we must rescind or alter them if we want to do otherwise, and why we can unilaterally make and unmake our commitments. We can make sense of personal commitments as directive representations, like commands, issued and obeyed by one and the same person.

This puts us in position to understand the other species of commitment: joint commitment. As Gilbert puts it, jointly committed individuals constitute the plural subject of the same commitment. While joint and personal commitments share homologous structures, the collectivity complicates matters in joint commitment. To see how, it is worth remembering how we started out, and how we got here. Gilbert's pair of cases, Walk 1 and Walk 2, brings out the need for an explanation of the togetherness we sense in Walk 2, but not Walk 1 (or for that matter, in Park 2, but not Park 1). With joint commitment, we can say what's needed to give an account.

So far, we have a working understanding of personal commitments: they are a normatively thick conative state, more demanding than intentions. Walk 1 shows

us two people with their own personal commitments going their own ways; nothing more than personal commitment is needed to understand Walk 1. The temptation now might be to analyze Walk 2 as a complex of Alice and Beatrice's personal commitments, so that, when Beatrice accepts Alice's invitation, each one thereby forms a personal commitment. Alice commits herself to going on a walk alongside Beatrice, and Beatrice commits herself to going on a walk alongside Alice. Their personal commitments enjoin each one to walk along Horsebarn Hill, for a time, until Beatrice breaks away. Remember, though, that personal commitment only binds the person who forms the commitment, and it remains binding until the person rescinds or alters their commitment.

This is why personal commitments alone, even complicated interlocking ones, cannot explain Walk 2. When Beatrice walks away in Walk 2, something has gone wrong. Her behavior is clueless at best, and we judge the case this way because Alice and Beatrice were walking together, unlike their activity in Walk 1. However, if the commitments of Walk 2 were simply personal commitments, then we have no way to explain what Beatrice has done wrong, for as long as Beatrice replaces her commitment to walk alongside Alice with a commitment to see the horses, she acts rationally. Of course, it would be irrational for her to suddenly ignore her commitment to taking a walk and veer off to the horses, but since she instead changes her mind, Beatrice acts rationally in walking away by rescinding her initial commitment and replacing it with a decision to go to the horses. On the directive model of personal commitment, she withdraws the self-directed to command to go on a walk, meaning her reasons to go on a walk vanish. As the author of that command, she

can rescind it at will; as a result, Alice has no grounds to criticize her. Personal commitment cannot give us the resources to understand how Beatrice might have done something wrong in Walk 2, but if not personal commitment, what?

When Alice and Beatrice commit to going on a walk together, they do not individually make personal commitments; rather, they jointly commit to going on a walk. The commitment is theirs. On the directive model of commitment, they issue a command to themselves, but in order for the commitment to be anything other than a personal commitment, it cannot simply be the case that Alice issues a command to herself, and Beatrice does likewise. Instead, they (Alice and Beatrice) must issue a command to themselves (again, Alice and Beatrice).

If that's the case, however, a charge of circularity arises. It seems that, in order to explain how Alice and Beatrice could do something together, we have to say that they make a commitment together. This is a problem if togetherness is what we're trying to explain, so the analysis needs care.

The circularity charge can be avoided. On the directive model of commitment, a person produces a command (or rather, a directive) towards him or herself, and is then the consumer of the command. When Alice decides to go for a walk, she issues the command and receives it. Insofar as she produces the command, she can hold the consumer accountable; since she herself is the consumer, she is accountable to herself. Recall that directives of this kind have two parts: a producer of the command, and a consumer. In personal commitments, the producer and consumer are one and the same person. Joint commitments have a homologous structure, where the committed parties are likewise the producers and consumers of the command.

23

They are not, however, joint producers and joint consumers of the command (that way lies circularity).

In order to understand their structure, consider the following example according to the directive model I'm proposing. Alice extends an invitation to Beatrice, which she accepts. At this point, two things happen. Alice issues a command both to herself and Beatrice to go on a walk, making Beatrice and herself consumers of Alice's command. At the same time, Beatrice produces a command to go on a walk directed both at herself and Alice, making Alice and herself consumers of Beatrice's command. On this picture, Alice and Beatrice do not jointly produce and jointly consume a command in order to become jointly committed. Rather, each one produces a command directed at an audience of consumers that includes themselves and others.

This means that successfully forming a joint commitment requires all parties to direct their commands to others (and themselves), and accept commands from others (and themselves). This in turn requires that all parties are aware of the joint commitment taking form. To see why, consider the following case.

> *Walk 3:* Alice asks Beatrice, "Shall we walk to Horsebarn Hill?" Beatrice, daydreaming about the horses and only half-listening to Alice, replies in the affirmative. They walk down Horsebarn Hill Road, but when they reach the horse barn, Beatrice says to Alice, "Time to see the horses — bye!" She then leaves Alice's side, walking towards the horses.

Something misfired here. When Beatrice replies in the affirmative to Alice, it's not clear that she understands what Alice is asking of her. Maybe she thinks the situation is like Walk 1, so that she is only expected to make conversation while their walking

routes overlap. This isn't the case, however; Alice expects them to walk together, not just fortuitously side by side. In Walk 3, one party expects a joint commitment, while the other does not. Successful joint commitment requires common knowledge among all committing parties that they are undertaking a joint commitment (Gilbert, 2013b, p. 46), and that all parties are ready for being committed.

The directive model understands the misfire in Walk 3 in the following way. Alice invites Beatrice on a walk, expressing her readiness to walk together. Beatrice does not recognize Alice's readiness, but instead takes her invitation to be a prompt to go for a walk. Thus, Beatrice forms a personal commitment to go see the horses (and share conversation with Alice for a moment), which Alice takes to be mutual readiness to walk together. Alice issues a command to herself and Beatrice, and (falsely) takes Beatrice to be accepting it. In taking Beatrice's "yes" as an expression of readiness, she treats it like a command-like directive representation, Beatrice's end of the joint commitment. Beatrice intends nothing of the sort, however; her affirmative response is simply a report of her personal commitment to see the horses. No joint commitment can form here, for the simple reason that the readiness of all the parties involved has not been established as common knowledge.

On the directive model of joint commitment, each party, having established their readiness to committing, issues a command to themselves and each other, which they then take up (from themselves and each other). This preserves a homologous self-directed command structure with personal commitment without circularity. It also preserves the normative features of personal commitment, and explains how they manifest in unique ways in joint commitment. In making a personal commitment

to cook tonight, I give myself sufficient reason to cook; I cannot rationally get out of cooking, unless I decide to do something else. This is because, as the consumer of the command, I am bound to do what is commanded, while as the producer of the command, I am free to rescind it. As both the producer and consumer of the command, it makes sense why personal commitments allow for unilateral change in plan, but joint commitments seem built to deny all parties unilateral powers. The following case illustrates this.

> *Walk 4:* Alice asks Beatrice, "Shall we walk together?" Beatrice replies in the affirmative. They walk down Horsebarn Hill Road, but when they reach the horse barn, Beatrice says to Alice, "Would it be alright if I go see the horses?" Alice nods, and they part ways.

Here we see the mutual obligations and permissions formed in the context of joint commitment. I take it that the normative situation in Walk 4 is different than Walk 2 or Walk 3, since Beatrice has asked permission to end their walk together, which Alice then grants. Beatrice wanted to do something else, but, having made a commitment with Alice, recognizes that she cannot simply walk away. She needs to make sure leaving would be okay with Alice, because the decision to walk away is not hers to make. She did make the commitment to go on the walk all by herself, after all; Alice co-authored the commitment. Unilateral revisions to one's personal commitments are permissible because the party bound by the command is also the party that made the command. In Walk 4, two parties make and take up the command, and because of this, both parties must agree to any revision. Everyone must be on board with the changes. In a sense, then, the basic structure that underwrites unilateral revision in

personal commitments has a homologous counterpart present in joint commitments. In any commitment, the only person who can permissibly change the plans is the person who made the plans in the first place, and in personal commitments, the producer and consumer of the command happens to be a single person, whereas in joint commitment, the producers and consumers are many.

To see why, consider how this might go on the directive model of commitment. When Beatrice accepts Alice's invitation, the command is produced to themselves and each other. That is, Alice's invitation and Beatrice's subsequent acceptance of it token self- and other-directed directives; these tokened self- and other-directed directives are like a command they issue to themselves and each other. If Beatrice wants to change that command, she cannot simply rescind it herself, because she was not the sole author of the command. Alice also issued it, so Alice would need to sign off on any changes. By changing what she has commanded to herself, she has done nothing more than release herself from her own self-directed command. That does nothing to change what Alice commanded. To illustrate the more general principle here, consider Mona, a child who wants some ice cream, but cannot have it without permission. She asks her mother if she can have some, and mom says no. Undeterred, Mona goes into the next room as asks her dad, who says yes. I take it that, as far as most families are concerned, Mona has not gotten permission to have ice cream, because only one of her parents agreed. For Mona to get permission, both of her parents need to be on board, because they set the policy that she can only have ice cream when they give her permission. The same considerations hold when someone has set rules with someone else that they themselves must follow. Beatrice,

along with Alice, gives herself rules. She cannot simply release herself from those rules, however, since she did not make the rules herself. She needs to ask permission from the other person responsible for said rules: Alice.

The mutual obligations and permissions that form out of joint commitments make it seem like a *sui generis* kind of behavior, but that appearance is an artifact of the comparison of joint commitments with personal commitments. While joint commitments are not reducible to personal commitments, the differences between joint and personal commitment (e.g. that personal commitment allows unilateral changes where joint commitment does not) appear natural and harmless when we view joint and personal commitment as two species of commitment.

1.5   Assessing the theory

Having explained and developed Gilbert's plural subject theory, we can now assess its viability. One test for a theory's success is to see how it fares at accounting for our pre-theoretic intuitions concerning the theory's subject matter; a theory's plausibility suffers to the extent that it cannot capture pre-theoretic intuitions. Two intuitions about collective intentionality seem particularly difficult to satisfy at once. On the one hand, collective intentionality seems to be a distinctive phenomenon, very different from individual intentionality. Call this the "irreducibility thesis." On the other hand, collective intentionality cannot be *too* different in kind from individual intentionality; they can't be completely untethered. Following Michael Bratman, we can call this the "continuity thesis." Each thesis seems reasonable on its own, but

they seem to be at odds with one another. If we accept the irreducibility thesis, we deny that collective phenomena are built out of individual phenomena, but that would seem to require some new primitives or group minds. If we want to reject *those* and instead accept the continuity thesis, then it looks like we must give up on the distinctiveness of the collective, and thus, the irreducibility thesis.

These are powerful intuitions; they are also strikingly at odds with one another. However, I will argue that when we adopt the directive model of commitment and understand commitment as a self-directed directive (modelled as a command), Gilbert's theory can satisfy these competing theses. Joint commitment can accommodate the intuition that collective intentionality is unique and cannot be reduced to the intentions of individuals, while also admitting that it is wholly explicable in more basic terms. First, however, we will consider in some detail the opposing positions that lead to the problem. I will take Michael Bratman to champion the more reductionist camp, and John Searle the nonreductionist camp.

Bratman motivates his view by appealing to what he calls the "continuity thesis" (Bratman, 2014). According to the continuity thesis, there is no discontinuity in nature between the intentions of collectives and the intentions of individuals. In Bratman's words (ibid, p. 8):

> My conjecture is that the conceptual, metaphysical, and normative structures central to such modest sociality are – in a sense I aim to explain – continuous with structures of individual planning agency. This is the continuity thesis. As we might try saying: once God created individual planning agents and placed those agents in a world in which they have

relevant knowledge of each other's minds, nothing fundamentally new — conceptually, metaphysically, or normatively – needs to be added for there to be modest sociality. ... The deep structure of at least a central form of modest sociality is constituted by elements that are continuous with those at work in the planning theory of our individual agency.

Bratman puts forward a three-pronged continuity thesis. Individual agency (understood according to his planning theory of intention) has conceptual, metaphysical, and normative structures. Collective intentionality does, too, and on the continuity thesis, such structures are continuous with those we find in individual planning agency. Stating the continuity thesis negatively: collective intentionality involves no new structures (conceptual, metaphysical, or normative) in addition to those we see in individual intentionality. To clarify the sense of continuity here, consider the continuity between the conceptual structures of individual and collective intentionality. Conceptual continuity can be viewed as a relation between theories. Collective and individual intentionality are conceptually continuous because a theory of the former is reducible to a theory of the latter.

There are a number of questions we might ask about the continuity thesis.

First, what is the argument for it? Second, even if some structures are continuous from the individual to the collective case, does that mean that all three kinds of structure will be? Could it turn out that individual and collective intentionality have the same underlying metaphysical nature, and yet our concepts for them are discontinuous? Bratman's answer to the first question informs his answer to the second. As he says in the quoted passage above, the continuity thesis is a conjec-

ture. It embodies a methodological principle for building as conservative a theory as possible without introducing new elements. If we really need new primitive elements to explain collective intentionality, then an argument must be given for that conclusion. Until then, there is a presumption in favor of plausible and parsimonious theories. As a result, there is a presumption in favor of theories that invoke continuity across all three structures. So Bratman takes the continuity thesis as a starting point for the construction of his theory. It is a theoretical desideratum: a theory is more plausible to the extent that it satisfies the continuity thesis. Even though we might discuss three separate continuity theses (one for each structure), I will discuss the all-inclusive thesis as Bratman does, especially the aspect concerning conceptual continuity.

Even with a focus just on conceptual continuity, we can draw out further distinctions between versions of the continuity thesis. On the continuity thesis, the difference in how we understand Walk 1 and Walk 2 is not a difference in conceptual resources. There is no deep difference between how we think of the intentions of Alice and Beatrice in the first case, and the intentions in the second case. The explanations of individual and collective intentions share the same underlying conceptual resources. Call this the moderate continuity thesis: collective and individual intentionality are continuous in that our resources for explaining collective intentionality are the same resources as those for explaining individual intentionality. In order to understand collective intentionality, we need not expand the conceptual repertoire we use for understanding individual intentionality. Thus, collective and individual intentionality are continuous in that they can be explained using the same conceptual

resources.

As we see in the above passage, Bratman holds the strong continuity thesis, where individual intentions are conceptually more fundamental than, and form the conceptual building blocks of, collective intentions. Compare this with relations between concepts in other theories. For example, in Euclidean geometry, points are conceptually more fundamental than lines, because lines are defined in terms of points (see Euclid's *Elements*, Book I, definitions 1-4). Some logical concepts are also more fundamental than others. On some views of logic, being a good argument is defined in terms of the absence of a counterexample, so absence of counterexample is more conceptually more basic than being a good argument. On the strong continuity thesis, the concepts involved in explaining collective intentionality are defined in terms of more basic concepts of individual intentionality. As a result, explanations of Walk 1 and Walk 2 invoke nothing other than individual intentions; the individual intentions of Walk 2 just turn out to have more complicated content than those in Walk 1. The strong continuity thesis entails the moderate continuity thesis, because it claims that the resources one needs to understand collective intentionality are also resources one needs to understand individual intentionality. However, it goes a step further in saying that, in fact, those resources just are the resources needed to understand individual intentionality. In order to understand what it is for Alice and Beatrice to intend to do something together, one first needs to understand what it is for Alice or Beatrice to intend *simpliciter*. Once you have that (along with some competency in understanding the minds of others), you have everything else. As Bratman might put it, God can completely describe the social world, at least

32

at the level of activities like Walk 2, with the conceptual resources of individual intentionality. Once God has made a world stocked with interacting people with their personal intentions, nothing else is needed. Thus, Bratman's thoroughgoing commitment to the strong continuity thesis makes for a reductionist theory.[6]

By contrast, Searle (1990) provides a nonreductionist theory of collective intentionality. Like Bratman and Gilbert, Searle starts from the intuition that Walk 1 and Walk 2 involve very different kinds of behavior; the collective behavior in the latter case is very different than what we find in the former. Searle's Park 1 and Park 2 cases are also meant to show that, even more strongly, we can have all the same individuals making the exact same movements, and still have a difference in their behavior regardless. Interestingly enough, Searle proposes two further theoretical constraints that Bratman would agree with. First, a theory of collective intentionality must be consistent with the fact that society only consists of individuals and individual minds; the theory must be consistent the denial of group minds (ibid, 406). Second, any account of collective intentionality must be consistent with the

---

[6]One concern is that acceptance of the strong continuity thesis makes reductionism too easy, because any continuity thesis is indistinguishable from reductionism. One can, however, maintain a nonreductionist theory that is compatible with the moderate continuity thesis, as we will see shortly. To show how two things can be continuous with one another, without one being reducible to the other, consider different kinds of conjunction. Sentential conjunction is familiar, e.g. "John went to the store, and Mary went to class." So are noun phrase conjunction ("John and Mary adopted a dog"), and verb phrase conjunction ("John cooked and ate some bacon"). A single rule makes them all count as conjunctions: a conjunction takes two expressions of the same semantic type, and produces a new expression of that type (Heim & Kratzer, 1998, pp. 182-183). Sentential conjunction takes two sentences and makes a new sentence; noun phrase conjunction takes two noun phrases and makes a new noun phrase; etc. Each kind of conjunction is continuous with the others, in that they are all understood using the same conceptual resources, even if no single kind of conjunction is more fundamental than the others. They are all generated by the same rule, but all the types of conjunction are not reducible to, say, sentential conjunction. So we can see how a moderate continuity thesis about some phenomenon can be compatible with rejecting reducibility. Thanks to Don Baxter for pushing me on this point.

possibility that we are all brains in vats. However, Searle departs dramatically from Bratman by saying that these collective behaviors cannot be understood using the same conceptual resources as individual intentionality, or "I-intentions". Instead, we must understand them in terms of "we-intentions." As Searle puts it (ibid, p. 406):

> The notion of a we-intention, of collective intentionality, implies the notion of cooperation. But the mere presence of I-intentions to achieve a goal that happens to be believed to be the same goal as that of other members of a group does not entail the presence of an intention to cooperate to achieve that goal.

According to Searle, reductionist approaches like Bratman's fail because they do not capture what is special about collective intentionality. Further, Searle argues, the only way to capture this robustly social notion is to admit it as a primitive phenomenon.

Searle proposes that collective intentionality is biologically primitive. What this means is unclear, but given some of his examples (birds building a nest together), perhaps he means that capacities for collective intentionality evolved independently of capacities for individual intentionality. For our purposes, though, we only need to focus on what it means for something to be primitive, and how it interacts with Searle's theoretical desiderata. Given that Searle's we-intentions purportedly cannot be understood as the participants' I-intentions (plus some common knowledge qualifications), I will take "primitive" here to mean conceptually primitive. Something is conceptually primitive when it cannot be analyzed in more basic terms. If

something is conceptually primitive, there is no true reduction of a theory that includes the primitive to a theory that does not include it. In this sense, we-intentions are conceptually primitive because the theory of collective intentionality cannot be reduced to a theory that does not invoke we-intentions, e.g. to a theory that only invokes I-intentions and common knowledge. Hence, for Searle, the strong continuity thesis is false. The understanding of collective intentionality involves conceptual resources other than those needed for understanding individual intentionality.

Searle's novel we-intentions are not meant to be the intentions of a group mind, however, so according to Searle, they are a special kind of intentional state had by individuals. While these states are had, *can only be had*, by individuals, they are not individual intentions, in the sense intended by Bratman. That is, they are not the intentions of individuals to do something themselves. Rather, they intentions of the form, "We are doing such-and-such." In Walk 1, for example, Alice is thinking, "I am going for a walk," whereas in Walk 2, she thinks, "We are going for a walk." What we have is not the intention of some new thing, a group, but a new kind of intention, a we-intention, for individuals like Alice.

As with Bratman, I will bracket any concerns about the viability of Searle's theory. Instead, I wish to flag the fact that, though they start from similar theoretical starting points, they end up in very different places. Bratman endorses a reductionist theory, where "we-intentions" are reducible to "I-intentions," while Searle endorses a nonreductionist theory where we-intentions are primitive. I propose that the directive model of commitment satisfies the guiding intuitions behind these very different views. While the understanding of collective intentionality involves no new

conceptual resources, it cannot be reduced to individual intentionality. This theory satisfies both the moderate continuity thesis and the rejection of reductionism.

On the moderate continuity thesis, the conceptual resources for understanding collective intentionality are homologous with those resources for understanding individual intentionality, but, contra the strong continuity thesis, individual intentionality need not have any conceptual or theoretical priority. The directive model of commitment satisfies this constraint. Any commitment at all involves both a producer of a directive and a consumer of the directive. This two-part structure explains the normative powers and liabilities of the parties concerned, even when the producer of the command is one and the same person as the consumer. The very same structure is present in joint commitments. The only difference is in how many producers and consumers there are, and how and by whom their commands are directed. The understanding of commitment, individual or joint, requires only the conceptual resources of the production and consumption of a command, and the normative statuses that result. The understanding of collective intentions, of joint commitments, involves no immodest expansion of our conceptual repertoire. Thus, the directive model of commitment satisfies the moderate continuity thesis.

It also satisfies Searle's nonreductionist intuitions, without making joint commitment primitive. In order to say whether a theory can or cannot be reduced, we need to identify which theory is purportedly being reduced, and which theory it is being reduced to. For Searle, no theory of collective intentionality can be reduced to a theory that only invokes individual intentions; we-intentions do not reduce to I-intentions. The directive model bears this out. On the directive model, two people

become jointly committed by producing directive representations directed at themselves and each other, and then consuming those commands. One person becomes committed on the directive model by producing a directive that they then consume alone (similar to a self-directed command). On the directive model, it is not possible to explain a joint commitment in terms of individual commitments alone, because the directives in question are both self-directed and other-directed; this feature of the commands is the core of the theory. If joint commitments were reducible to individual commitments, then some people would be jointly committed once the people involved had the right individual commitments. Individual commitments, however, can only enjoin oneself to action; a person is only answerable to themselves in individual commitments. Since joint commitment requires that one is answerable to more than oneself, something more is needed. The directives must be produced for other people, as well as oneself. But then they are no longer individual commitments. So the reduction is not possible. However, we can see how both individual and collective intentionality (or individual and joint commitment) are explicable in terms of the same conceptual resources. The directive model satisfies the moderate continuity thesis, but embraces nonreductionist aspirations while denying that joint commitment involves any sort of primitive notion.

Chapter 2

E Pluribus Unum?

Groups (plural subjects, to be specific) of all kinds populate the social landscape, from the maximally informal (like Gilbert's walking partners) to perfect pictures of formality (the Supreme Court), and everything in between. Like other constituents of commonsense ontology, groups present interesting metaphysical puzzles. Like ordinary objects, they exist contingently, trace a path through space-time, and persist through change. Unlike ordinary objects, they can coincide with each other, and sometimes, they don't always exist where their parts exist. This is the *group problem*: how does the metaphysical nature of groups explain these unusual features?

In tackling the group problem, we need to ask the right questions. There is not one question here (e.g. what are groups?), but several. There is the question of group identity: what makes one group the same group over time? But there is also the question of group ontology: what is the metaphysical nature of groups? What kind of thing is a group? To recognize the difference in these questions, we will

examine parallel issues in personal identity. We can then see how this distinction goes to work: some of the views offered about the metaphysics of groups do not really disagree with each other.

However, a problem arises for responses to the group problem. Cases involving reconstituted groups do not fit well with various approaches to the metaphysics of groups. Reconstitution cases suggest a kind of nihilism about groups; strictly speaking, groups have no deep metaphysical nature. Groups are jointly committed people, and what matters in the persistence of a group is the persistence of its joint commitments. Thus, the answer to the group identity question is that group identity is a loose and popular notion, and tracks the persistence of joint commitments. The answer to the group ontology question is that there are no groups, but only people and their commitments.

2.1   Persons and groups

Responding to the group problem requires some finesse in asking the right questions. What looks initially like a single question about groups might be better asked as two questions, as two approaches to different metaphysical issues in the same neighborhood. To illustrate the method, it helps to consider how it has been used in the literature on the metaphysics of persons, and how it has paid dividends there.

Much of the personal identity literature is concerned with answering the persistence question (Olson, 2014): what does it take for some past or future being to be me? The following gives us the most common formulation of the persistence

question, which we can call (Olson, 1997, p.25):

> **Narrow Question:** What are the conditions under which something that is a person at one time is identical with something that is a *person* at another time?

An answer to the Narrow Question is usually taken to be sufficient for answering the persistence question generally. Once we provide the necessary and sufficient conditions for a person at one time to be the same person as a person at an earlier or later time, we answer the Narrow Question and, *ipso facto*, the persistence question.

Even if we answer the Narrow Question, we may be left with unanswered questions about our persistence conditions. Suppose that some psychological criterion suffices to answer the Narrow Question, so that a person A at t is the same person as B at t* if and only if the right psychological relation holds between A and B. This tells us less than we might expect about what it takes for us to persist through time. If it's true, it does tell us what it takes for persons to persist, *qua* persons, but there are many different kinds of persons. Some believe that God is a person; others believe that robots (like Data from *Star Trek*) might one day be sophisticated enough to be persons. It seems unlikely that a psychological criterion could tell us enough about our persistence conditions, unless its advocates intend to say that we could share our persistence conditions with divine or artificial persons. If we wish to say more about the persistence conditions of human persons, we can broaden the scope of the persistence question, and ask the following (ibid.):

> **Broad Question:** What are the conditions under which something that

is a person at one time is identical with *anything at all* that exists at another time?

The Broad Question affords more flexibility in how we understand our persistence conditions. Note that, on psychological answers to the Narrow Question, there is no way to make sense of saying that my 5-year-old niece, who is now a person, was once a fetus with no capacity for psychological states. Since she has no psychological relation with any fetus at all (except, perhaps, one very late in development (Olson, 1997, p. 75), she is not identical with any fetus. There were never any ultrasound pictures of my niece, because it is not possible for the fetus depicted to be my niece — at least, according to psychological approaches to the Narrow Question about persistence.

This is a surprising result, and one that motivates the Broad Question understanding of our persistence conditions (see chapter 4 of (Olson, 1997)). I suspect that my niece was once a fetus, that I literally saw her in ultrasound pictures, perhaps before she was a person. Loose and popular talk supports such intuitions, in any event. To be fair, there may be a good argument for the thesis that, strictly and philosophically, anything that is a person is a person essentially. In that case, any questions about our persistence conditions would be answered with an answer to the Narrow Question, with an answer that specifies necessary and sufficient conditions for a person at one time to be identical with a person at another time. But this position – that any person is necessarily a person – requires an argument.

Though I agree with Olson, the purpose of this discussion is not to argue that his fetus argument is a conclusive argument for rejecting psychological approaches

to personal identity, or to reject the understanding of our persistence conditions in Narrow terms. Instead, I take the debate to stake out distinct questions about the metaphysical nature of persons: questions about personal ontology, and questions about personal identity.[1] It is one thing to ask what it takes for me to count as the same person over time; it is another to ask what kind of thing a person is (Olson, 2007, p.16). Suppose some psychological answer to the Narrow Question is true, so that the person I am today is identical with the person I will be tomorrow because some suitable psychological relation obtains between me today and me tomorrow. This leaves open an important question: what kind of thing am I? What are persons like us? For all the psychological approach says, I could be an immaterial mental substance, or a bundle of mental states, or my brain (Thomson, 1997, p.205). The psychological approach cannot tell me whether I am a material thing.

Thomson puts the point another way. Suppose we adopt the personal ontology thesis that people are their bodies (Thomson, 1997, p.204). This entails a personal identity criterion: x = y if and only if x's body = y's body. The entailment does not hold in the opposite direction, however. Even if x and y are identical because x's body is y's, x need not be identical with x's body. For the sake of analogy, consider that:

$$\{a\} = \{b\} \text{ if and only if } a = b.$$

The identity of a singleton is determined by the identity of its member, so that the singleton of a is the singleton of b when a is identical with b. Even so, $\{a\} \neq a$.

---

[1]Olson (2007, p. v) endorses Thomson's (1997, pp.204–205) way of framing the issue, and I follow their lead here.

Even though the criterion of identity for sets is the sameness of their members, sets are not identical with their members. The criterion of identity does not fully settle what sets are. We should not expect criteria of personal identity to settle fully what we are either.[2]

Following Olson and Thomson, I will use "personal ontology" to refer to the broad metaphysical questions about the nature of persons. What sort of thing are we? Does that tell us the conditions under which we exist, even if not as a person?[3] Questions of personal *identity*, on the other hand, concern the narrower questions about persons, like what it takes to count as a person, or what it takes for a person to persist *qua* person through time. I will argue in the next section that a similar distinction holds for groups, the distinction between questions of group ontology and group identity.

2.2   Two questions about groups

The metaphysics of groups is often understood as the singular, unified group problem, but it's not hard to see how the distinctions found in the personal identity literature would extend to questions about groups. To bring out such differences, consider a familiar scenario. Suppose that the graduate students at State University have met

---

[2]Compare these questions about sets with questions about the nature of propositions. What are the criteria of identity for propositions, i.e. when do you have one proposition rather than two? Here's one proposal. Let p and q be propositions, and P and Q sentences that express them respectively. Here's a proposed criterion of propositional identity: $p = q$ if and only if $P \equiv Q$. It tells us when we have one proposition rather than two, but doesn't tell us anything about what propositions might be. Thanks to Andrew Parisi for discussion on this point.

[3]Is there something that was once a fetus and is now writing this paper? If so, what is it that does so? The answer to this question tells us our deep metaphysical nature (Wiggins, 1980).

to discuss unionization, and State U's administrators hear of the meeting. It would be sensible for one administrator to say, "They could be a union by next year." Presumably, "they" are the graduate students that met to discuss unionization. The administrators might want to prevent those students from being a union. If the students do become a union, they will have gone from not being a union to being a union. In an address to their fellow union members, the leaders could say, "Once upon a time, we were not a union, and now we are."

This is similar to the intuitions driving Olson's fetus argument. In that argument, we found that standard views about persons clashed with the intuition that we were once fetuses. If our existence is inextricably bound up in our psychology, then we could never have been a fetus, since we never bore any psychological relationship with such a fetus (or at least, not until very late in development). The union case trades on similar intuitions. The union bears a significant kind of continuity with the pre-union graduate students, similar to the continuity I bear with a particular fetus. However, this intuition is in tension with predominant views about the metaphysics of groups. This tension in turn puts pressure on us to recognize two questions we can ask about groups. The first corresponds with standard questions about group identity:

Narrow Group Question: What are the conditions under which some things that are a group at one time are identical with some things that are a *group* at another time?

This contrasts with a related, but distinct question:

Broad Group Question: What are the conditions under which some things that are a group at one time are identical with *any things at all* that exist at another time?[4]

If the Broad Group Question is legitimate, it permits us to think that some things in the union case start out as graduate students, and are later a union, not unlike how the Broad Question about persons allows us to think that something can start out as a fetus and later be a person. The parallel, of course, is not perfect. If the fetus argument is successful, it shows that persons are not persons essentially. Instead, "person" might be a phase sortal, something that only applies to a given substance during a particular phase of its existence. If we expect the parallel to hold strictly, we should expect "union" to behave like a phase sortal too, applying only to a particular phase of an object's career. In the case of groups, however, it is not clear that there is a corresponding substance that the sortals apply to. What substance could start as students and later be a union?[5]

The parallel does not need to apply that strictly in order for the contrast between distinct issues to come to the fore. The Broad and Narrow questions bring out contrasting questions to ask about groups. First, what makes one group the same group over time? What determines when you have one group, rather than many? Second, when do some things become a group, or stop being a group? What sorts of things could meet the requirements to be a group? Call the first kind of question

---

[4]Formulating the questions this way may threaten to prejudge the issue; if some things are a group, rather than something, then plural constructions like sets or fusions are ruled out. However, "some things" should be read as neutral between singular and plural reference; here I follow the reading of plural variables in Boolos (1984).

[5]Interestingly, Ruben (1985) is perhaps the only philosopher to frame the issue in terms of social substances.

a question about group identity; the second kind is a matter of group ontology.

The distinction between these questions is not much deployed by those working on the metaphysics of groups. Instead, they tackle the group problem by finding some sort of thing that could satisfy a cluster of theoretical desiderata, based on the canonical features of groups that need to be explained.

The first thing to sort out is the relationship that groups bear to their members. At a glance, it seems "obvious that there is some close relation between social entities and human individuals. If social entities exist, it cannot just be a contingent fact that some individuals do so as well" (Ruben, 1985, p.45). However we understand this relation, and however strong it is, the group's existence is intuitively bound up with the existence of its members. Even so, groups can change their members across time and across possible worlds, while remaining the same group. We can say, truly, that the Supreme Court ruled on Roe v. Wade in 1973, though none of the Justices that sit on the Supreme Court now ruled on Roe v. Wade in 1973 (Uzquiano, 2004). The present Court and the 1973 Court are the same group, despite their non-overlapping membership. Likewise, we can say of the Court that it could have included Harriet Miers instead of Samuel Alito; if we can truly say that, then the Court's membership is a contingent matter. The Court can have different members at different times, and different members at different possible worlds.

Likewise, the Court itself can exist (or not exist) at different times, and could possibly have existed (or not existed) at different possible worlds. In the actual world, the U.S. Supreme Court has existed since circa 1789 (the precise date is murky; see chapter 11 of (Epstein, 2015)). In another possible world, the Court does not exist

at all. Indeed, the Court could have failed to exist, even though all of its would-be Justices existed. Further, just as the Court exists in time, it exists in space as well, at least sometimes. When the Court hears oral arguments, they do so in Washington, D.C. When the Court is not in session, and none of the Justices are in D.C., its physical location is unclear. Even if all the Justices go to the opera together one night, it would be counterintuitive to say that the Supreme Court went to the opera.[6] Hindriks (2013) calls this the *location problem*: where exactly are social entities? The location problem aside, it is commonly assumed that groups sometimes have spatial location.

Finally, groups like the Supreme Court can coincide with distinct groups, i.e. two groups can be non-identical and have exactly the same members. The chess club and the nature club can have exactly the same members, yet be distinct groups (Ritchie, 2013, p.258); likewise for the Supreme Court and the club made up of its members (Uzquiano, 2004, pp.146–147). This makes groups unusual among the other, putatively concrete things in the world. Some philosophers deny outright that distinct objects can occupy the same space at the same time. Others are more permissive, and allow objects to coincide so long as the objects belong to different sorts. As permissive as these views are, however, they still deny that objects of the same sort can coincide.[7] For example, while a statue could share all of its parts with a lump of clay, that statue could never completely coincide with another statue. Groups do not face this restriction. Chess clubs and nature clubs are of the same

---

[6]Ruben (1983, 1985) provides arguments along these lines.
[7]Sidelle (1998) provides a good overview of the issues here.

sort, and yet they completely share membership. They are thus said to coincide.[8]

By taking the aforementioned to be desiderata in forming a theory of groups, the many theories on offer put forward answers to what they take to be a single problem — the group problem — when in fact that they are trying to answer several questions at once. To see how this plays out, we'll look at two sorts of popular answers that try to answer the above desiderata. The first family of answers take groups to be *identical* with some sort of plural construction, e.g. a set or mereological fusion. The second family of answers take groups to be *constituted* out of, but not identical with, their members, in a way similar to how statues are constituted out of clay. In the remaining sections, I will show that some of these views are better suited to the group ontology issue, but not the group identity issue, and vice versa. But first, we must look at the details of such views.

_____

[8]The case for coinciding groups is sometimes made too quickly, by pointing out that one group's meeting need not be a meeting of a distinct group, even if both groups completely share their members. For example, a meeting of the chess club need not be a meeting of the nature club, even if they share all their members. But this isn't enough for coincidence of the metaphysically interesting kind, where distinct things exist at exactly the same place and time. If I disassemble a chair, and use all and only those materials to make a table, the chair and table might share all their parts, in a very weak sense, but this doesn't make for a case of coincidence, since the table and chair do not occupy the same place at the same time. A genuine case of coincidence would show not only that two groups can have the exact same members, but also that two groups could be wholly present at the same place and time. Here's such a case. Suppose all the student groups on campus gather at the student union to protest budget cuts to student organizations. All the members of the chess club and nature club show up; as it turns out, the chess club and nature club share all the same members. A count of the student organizations that attended the protest could reasonably count the chess club and nature club separately, without engaging in double counting. In that case, the chess club and nature club would coincide in the controversial way that lumps and statues are said to coincide. Moreover, though, they coincide in a way that statues and lumps do not. Statues and lumps belong to different metaphysical kinds, which is why they have different persistence conditions. The chess club and nature club, on the other hand, belong to the same metaphysical kind. They are both groups, and presumably, groups have the same persistence conditions by virtue of being groups. It is not uncommon to hold that coinciding objects of different kinds is uncontroversial, but objects of the same kind cannot coincide. Yet groups do just that. This makes coinciding groups truly unusual.

Consider the identity views first. One straightforward and attractive approach to groups is to identify them with sets (Uzquiano, 2004, p.135). The Supreme Court is the set containing only the Supreme Court Justices, and the chess club is the set of chess club members, and so on. Group membership is understood in terms of set membership. Call this the "naive setist" view. As noted by Uzquiano and Ritchie (2013, p.266), the naive setist faces problems immediately. Since sets are identical when they have all the same elements, groups that have all the same members would be identical. The chess club and the nature club, having exactly the same members, would be the same group, since they are the same set. Thus, the naive setist fails to explain coincidence. They also fail to explain how groups can exist in space (e.g. how the chess club can meet in the student union, rather than the park), since sets are abstract, non-spatial entities. The change in membership across time and possible worlds is also unexplained, since sets have their members essentially, and groups do not. Needless to say, the naive setist faces serious challenges.

Luckily, Effingham (2010) has offered a more sophisticated setist view that is meant to avoid these issues. On Effingham's view, the Supreme Court is not to be identified simply with the set of its current Justices, because, as we've seen, there are different sets of Justices at different times and different worlds, but the same Supreme Court throughout. Instead, the sophisticated setist proposes that the Supreme Court is identical with a set of ordered pairs. The pairs in question have a possible world as their first member; as their second member, a set of ordered pairs of times and sets of individuals (or the empty set). In Effingham's (2010, p.259) formulation, "g is a group" is defined as follows:

(a) g is a set with only ordered pairs as members$_S$; (b) every possible world is the first member of exactly one ordered pair that is a member$_S$ of g; and (c) the second member of each such ordered pair is itself a set of ordered pairs such that (i) every instant is the first member$_S$ of exactly one of those latter ordered pairs, and (ii) the second member$_S$ of each of the latter ordered pairs is either the empty set or a set of individuals.

The sets in question seem to build in changes in existence and membership across times and worlds. While the Supreme Court is not identical with the set of its current Justices, it is a set all the same. We could read off that set information about where the Supreme Court exists, and when it exists (if at all) in those worlds. At a glance, Effingham's proposal does seem to capture the modal and temporal aspects of the existence of groups and group membership, since one can read off g's ordered pairs the membership of g at a time and world. However, looking at g's ordered pairs does not tell us the location of a *group* in a world, at a time. We might think that the Supreme Court is wholly present when all nine Justices convene in the courthouse to hear a case, but this is not so, according to the sophisticated setist. When the Justices convene, only the set of members at the time in that world are present, but not the Supreme Court itself. The Supreme Court itself is a set whose members are ordered pairs of worlds and sets, which have pairs of times and sets as members. Strictly speaking, nobody ever sees the Supreme Court, or a band, or any group; they only see the membership at a given world and time. This is a counterintuitive result, insofar as we think that, say, when The Beatles performed on American Bandstand, the band itself performed, rather than the elements of a subset of a subset.

50

The sophisticated setist faces further counterintuitive results that stem from the nature of sets themselves. Since sets are abstract objects, they do not come into being. We might say that The Beatles, for instance, did not exist before 1957 (ignoring complications involving the formation of The Quarrymen). The sophisticated setist would have to say this is false, however, since the set which *is* The Beatles has always existed (or exists timelessly, perhaps). So again, the sophisticated setist has to deny commonsense, or heavily qualify it.

The sophisticated setist approaches the group problem by finding a familiar plural construction (in this case, sets) and seeing how it can have the features commonly attributed to groups. Similar approaches exist for those skeptical of sets. Some propose that groups are mereological fusions of their members, and group membership is understood as parthood. Still others suggest that groups are non-singular pluralities that are not unified in the sense that sets or fusions are unified; these suggestions appeal to the machinery of plural logic (such as that found in McKay (2006), Boolos (1984), or Oliver and Smiley (2013)) to handle talk of groups. Likewise, the view of many-one identity developed by Baxter (1988) could explain how a group can be literally identical with the members.

Still other approaches reject the notion that we should identify groups with some plural construction or other. The relation between groups and their members is not one of identity, but constitution. Constitution has been used to explain how objects could coincide (e.g. a statue and the lump of clay that makes it up (Thomson, 1998)), and it does similar work here. The identity approaches have to show how plural constructions could bear the properties typically attributed to groups, like

the ability to persist through change in membership. But many plural constructions have a tight connection with the individuals that make them up. Sets (or fusions, or non-singular pluralities) have their members essentially. In this way, they are not unlike lumps or heaps, which have their constitutive particles essentially. As Locke (*Essay* II.xxvii.4) puts it:

> ...if two or more atoms be joined together into the same mass, every one of those atoms will be the same, by the foregoing rule: and whilst they exist united together, the mass, consisting of the same atoms, must be the same mass, or the same body, let the parts be ever so differently jumbled. But if one of these atoms be taken away, or one new one added, it is no longer the same mass, or the same body.

While lumpy kinds have their parts necessarily, living things do not. Locke goes on to say:

> In the state of living creatures, their identity depends not on a mass of the same particles, but on something else. For in them the variation of great parcels of matters alters not the identity: an oak growing from a plant to a great tree, and then lopped, is still the same oak; and a colt grown up to a horse, sometimes fat, sometimes lean, is all the while the same horse: though in both these cases, there may be a manifest change of the parts; so that truly they are not either of them the same masses of matter, though they be truly one of them the same oak, and the other the same horse. The reason whereof is, that in these two cases, a mass of matter, and a living body, identity is not applied to the same thing.

Groups do not have their members necessarily, just as trees, horses, and perhaps statues do not have their parts necessarily. Plural constructions, like lumpy kinds, do have their constituents necessarily.[9] If a group is to be identical with a plural construction, it will inherit the plural construction's modal properties, including its tight connection to its constituents, thus making it a poor candidate to be a group.

This makes constitution, rather than identity, an appealing explanation for how groups are tied to their members, because it captures the sense in which two things can share their parts while having different modal connections to their parts. Thus, Uzquiano provides the following analysis of what he calls "group-constitution" (Uzquiano, 2004, p.150). Anything that is a member of the set S is a member of group G, and vice versa. Further, while a member of G is necessarily an element of S, it is not necessarily a member of G. Likewise, while current non-members of G could not possibly be elements of S, they could be members of G. Hence, while G is "made up" of S in some sense, it is not identical with S.

Epstein also promotes a constitution analysis of groups, but he holds that constitution is to be understood in terms of grounding. As he puts it, "constitution is not about one thing being ontologically related in the right way to another thing's existence, or to another thing's persistence. Rather, it is about the stuff of one thing being part of the metaphysical reason that another thing is made up of the stuff it is" (Epstein, 2015, p. 149). For Epstein, the individuals constituting the Supreme Court do not fully ground facts about the Supreme Court, but they are part of the

---

[9]As Uzquiano (2004, p.150) makes the point, "The insight behind Thomson's account of material constitution lies in the observation that while it is indeed the case that a statue and the portion of clay of which it is made share all of their parts at a time, the portion of clay is more closely tied to them than the statue is."

explanation for why the Supreme Court is made up of those individuals in the first place. Ritchie (2013) departs from Epstein and Uzquiano's constitution views, but her structural realization view is fairly close in spirit. Like Epstein and Uzquiano, she denies that groups are identical with plural constructions. Instead, people "make up" groups by realizing group structures. According to Ritchie:

> The structure of a group can be represented with nodes (or places) and edges connecting nodes to other nodes. The edges of a structure capture the relations that hold between nodes. Since all members of a group are related to some degree, each node in structure S is connected to every other node in S. Functional relations connect nodes in group structures. Some of these functional relations might be hierarchical in nature. (Ritchie, 2013, p.268)

A group exists once its structure is realized, i.e. once all of the structure's nodes are occupied. This yields conditions for membership in a group, and conditions for synchronic identity (while Ritchie does not give the explicit statement for persistence through time like she does for membership and synchronic identity, she does say that group persistence consists in the continuity of the realization of a group's structure). Like the constitution views, Ritchie's structural view tells us the conditions under which groups come to be and continue to be, and what group members have to do with those facts.

Constitution views do not tell us much about what it takes for some things to be a group, however. The peculiar features of groups, like their ability to coincide and persist through change in membership, are explained by facts that have little

to do with plural constructions made out of the group members *per se*. Ritchie's structural realization view, for example, says that individuals must jointly realize a structure for a group to exist; persistence consists in continuous joint realization of the structure. It is the structure that does most of the explanatory work here, that unifies the group and accounts for its persistence conditions. The individuals matter insofar as they stand in the relations described by the structure; otherwise, they do not contribute much. Uzquiano has sets of individuals constitute groups (e.g. the set of the current Justices constitutes the Court), but while the use of sets allows some precision in specifying relations between individuals and the collections they are in (through familiar set-theoretic relations like membership in a set), sets do not do any theoretical work that could not have been done by mereological fusions (where we use part-whole relations) or non-singular pluralities (where we use plural reference and the "among" relation). And like many constitution theories, Uzquiano does not have much to say about why constitution occurs when it does. While the analysis of constitution he gives lays out the modal properties of sets, groups, and their relationship in clear terms, it does not explain how that relation comes to obtain when it does obtain. He does say that groups are *sui generis* objects constituted out of sets of their members, hinting at a theory of group ontology. But saying that groups belong to a *sui generis* category does not say whether groups are material, or have spatial location. He compares groups to artifacts, because like artifacts, there is more to a group's existence than its constituent parts. However, while we know that artifacts are concrete, material particulars, we do not know where, according to Uzquiano, groups fit into our ontology. They are *sui generis*, but the constitution

relation on its own does not explain why. For all Uzquiano says, they might be material particulars, like Sheehy (2006) claims.

To put the point another way, constitution views only answer the Narrow Group Question. Consider a reading group that has met for the past five years and has gained and lost members multiple times. If there is a kind of continuity between the reading group and the people who *became* the reading group, or the reading group and the people left over should it dissolve, constitution views have nothing to say about that continuity. There is no room in constitution views to say anything about groups before or after their being groups; there are no resources to make sense of some people becoming or ceasing to be a group. In fairness to constitution views, it might be that, if something is a group, it is a group essentially, in which case, constitution views do not need to say anything about continuity between groups and mere pluralities. If groups are essentially groups while they exist, then asking what it is that is a group at one time but not another is asking an ill-formed question. If some x is essentially F, then asking what x was before it was F is to ask what x was when x did not exist. However, the claim that groups are essentially groups, that "group" is like a Wigginsian (Wiggins, 1980) substance concept, must be argued for. If constitution views only tell us about group identity (that is, they only answer the Narrow Group Question), then questions about group ontology are left open.

2.3   The Broad Group Question

The Broad Group Question asks, roughly, what groups are made of; what kind of thing are they? Answers to the Narrow Group Question specify the relationship between groups and their members, and between different temporal stages of a group. They provide individuation criteria, telling us when we have one group rather than several. Answers to the Broad Group Question, on the other hand, tell us what groups are; they tell us what sorts of things can be groups at all.

Why does this question matter, though? Answers to the Narrow Group Question purport to tell us when a group G_a at t_1 is the same group as G_b at t_2. However, there are *reconstitution cases* where it is unclear whether we have the same group across times. If an answer to the Narrow Group Question cannot say whether a group continues to exist in a reconstitution case, then perhaps an answer to the Broad Group Question can fill in the gaps. It could tell us what sort of things become groups, and the conditions under which they do.

To start, consider the first reconstitution cases:

Hiatus: In 2006, Sleater-Kinney announces they are going on indefinite hiatus, meaning that they have no intentions to record another album or go on another tour. Instead, members will pursue their own projects. In 2014, they reunite with the original line-up.

Breakup: In 1993, the Pixies announce the end of their band. They will release no new albums, or go on tour together. Instead, members will pursue their own projects. In 2003, they begin rehearsing together again,

and decide to reunite with the original line-up.

Indefinite hiatuses and breakups have two things in common: the cessation of continued activity as a group, and the lack of intention to continue acting as a group. But is there a deep, metaphysical difference, so that something is true of Sleater-Kinney on hiatus, but not true of the broken-up Pixies? Does Sleater-Kinney survive in suspended animation, whereas the Pixies undergo resurrection, going completely out of existence only to come back later?

One could plausibly respond to these cases by saying that, in *Hiatus*, the band in 2006 survives during their hiatus, but in *Breakup*, the band goes out of existence. The Pixies do not survive their breakup; they go out of existence. When they break up, the band goes out of existence. By contrast, going on hiatus does not bring about Sleater-Kinney's end. Some theories of group persistence can explain why this might be. Ritchie, for instance, says that "[t]he persistence of [a group] G requires the continuity of the realization of [G's structure] S" (Ritchie, 2013, p.270). During a hiatus, a band does not stop realizing the structure of their group; the band's structure is continuously realized, albeit in a very attenuated way. However, when the Pixies break up, they no longer jointly realize their band's structure in any way, breaking the continuity in its realization. Hence, while Sleater-Kinney survives their hiatus, the Pixies' breakup brings their existence to an end.

There's something intuitive to this, since hiatuses are not meant to be permanent pauses in activity, even when they are indefinite. This does not yet explain what happens to Sleater-Kinney in 2014, or to the Pixies in 2003, since we do not yet

know what continuity is in Ritchie's theory of group persistence.[10] If continuous realization is *gapless* realization, then the Pixies' breakup is permanent. In 2003, a brand new band, exactly like the Pixies of 1993 with exactly the same members, comes into existence. This understanding of continuous realization is compatible with how Ritchie elaborates on group persistence: "[O]ccupiers of the nodes of S which form the realization of group G (i.e., members of G) must continue to bear the functional relations required for a realization S" (ibid.). G goes out of existence permanently when the realization of its structure ceases. Hence, the Pixies go out of existence permanently following their breakup in 1993, and a new band with the same members and name is formed in 2003. They do not literally get the band back together, because the band is gone. If the Pixies of 2003 are the Pixies of 1993, they are only identical in the *loose and popular* sense, not the *strict and philosophical* sense (to borrow Butler's coinage (Baxter, 1988)).

This is a surprising result if we think that it is *literally* possible for the band to get back together, to reconstitute the same band, strictly and philosophically[11]. A notion of continuity weaker than gapless realization could explain how the Pixies in 2003 are the Pixies in 1993, despite the gap in their existence. Continuous realization here amounts to continued, or merely resumed, realization, so that a group G_a with structure S at time t_1 is G_b at a later time t_2 only if G_b is a realization of S. A group exists whenever its structure is realized, regardless of whether that structure is unrealized at some times. The Pixies return in 2003 because they realize the Pixies'

---

[10]Uzquiano's view is worse off in this regard, since he does not specify conditions under which constitution happens.

[11]After all, if you were a Pixies fan going to a Pixies show in 2004 advertised as a reunion show, wouldn't you want to see the Pixies play, and not just a Parfitian successor?

1993 structure.

However, if continuous realization is simply resumed, not gapless, realization, then we get an unclear, if not counterintuitive, result in the following case:

> *Student Association*: The Philosophy Graduate Student Association (PGSA) has a constitution C that defines the roles of its officers and how meetings proceed. C is the PGSA's structure, and its members in 1995 realize C. However, subsequent years of students cease to have PGSA meetings, leaving C unrealized, ending the existence of the PGSA. In 2015, a new student finds a physical copy of C tucked into a book in the department library. The graduate students in 2015, who were not students in 1995, decide to form the PGSA in accordance with C.[12]

Does the PGSA reconstitute in 2015, or is it merely a new group with the same name and structure? Suppose that, by 1995, the PGSA had existed for years, had passed several budgets, and several drafts of its constitution. In 1990, the PGSA drafts its constitution, C, which, unlike previous constitutions, does not include a role for a treasurer. It is true of the PGSA in 1995 that it once had a treasurer. Suppose that the PGSA's president in 1995, Sophie, was not a student in 1990, and did not become a member of the PGSA until 1991, by which time the PGSA had eliminated the treasurer role. When she becomes president in 1995, Sophie could say truly, "We once had a treasurer." She says, of the PGSA (a group of which she counts herself as a member), that it once had a treasurer, even though she did not belogn to the PGSA in 1990, the last time the group had a treasurer.

---

[12]This case bears some similarities to Parfit's club case in Part III of (Parfit, 1984).

However, it is unclear whether the PGSA in 2015 ever had a treasurer. If Sam, the president of the PGSA in 2015 says, "We once had a treasurer," it is unclear whether he speaks the truth. Importantly, it is not unclear because he refers to a time when he was not yet a member; his statement is not false simply because he uses the first-person plural pronoun to refer to a group that did not yet include him. Sophie, after all, uses the first-person plural pronoun to refer to a group that did not at the time include her either. So if Sam says something false, it is not because he speaks of himself as if he was a member of the group when they had a treasurer. Rather, the mistake is in the assumption that he is a member of a group that once had a treasurer.

If the PGSA in 2015 is not the PGSA in 1995, even though it realizes the structure C, then realization of the same structure is not sufficient for a group to persist. A group cannot reconstitute after a period of non-existence merely by having its structure realized again. *Breakup* shows that gapless realization is not necessary for group persistence, since the Pixies reconstituted, while Student Association shows that resumed realization is not sufficient for group persistence, since the 1995 PGSA is not the 2015 PGSA, despite having the same structure. If, as Ritchie says, group persistence requires continuous realization of structure, then merely resumed (not gapless) realization would be necessary, but not sufficient, for group persistence. But what explains why the Pixies reconstitute in *Breakup*, whereas the PGSA in *Student Association* does not? What further feature besides resumed realization accounts for the Pixies' reunion, which the PGSA lacks?

When the Pixies reunite, they do so with the band's line-up from 1993. By

contrast, no member of the 2015 PGSA was a member of the 1995 PGSA. This suggests that the sameness of members, in addition to sameness of structure, enables the Pixies' reunion. But recall that, for Ritchie, "occupiers of the nodes of S which form the realization of group G (i.e., members of G) must continue to bear the functional relations required for a realization S" in order for G to persist (Ritchie, 2013, p.270). In *Breakup*, the occupiers of the nodes in 1993 occupy them again in 2003, whereas in Student Association, no occupier in 1995 occupies a node in 2015. The change (or lack thereof) in membership alone does not explain the reconstitution facts in these cases, since groups change their members all the time. Indeed, survival through change in membership is a canonical feature of groups, embodied in Ritchie's view and Uzquiano's analysis of group constitution.

However, membership in *StudentAssociation* does not just change — the PGSA undergoes a *total replacement* of members. The Pixies genuinely reunite in *Breakup*, and do not merely form an exactly similar band, because those very people continue to realize the structure ten years later — the members pick up where they left off. The 2015 PGSA does not. While gapless realization is not necessary, merely resumed realization is not sufficient. The right relation must obtain between an earlier realization and the later realization that resumes realization. This relation would obtain when the same members resume a structure's realization, but fails when there is a total replacement. However, the relation cannot preclude group persistence through change in membership. While reconstitution cannot occur in cases of total replacement, it could occur in cases of gradual replacement, or in cases where there is no replacement at all.

Since replacement occurs by degrees, let R be the degree of replacement that is sufficient for a later realization of a structure S to count as a resumed realizer of a group G with structure S. *Breakup* and *Student Association* present the limit cases of R, where R is determined by how much overlap occurs between the membership of the earlier and later realizers. In *Breakup*, there is total overlap (there is no replacement), whereas *Student Association* has no overlap whatsoever (there is total replacement). Suppose there is a group G with a structure S that has exactly four nodes. At t, for every node in S, there is exactly one member, and no member occupies more than one node. At a later time t*, one of those members leaves and is replaced by a new person. From t to t*, G's degree of replacement R is .25: one in four node occupants have been replaced. Presumably, this degree of replacement would be low enough for G to persist from t to t*. Accounting for the difference between *Breakup* and *Student Association* is a matter of finding the right value for R, somewhere between 1 (total replacement) and 0 (none).

Could R be .5, such that, when more than half of a group is replaced in a resumed realizer, the group does not survive? This presents us with a difficulty in the following case.

> *Breakup 2*: In 1993, infighting among the band members causes the Pixies
> to break up. In 2003, Frank and Joey decide to continue playing music
> together, and Kim and David do likewise. Call Frank and Joey's band A,
> and Kim and David's B. Frank and Joey find a new bassist and drummer,
> and Kim and David find two new guitarists. Both bands, A and B, claim
> to be the Pixies reunited.

63

In *Breakup 2*, which band in 2003, A or B, is identical with the Pixies of 1993? On the assumption that .5 is a sufficient value of R for being a resumed realizer, A would seem to be a resumed realizer for the Pixies, and so would be the Pixies. However, by that token, B would be the Pixies as well, and this poses a problem. Once we say that a value of .5 for R suffices for identity between resumed realizers, we must say that the Pixies of 1993 are identical with both A and B, even though A and B are not identical. Thus, the Pixies of 1993 are identical with two distinct bands in 2003: identity is no longer a one-one, but one-many relation. So what are the reconstitution facts in *Breakup 2*? We have four options: (1) the Pixies do not reconstitute, (2) the Pixies reconstitute as A, not B, (3) the Pixies reconstitute as B, not A, (4) the Pixies reconstitute as both A and B. Choosing (4) would give up on identity as a one-one relation, and choosing (2) instead of (3) (or vice versa) would seem unmotivated, since, by hypothesis, we have as much reason to choose either one. The only option left is (1): neither A nor B are the Pixies. The Pixies do not reconstitute in 2003.

What if the problem here arises merely because the Pixies' resumed realizations branch out, i.e. they have more than one resumed realizer? What if things had not been like *Breakup 2*, but instead like the following?

> *Breakup 3*: In 1993, infighting among the band members causes the Pixies to break up. In 2003, Frank and Joey decide to continue playing music together, while Kim and David stop playing music entirely. Frank and Joey recruit a new bassist and drummer, and begin recording and touring as the Pixies.

What are the reconstitution facts in *Breakup 3*? It is *prima facie* plausible that Frank and Joey, playing with the new bassist and drummer in 2003, make up the very same band that dissolved in 1993. This might suggest that the Pixies reconstitute in 2003 because they are a resumed realizer with a .5 value for R, and there are no other resumed realizers. However, arguments from Parfit (1984, 1971) show that "no branching" clauses do not help here. The mere fact that there are no other resumed realizers besides Frank and Joey's band cannot make their band a reconstitution of the Pixies of 1993. Since the identity between a group and its resumed realizer would presumably consist in relations that the group bears to itself alone, it is hard to see how states of affairs external to the resumed realizer in question could bear on whether the realizer is identical with the earlier group, so non-branching cannot explain the reconstitution facts in *Breakup 2* and *Breakup 3*.

Determining why reconstitution happens (or seems to happen) in some cases (like *Hiatus* and *Breakup*) but not others (like *Student Association* and *Breakup 2*) seems to depend on what it is to continue to realize the same structure. Gapless realization seems to be too strong a condition, while resumed realization seems too weak. But it does seem as though groups persist, in a sense, by "passing the torch" to their successors, and how else to do that but by having successors resume realizing the group's structure?

2.4 Joint commitment and group ontology

As noted before, the Narrow Group Question tells us about relations that obtain between groups and their members (to be more specific, their members at a time), and between a group and itself across time. However, Narrow answers like Ritchie's do not tell us enough about reconstitution cases, about whether there is any difference between *Hiatus* and *Breakup*, and if so, why. Consider Ritchie's theory, that a group exists when some people jointly realize the structure. Nothing in the theory tells us what it takes for people to realize a structure in the first place. Her theory could be supplemented with a theory of realization, or a theory of building relations generally (Bennett, 2011), but even that supplemented theory would not differentiate between mere collections of people who happen to realize a structure by accident, and groups of people who realize a structure in a way sufficient for making a group. Suppose G exists once its structure, S, is jointly realized by some people, $M_1...M_n$. For all Ritchie says (and Uzquiano, for that matter), $M_1...M_n$ could realize S unintentionally, and G could exist unbeknownst to them. To rule out this possibility, we could say that jointly realizing a structure, or becoming a group, is something done on purpose. For some people to become a group, they must jointly commit to doing so.

Some people have a joint commitment when they have committed themselves and each other to doing some activity. Suppose, then, that joint commitment plays a role in explaining how some people become a group. For example, people could become a group by jointly committing to realizing a group structure, and then jointly realize it by virtue of their joint commitment. This does not merely restate Ritchie's view. It

explains how some people come to realize a structure in the first place, and the sort of continuity that obtains between people before they form a group, and the group that forms after. This looks more like an answer to the Broad Group Question, rather than the Narrow. Provisionally, call this the *Joint Commitment* view: regardless of the right answer to the Narrow question, joint commitment is needed to explain how people become members of groups, and the relevant reconstitution phenomena.

To see how the Joint Commitment view works, consider the reconstitution intuitions in *Hiatus* and *Breakup*. The difference between a hiatus and a breakup, on the Joint Commitment view, is that the former does not dissolve the joint commitment that keeps the band together, whereas a breakup does. An indefinite hiatus leaves the possibility of further group activity open; this plural subject might share activity again. A breakup, by contrast, dissolves the normative structure of joint commitment altogether. Hence, it is plausible that Sleater-Kinney does not go out of existence when they go on hiatus, but the Pixies do go out of existence, because their joint commitment dissolves when Frank declares on the radio that the band is over.

However, while it is no surprise that Sleater-Kinney survives their hiatus (their joint commitment never went away), the Pixies' survival might come as a surprise, since the plural subject that, say, realized the Pixies' structure, dissolved. How could a plural subject dissolve and come back into existence? That is, why do the Pixies come back in 2003 after their dissolution in 1993, but the PGSA does *not* come back in 2015 after their dissolution in 1995? While Narrow answers could not answer this, the Joint Commitment view can. The difference between the Pixies and the

PGSA comes to a difference in how the two groups are related to the earlier joint commitments.

A plural subject G_1 at t_1 is J-related to plural subject G_2 at t_2 when G_2's token joint commitment is inherited from G_1's token joint commitment. The Pixies of 2003 are the Pixies of 1993 because the 2003 Pixies inherited their joint commitment from the 1993 Pixies, and are thereby J-related to the 1993 Pixies. The 2015 PGSA is *not* the 1995 PGSA, however, because although the 2015 PGSA's joint commitment has the same content as the 1995 PGSA's (namely, the commitment to realize the structure defined by C), they do not inherit their commitment from the 1995 PGSA. Thus, the 2015 PGSA and 1995 PGSA are not J-related. The lack of a J-relation explains why some groups seem not to be continuations of an earlier group (as in *Student Association*), while the presence of a J-relation explains why some groups *are* continuations.

However, the notion of J-relatedness relies on the notion of "inheritance," which needs definition. To capture the intuitive idea, consider the following case:

> *Family Heirloom*: The Smith family has an antique pocket watch, "Goldie," that passes down from father to oldest son. When John Jr. comes of age, John Sr. gives him Goldie. John Jr., in a moment of carelessness, loses Goldie on a walk downtown, where Adam Jones picks it up. When Jones's son, Adam Jr., comes of age, he gives his son Goldie, inaugurating a tradition of passing Goldie down from generation to generation.

There is a sense in which the Smiths and the Joneses participate in the "same" tradition: that of passing Goldie down from father to eldest son. It would be a

mistake, however, to say that Adam and Adam Jr. are carrying on the *Smith* family tradition. Engaging in the same activity (here, passing down Goldie, that particular watch, from father to eldest son) does not suffice for carrying on the tradition. A further feature is needed: the right relation between who passes the torch and who receives it. In *Family Heirloom*, John Jr. does not bequeath Adam with Goldie in the way required for Adam to carry on the tradition. When John Jr. permanently loses Goldie, the Smith family tradition is lost.

Inheriting a joint commitment is like inheriting a pocket watch in the family tradition. The tradition's continuation relies on more than the identity of the particular item being passed down; it also depends on the right relations obtaining between the people involved. Likewise, a joint commitment's continuation relies on more than preserving the commitment's content, like the commitment to jointly realizing C; the right relation has to obtain between the people who pass down and inherit the commitment. One cannot be included in the tradition unless one is inducted into it by people who are involved themselves.

Of course, inheriting a joint commitment differs from inheriting, say, a pocket watch, in significant and important ways. One can inherit a pocket watch by having it bequeathed to one in a will. The executor of John's estate can pass on Goldie to John Jr., rather than John himself, and this could be so even if John Jr. were entirely ignorant of Goldie's existence prior to his father's death. This is not so with joint commitments; they cannot be bequeathed in this way. John, Paul, George, and Ringo could not bequeath the joint commitment to performing as The Beatles to their children, Sean, Stella, Dhani, and Zak, merely by "giving it" to them, against their

children's wishes. The difference between these two senses of "inheritance" lies in the relations between the parties concerned. Inheritance, in the simple bequeathing sense, involves no particular uptake conditions, or at least very loose conditions, by the receiving party. The transfer of property rights can take place regardless of the recipient's input. The inheritance of a joint commitment, however, involves the same stringent conditions as the initiation of a commitment. Inheriting a joint commitment involves undertaking the commitment oneself, giving to oneself and others a directive that one also receives from others. No such exchange occurs in ordinary inheritance, in the bequeathing sense. For this reason, when I discuss "inheritance" in the J-relatedness sense, it should be interpreted as a term of art, much like "joint commitment" itself.

Whatever the inheritance (*qua* J-relatedness) of a joint commitment consists in, it must account for reconstitution facts, and the possibility of a group replacing its members completely over time. That is, it should explain why the 1995 PGSA did not survive in 2015, and why the 1995 PGSA counts as a continuation of the same student organization from 1990, even though no member in 1995 was a member in 1990.

To start, it helps to see how we individuate *token* joint commitments. A joint commitment *type* is individuated by its content; what distinguishes a joint commitment to X-ing from a joint commitment to Y-ing (supposing X-ing is not Y-ing) is the difference in the activities they commit people to doing. Joint commitment *tokens* are individuated not only by their content, but by who makes the commitment. Alicia and Bill's joint commitment to go to the store is the same type as Cassie and

David's joint commitment, but Alicia and Bill's joint commitment is different from Cassie and David's, because Alicia and Bill make the commitment to themselves and each other alone, whereas Cassie and David make *theirs* to themselves and each other alone. Joint commitment tokens are individuated not only by their content, but by their authors as well.

Since joint commitments are directive-like representations, it helps to compare joint commitment tokens to tokens of ordinary directives. Suppose Alicia directs Bill at t_1 to buy green beans; Cassie likewise directs David at t_2 to buy green beans. These directives might be the same directive type, by virtue of having the same content (namely, "buy green beans!"), but they are not the same token, because they are made at different times, and they are directed by different speakers to distinct hearers. Joint commitments are likewise individuated by when they are made and by who makes them, as well as their content.

However, unlike ordinary directive tokens (like Alicia's directive to Bill), joint commitments involve a more complicated process of production and consumption: each party produces a directive to be consumed by all parties involved (including themselves), and each party also consumes those directives produced by all parties involved. In this case, there are not multiple joint commitment tokens, where each party involved tokens the commitment. That is, it is *not* the case that, in a two-party joint commitment, there are actually two commitments tokened. Instead, the joint commitment is tokened once each party produces and consumes the same content from themselves and each other. That is, there is no joint commitment token until all the parties involved have taken up the directives from themselves and each

71

other. In this way, the tokening of a joint commitment is like the tokening of a handshake. An act does not count as a handshake token when one party extends their hand to another. No handshake occurs (it is not properly tokened) until the other party accepts the invitation and they shake hands. Likewise, individually-produced directives do not themselves token joint commitments; the tokening of a joint commitment is in the reciprocal production and uptake between all parties involved.

However, people can join or leave plural subjects – membership is not fixed – so the persistence of a joint commitment token across time must account for this. To see how it might, consider the following:

> *Walk 5*: Alice and Beatrice regularly go on walks together, initiated by an invitation Alice extended to Beatrice long ago. One day, they ask Carl if he'd like to join, and he happily accepts. The three of them walk together regularly, and when Alice regrettably moves away, Beatrice and Carl continue walking together.

Presumably, the joint commitment that moves Beatrice and Carl to walk together after Alice moves away is a continuation of the commitment that Alice and Beatrice had (and, later, the one that Alice, Beatrice, and Carl had). That is, the joint commitment token that moves Beatrice and Carl when Alice moves away is a continuation of the joint commitment tokened first by Alice and Beatrice. People make a joint commitment when they issue a directive-like representation to themselves and each other; that commitment is still in force so long as it has not been rescinded

or otherwise given up. So, the commitment can be extended when jointly committed people re-issue the directives which committed them to someone else, bringing a new member into the plural subject. Likewise, someone can leave a plural subject when the commitment is partially rescinded for them, and they are released from their obligations to the group. Thus, in *Walk 5*, Alice and Beatrice jointly commit to going for walks; they take on the obligations to one another by virtue of their shared commitment. Call this commitment W, the commitment made by Alice and Beatrice to themselves and each other. W explains their regular activity together, and generates the norms that regulate their shared activity.

When Carl joins, the token commitment that governs all three's shared activity, W', is a continuation of W. This does not mean that W' is necessarily the "same token" as W; indeed, it cannot be, if tokens are individuated by their authors and when they are made. W is tokened first by Alice and Beatrice, and W' is tokened later by Alice, Beatrice, and Carl. We would not say, though, that, when Alice and Beatrice bring Carl in, they effectively dissolve W and replace it with the nearly-identical W'. What makes W' a continuation or successor of W is that the tokening of W' bears the right causal history with the tokening of W. When Alice, Beatrice, and Carl token W', they do so because of Alice and Beatrice's tokening of W. In this instance (*Walk 5*), in order for W' to have the right causal history, the proper inheritance of a joint commitment, the authors of W' must overlap partially with W.

To further drive home the point that token-continuity is not token-identity, we can consider cases where there is no overlap between authors at one stage and at a later stage:

Walk 6: Alice and Beatrice walk together by virtue of their commitment W. They bring Carl into their shared activity, tokening W', a continuation of W. Alice moves away, leaving Beatrice and Carl to walk together by virtue of a new token, W''. After a time, they find a new third partner, Don, thereby tokening W''', and later, Beatrice leaves too, leaving Carl and Don as walking partners, tokening the commitment W''''.

In *Walk 6*, we see that W' is a continuation of W; W'' is a continuation of W'; and so on, until W'''' is a continuation of W'''. The authors of W'''', Carl and Don, are not the authors of W. Even so, the right causal connections obtain between different steps in the series that lead from W to W''''. We can now see how the inheritance of a joint commitment works: a plural subject G_1 inherits its joint commitment X_1 from plural subject G_2 when X_1 is a continuation of G_2's commitment X_2 by virtue of the right causal connections obtaining between X_1's tokening and X_2's tokening. Thus, a plural subject G_1 is J-related to G_2 when G_1 inherits its joint commitment from G_2. Again, a parallel with personal identity is instructive here. J-relatedness bears a similarity with Parfit's Relation R, where "R is psychological connectedness and/or psychological continuity, with the right kind of cause" (Parfit, 1984, p.262). Inheriting a joint commitment from an earlier plural subject is like inheriting memories and the like *via* Relation R to an earlier person. The role played by psychological continuity in some theories of personal identity is analogous to the role played here by inheritance and J-relatedness.

The test of this view is in how well it explains reconstitution cases. First, consider *Hiatus* and *Breakup*. Intuitively, there is a difference between the cases, and this

view says that the difference lies in how strongly the earlier and later versions of the bands are J-related. Sleater-Kinney's joint commitment is never dissolved, so when they start playing together again, they are resuming activity by virtue of their old commitment. With the Pixies, the joint commitment does dissolve in 1993, but it is J-related to their commitment in 2003 in a very attenuated way, in that those people in 2003 would not be so committed unless they had been committed in 1993. The 2003 Pixies inherit the 1993 Pixies' commitment in an unusual way, but it is inheritance all the same, helped by the fact that the authors of the earlier token and the later token are identical. This is also why the total replacement of members in *Student Association* counts against J-relatedness between the 1995 PGSA and the 2015 PGSA. There is no overlap between 1995 and 2015 in the authors of the joint commitment tokens. Even though they are both committed to realizing C, the 1995 token lacks the right causal connections to the 2015 token.

The Narrow Group Question asks about what distinguishes one group from another; answers to it tell us when we have one group instead of many. The Broad Group Question, by contrast, asks about the deep metaphysical nature of groups; what kind of thing can be a group at one stage of its career, and a non-group at another? I suggest that what brings groups into existence are the makings of joint commitments; that groups continue to exist, to expand or contract in membership, when different stages are J-related; that they go out of existence when the commitments are dissolved.

2.5 The ontological innocence of plural subjects

Suppose that groups are jointly committed people, and that they are individuated by the structures they commit to realizing, like Ritchie claims. The various cases considered suggest that group identity over time largely consists in chains of strong J-relatedness. However, as *Breakup 2* and *Breakup 3* show, J-relatedness can take a branching form. Consider the former case: A is J-related to the Pixies, and B is J-related to the Pixies. Both are equally J-related to the Pixies. They cannot *both* be the Pixies, though, since, by transitivity of identity, they would be identical with each other, and they are not. This means that, while J-relatedness might be necessary for identity over time, it is not sufficient.

One concern about a branching form of the J-relation in *Breakup 2*, or the J-relation between the full band and mostly replaced band in *Breakup 3*, is that the successors do not really inherit their commitments from the initial band at all, but in fact do form entirely new commitments. These new commitments do bear an important historical connection to the one which held together the 1993 Pixies, but the kinds of normative structures generated by these commitments show that they are replacements, but not proper successors to the 1993 Pixies' commitment. One way we individuate commitments, after all, is by looking at *who is committed* to them. When the Pixies "branch off" into Frank and Joey's band A, and Kim and David's band B, Frank and Joey become accountable only to each other, and likewise for Kim and David. As Pixies, all four were accountable to one another, but as A and B, the A members are only accountable to A members, and likewise for B. One concern, then, is that the Pixies' commitment does not branch, but simply dissolves

entirely. Once all four lose all expectations about one another, the commitment is gone, especially in an acrimonious breakup. The same goes for *Breakup 3*: the band does not survive the sudden departure of half its members, but simply dissolves. The band that forms from the remainder is not a proper J-related successor, but a new band. In response, consider the following case, which seems less controversial than *Breakup 2* and 3:

> *Game Night*: A group of students, G, gets together two nights per week to play games: Monopoly on Monday, poker on Friday. Attendance on both nights is expected, barring special circumstances. Later, however, this expectation becomes untenable, as the school year imposes more difficulties on their schedules; moreover, there are new people would like to join in on playing Monopoly or poker with G, but cannot go to both nights. So G decides to continue on with Monday and Friday being separate activities. G_M is the group that plays Monday Monopoly; G_F plays Friday poker. Over time, G_M and G_F eventually have entirely disjoint memberships, and nobody in either group knows that they share a history with G.

While *Breakup 2* and *Breakup 3* arguably present cases where the right J-relation does not obtain between the earlier and later commitments, *Game Night* lacks the enmity and lapse in all normative expectations which dissolves the Pixies' initial commitment. G's commitment does not dissolve, but rather seems to undergo a revision that preserves it enough to be J-related to G_M's commitment and G_F's commitment. If *Game Night* shows that two groups, G_M and G_F, can be appropriately J-related

to an earlier group G, this shows that J-relatedness might be necessary, but cannot suffice, for identity over time.[13]

We might follow Parfit in saying that, in branching cases, it might be an empty question whether, say, A or B are *really* the Pixies. More strongly, he would say that, if we asked, in the *Breakup* case, whether the 2003 Pixies are really the *same band* as the 1993 Pixies, we would be asking an empty question, and that there is no fact of the matter or determinate answer to it. While this might be too strong a verdict for *Breakup* or *Reunion*, it is more plausible for cases like *Breakup 2* or *Breakup 3*. Following Parfit's reasoning, we would be reductionists about groups. In some cases, there is no determinate answer about whether a group has survived some adventure, like a breakup. Once you know all the facts about who is J-related to who in a given case, you have all the information you need to know. No further fact about the existence of a group is needed. As Parfit might put it, to ask about the Pixies of 2003 whether they are the same band as the 1993 Pixies, or merely a cover band with exactly the same members, is to ask an empty question.

This seems to be at odds with statements that mention groups, like, "Sleater-Kinney played Lollapalooza in 2006." If we say that a band played a particular gig, or that a poker game meets every Tuesday, we seem to make an ontological commitment to groups, to plural subjects as *bona fide* entities. This is the reification worry, that an ontological commitment to group entities comes from true statements about *them* which do not seem to be about the individuals. We face an inconsistent triad: (1) Something is X-ing, and (2) no single individual is X-ing, yet (3) there is

---

[13]Thanks to Don Baxter for pressing me on this point.

nothing there but the individuals.[14] To maintain the strategy here, I will use the resources of plural quantification to show that talking about groups is ontologically innocent, thus avoiding the reification worry. There is no need to worry about groups as entities distinct from their members. To see why, we will turn to the mechanics of plural predication and quantification.

Quantifiers sometimes range over singular variables, as in, "At least one crocodile is albino," which has the logical form "There is some x such that x is a crocodile and x is albino." But there are also plural variables, as in, "There are some crocodiles that are albino," which has the logical form "There are some xs such that the xs are crocodiles and the xs are albino." [15] [16]

A rough and ready analysis of "There is some x such that x is a crocodile and x is albino" treats the sentence as a long disjunction. Suppose our universe of discourse contains one dog (named Dog) and four crocodiles (named a, b, c, and d), two of which (c and d) are albino. The disjunction yielded by such an analysis then looks like:

[Dog is a crocodile and Dog is albino] or [a is a crocodile and a is albino] or [b is a crocodile and b is albino] or [c is a crocodile and c is albino] or [d is a crocodile and d is albino].

This is true, since two of the disjuncts are true, so "At least one crocodile is albino" is true. We determine the truth of the sentence by seeing whether the

---

[14]I owe this inconsistent triad formulation of the problem to Donald Baxter.

[15]My exposition of plural quantification here comes mainly from (Linnebo, 2012), but also from (McKay, 2006).

[16]Here I use Yi's (Yi, 2014) preferred notation for plural variables, e.g. "the xs," as opposed to "xx" or "X".

predicates "...is a crocodile" and "...is albino" both apply to at least one value of the variable.

Now consider: "The crocodiles in the zoo weigh eight tons." Suppose there are four ordinary crocodiles in the zoo, and each individual crocodile weighs two tons. The truth value of the sentence depends on how we read "...weighs eight tons." If we determine the truth value as above by looking at each individual crocodile and asking whether it weighs eight tons, the sentence is false, since each crocodile weighs two tons. On the other hand, if we determine the truth value by looking at all of the crocodiles together and asking whether *they* are eight tons, the sentence is true, because their weight added up is eight tons.

We learn from this example that for them to weigh eight tons is not always a matter of every individual in the plurality weighing eight tons. That is, "...weighs eight tons" is not always a distributive predicate; predicating weight of a plurality does not mean automatically that every individual in the plurality has some particular weight. Some predicates are distributive, like "...is a crocodile". If there are crocodiles in the reptile house, the plurality is not itself a crocodile. Some predicates cannot distribute, like "...surround the - - -". If the crocodiles surround the pond in their enclosure, that does not mean each crocodile individually surrounds the pond, because surrounding is something done by a plurality.[17] So there are distributive predicates (like "...is a crocodile"), non-distributive predicates (like "...surround the - - -"), and predicates that can be either, depending on context ("...weigh - - -") (Linnebo, 2012, section 1.1).

---

[17] For the most part, anyway. One long piece of rope can be laid on the ground to encircle, and thereby surround, the enclosure, but things like crocodiles cannot do that.

80

A further question arises about non-distributive predicates. Take the predicate "...form a queue." It looks like only pluralities can satisfy that predicate, since one person cannot form a queue. But does this really mean that forming a queue requires pluralities? Consider the following (rough and ready) analysis of "...form a queue":

> The xs form a queue $=_{df}$ there is an x_1 and there is an x_2 and ...and there is an x_n, such that x_1 stands in front of x_2, ..., x_{n-1} stands in front of x_n.

This definition tells us under what conditions some individuals can come to stand in such a way as to form a queue. So maybe we can analyze away these non-distributive predicates by giving the truth conditions of sentences involving something like ordinary distributive predicates and singular quantification. That way, we only need to deal with non-distributive predicates as paraphrases of longer predicates involving ordinary distributive predicates and singular variables. We would effectively be reducing the non-distributive to the distributive, so that, for example, these xs are F (non-distributively) if and only if this x is G (distributively), this other x is G, and so on.

Unfortunately, these approaches lack the expressibility of non-distributive predication. For example, one of such approach uses Davidsonian event semantics (McKay, 2006). For this approach, to say that Emily, Amy, and Sally form a queue is to say that there is an event, a queue formation, such that Emily is an agent, Amy is an agent, and Sally is an agent. But as McKay points out (McKay, 2006, p.242), for sentences that are ambiguous between distributive and non-distributive predication, this first pass cannot satisfactorily handle the ambiguity. Consider one of McKay's

examples, the sentence "The three students lifted 200 pounds." This sentence contains a scope ambiguity. Either there is a lifting of 200 pounds such that three students are agents, or there are three students such that, for some lifting of 200 pounds, each student is an agent. The first gloss captures the form of the non-distributive reading (some event has three agents: they cooperated in lifting the 200 pounds together), while the second captures the distributive reading (there are three students, and each one is an agent of some event: each student took turns performing a feat of strength). However, distributive readings ought not follow logically from non-distributive readings, and in the case of the three students, the distributive follows from the non-distributive, unless the gloss captures the sense that these agents are the unique agents of the event, and not just some agent or other.

Attempts to build in uniqueness of agency encounter another problem. To say that Emily, Amy, and Sally formed a queue, we say that there is a queueing, Emily, Amy, and Sally are agents of the queueing, and for any other agent of the queueing, it is identical to either Emily, or Amy, or Sally. However, this analysis of "Emily, Amy, and Sally form a queue" is different in logical form from an analysis of "Emily, Amy, Sally, and Jenny form a queue," since the logical form of the latter requires a fourth disjunct in the uniqueness clause. This is to say that "...form a queue" is not one predicate, but many which differ in virtue of how many argument places they have. That becomes problematic when the plurality referenced in the sentence is of indeterminate number, as in "They form a queue." In that case, it is not clear which predicate in the "...form a queue" family is in use. A uniform treatment is preferable to that analysis. Our treatment will take non-distributive predicates to

be irreducible to distributive predicates.

Consider the predicate ". . . are a plural subject" (or perhaps ". . . compose a plural subject"). Individuals cannot be plural subjects; plural subjecthood can only be predicated of pluralities of people. This makes ". . . are a plural subject" a non-distributive predicate; when predicated of a plurality, it is predicated of the plurality as such. So, for instance, "Alice and Beatrice are a plural subject" does not entail a sentence (or sentences) about Alice and Beatrice involving only distributive predicates. Even so, no special ontological commitments to anything beyond Alice and Beatrice are involved in the use of non-distributive predicates. On the non-distributive predicate reading of ". . . are a plural subject," "Alice and Beatrice are a plural subject" is semantically on a par with "The apples are stacked in a pyramid shape." Saying the latter sentence incurs no commitments to anything beyond the apples stacked pyramid-wise; likewise, the former sentence commits one to nothing more than Alice and Beatrice. The same treatment can be given for non-distributive predicates used in describing any plural subject phenomena, so that talking about groups commits one to the group members and nothing more.

One might be concerned that the reification worry arises from purely semantic concerns. We might think that employing plural quantification in this way (following (McKay, 2006)) assumes that the plurality in the subject position of a sentence with a non-distributive predicate is *not* a single object. Thus, the worry goes, the position advocated here presupposes the falsity of singularism, according to which non-distributive predicates only apply to single entities composed by the members of the plurality. If singularism is true, then the subject of "Amy and Bob are class-

mates" is some sort of object composed of Amy and Bob. This condition applies for any non-distributive predicate. It would be a fact of natural language semantics that non-distributive predication requires composites like Amy-and-Bob sums, or a set with only Amy and Bob as its members. This assumption goes back to Frege and "telescoping" sentences(Ben-Yami, 2004). Consider the following sentence pairs:

1a. Goethe and Schilling are poets.

1b. Goethe is a poet and Schilling is a poet.

2a. Larry, Jerry, and Barry are surrounding the fire pit.

2b. Larry is surrounding the fire pit, and Jerry is surrounding the fire pit, and Barry is surrounding the fire pit.

The conjunction in the first pair allows the sentence to "telescope" the two sentences of 1b into the single sentence of 1a. However, the conjunction in the second pair does not telescope, since 2b is ungrammatical. The predicate "...are surrounding the fire pit" requires a plurality in the object place. Frege assumed from this that the reference in 2a must be to an object composed of Larry, Jerry, and Barry. So one might be concerned that this brings back the reification worry. If singularism is true, there must be groups, if we are to talk about them at all.

However, the plural quantification approach advocated here need not presuppose that singularism is false. Singularism is compatible with the ontological innocence of group talk. Indeed, the truth or falsity of singularism is besides the point, so a criticism of the project here that argues from the truth of singularism will not go very far. According to singularism, the ontological commitments of "Alice and

Beatrice are a plural subject" are on a par with "The apples are stacked in a pyramid shape." The proper semantics of these sentences requires that the predicates apply to single objects, like a fusion of Alice and Beatrice, or a fusion of the apples which are stacked pyramid-wise. Mereological fusions of apples might be strange objects, but singularists are willing to countenance them, and if the fusion of apples is not metaphysically problematic for singularists, then the fusion of Alice and Beatrice is not either. But presumably, the singularist critique wants to say more than that Alice and Beatrice must be a single object in order to get the semantics right. Groups somehow be *distinct* from their members (if talking about them is not just talking about the members in a certain way), or are metaphysically strange in some other way. Their distinctiveness or strangeness should be unique to groups for the criticism to stick; *groups*, not merely fusions of people, should be suspect entities. But nothing like that seems to follow from singularism. If fusions of people are no stranger that fusions of apples, then there is nothing distinctively, ontologically profligate about group talk.

2.6   Conclusion, and the Fate of Groups

The Joint Commitment view developed here makes a number of claims. First, however we answer the Narrow Group Question, that answer needs to be supplemented by the theory of joint commitment. But plural subjects seem to be nothing over and above their members, which we can appreciate by reflecting on Parfit-style cases where there seems to be no good answer as to whether the same group has continued

existing. The plural quantification approach, which holds that plural subjects are pluralities of people, gives a semantic backing to what looks like reference to group entities. Thus, we get a sense for what it means to say that groups are no more than jointly committed people.

The project pursued in this chapter is reductionist, broadly speaking. A further question remains, however: do groups have a place in our ontology? Should we say that there are indeed groups, which are nevertheless no more than jointly committed people? Or should we instead say that there are jointly committed people, and there are no groups? How should we decide? What's the difference?

To see the difference between these two positions, consider a tale of two "reductions," familiar from the eliminative materialist literature (see the first section of (Ramsey, 2008)). After some scientific inquiry, we found that water is H2O; the potable stuff that fills the rivers, lakes, and streams turns out to be H2O. This is a successful case of reduction. Now consider the following. In the past, people believed that demonic possession caused seizures. We know better now; seizures actually result from abnormal electrical activity in the brain. It would be a mistake to infer from this that, when we found out that electrical activity in the brain causes seizures, we also learn something about the nature of demonic possession. We do not learn that demonic possession is a disturbance in the brain, in the same way that water is H2O. What we learn is that there is no such thing as demonic possession. The "possession" theory of seizures holds that they result from an immaterial demon manipulating a human being; since nothing like that is actually happening, the possession theory is replaced. The "brain malfunction" theory is not the result of building upon previous

work in the "possession" theory; it is an outright replacement. The ontology of the older theory is given up completely in favor of the new one.

Some reductions, then, are clean. When we learn that water is H2O, we do not have to discard the concept of water. Future work on H2O is not hampered by the continued use of "water." Some reductions, however, require an elimination of the old concepts in favor of the new; the old vocabulary must go. Reductionist theories come in two varieties, then. Ontologically conservative theories do not eliminate the reduced kinds; for example, identifying water with H2O does not require rejecting the concept of water. Ontologically radical theories, on the other hand, do require a rejection of concepts that no longer apply; such is the fate of demonic possession explanations of seizures.

With our options now given, what kind of reductionist theory have I defended in this chapter? Is it ontologically conservative, or ontologically radical? One trouble here is that, in the eliminative materialist literature, there is no widely accepted criterion that tells the difference between the two types of reductionist theories. While there looks to be an intuitive difference between the reduction of water to H2O and the "reduction" of demonic possession to epilepsy, there is no obvious principle that determines where precisely the difference is.

Nevertheless, some features of a reduction might push towards eliminative, ontologically radical reduction. For instance, if there is a deep enough mismatch between how the new theory (the "reducing" theory) understands the phenomena in question, and how the old theory ("reduced" theory) does, that might be a reason to reject the reduced theory outright, including its ontology. Theories would be mismatched

if they disagreed over broad features of the phenomena they both purportedly study. The phlogiston theory of combustion is mismatched with the oxidation theory, because they disagree over the direction of causation in combustion. Phlogiston theory posits that burning materials *release* something, phlogiston, while oxidation theory posits that burning materials *gain* something, oxygen (Bickle, 1997, p. 31). Likewise, the demonic possession theory posits that an external agent's manipulation of a seizure victim explains seizures; brain malfunction theories deny that there is an external agent. The theories disagree fundamentally over broad features of what they each presume to explain. No such disagreement prevents the reduction of water to H2O.

If there is a fundamental mismatch between the commonsense metaphysics of social groups and joint commitment theory, it isn't obvious. Some assumptions do have to be modified or rejected. For instance, we might initially think that groups have determinate identity conditions, like any other concrete particular. I've argued that this is a mistake, but even so, that does not entail that our folk theory of groups is radically false, or that there is such a deep mismatch between the theories that the "folk theory" has to go. For this reason, the reductionist view that I advance here leans towards ontological conservatism. There does not seem to be any compelling reason to be an eliminativist about groups, or to adopt an ontologically radical view. As a result, the reductionism I've defended leaves groups in our ontology.[18]

---

[18]Many thanks to Michael Lynch, Austen Clark, Don Baxter, and Mitch Green for discussion on these points.

Chapter 3

Group Belief

3.1 Explaining behavior

Suppose that, today, I decide to make croissants; I will see to it that, by the end of the day, I pull fresh croissants out of the oven, which I have made from scratch. This is a complicated endeavor, requiring the success of a number of sub-plans. For example, I have to plan on measuring out the right kinds of flour for the dough, and plan on cutting up sheets of parchment paper for rolling out the butter block. To achieve flaky, buttery goodness, I need the right intentions, goals, and desires. But in addition to having the right outcomes in mind, I need to have answers to some basic questions. For example, where is the flour, and what kind do I need? If I don't have any, where can I get it? Is my rolling pin up to the task? (And where did I leave it?) Without answers to these questions, I can't even take the first step towards making croissants. I have decided what is to be done, but without some beliefs about how

those goals are to be achieved (among other things), nothing will get going.

My croissant adventure illustrates a familiar understanding of behavior. Suppose you see me pulling the pan of croissants out of the oven, a pleased smile on my face. Why am I doing that? I directed all my activity towards this moment because I wanted to make my own croissants, and I believed that the steps I took along the way, the resources I used, would be instrumental towards that end. The explanation of my baking activities relies on a few basic conceptual tools: conative states (like desires, intentions, commitments, etc.), and cognitive states (beliefs, acceptances, etc.). With these tools in hand, you can understand my actions.

The same considerations extend to joint actions. If I want to make croissants with Jourdan, rather than making them alone, then we need to make a joint commitment to baking croissants. While this is a central part of the story, we need more than our joined wills to carry out our project. Just as I needed the right kind of beliefs to make croissants alone, we need the right kind of beliefs to make croissants together. If our joint commitment is to be effective in bringing delicious croissants out of the oven, we need a shared pool of information to act on, or else our project won't happen. Without beliefs to inform the desires, we're at a standstill.

In previous chapters, I've defended the directive model of Gilbertian joint commitment, which explains how people act together. If the belief-desire model is right, then explaining the behavior of groups needs more than joint commitment; it needs group belief as well. But what is group belief? What does it look like? This chapter provides (1) the rationale for taking group belief seriously, (2) arguments that summative accounts fail, (3) the details of a positive account, (4) problems for such an

90

account, and (5) responses to these worries.

3.2   Why group belief?

Familiar philosophical practice proceeds along the following lines. First, we notice
the pervasiveness of some feature of a very important mode of discourse, usually (but
not limited to) everyday talk. For example, we often employ the concept of joint
action, which supposedly picks out actions distinct from actions performed merely
in concert. Next, we start to critically examine that feature of discourse. We might
ask, for example, "If there is a difference between joint action and merely concurrent
action, what is it?" Presumably, the question is pressing because intelligent use of
the concept requires that we have some idea of what we're talking about, and thus,
some idea (to be made precise) of where the differences are. With some philosophical
labor, we get a satisfactory theory that accounts for how the concept feature in its
home discourse. Or perhaps we fall into aporia. Either way, that's how philosophical
procedure can play out.

This is our situation with respect to joint action, which joint commitment views
are supposed to explain. Group belief has a similar place in everyday talk, and
given the pervasiveness of the concept, it needs to be explained. How pervasive is it,
exactly? It is invoked often enough that many philosophers have commented on it
(Clark, 1994; Tollefsen, 2015a; Quinton, 1976).

Quinton defends summativism, or the view that belief attributions to groups are
indirect ways of attributing beliefs to the members, but he makes two claims about

the sense in which the attributions to groups indirectly take us to attributions to the relevant individuals. On the one hand, he says that attributions of beliefs to groups are "plainly metaphorical." This would be to say that group beliefs need not be literally true, or that such attributions need not be taken at face value. We interpret statements like "The Cubs believed they would win the World Series" in the same way we interpret statements like "John was over the moon about the Cubs victory." With the latter statement, we're not meant to think of John's position relative to the moon, but instead about John's emotional response to the Cubs' victory. Likewise, when it comes to statements about what the Cubs believe, the belief of "the Cubs" is not what's at issue, but rather what the members of the Cubs organization believe. The literal truth of the matter is unimportant.

On the other hand, Quinton also suggests that group belief attributions just are attributions of beliefs to group members, or to authorized members. This summative reduction view states that there is a straightforward logical entailment from state-ments about what a group believes to statements about what the members believe. This means that we can take group belief attributions at face value; they can be true in virtue of what they literally mean. This makes the summative reduction view dis-tinct from the summative metaphor view. Strictly speaking, then, there are distinct summative views in Quinton: the summative metaphor view, and the summative reduction view. Since the views are logically independent, they need to be addressed separately.

How well does the summative metaphor view fare? If there is a robust parity between our practices with respect to individuals and with respect to groups, then

there is some reason to treat them alike semantically, i.e. with respect to how to interpret these attributions. There seems to be parity between how we attribute beliefs to individuals, and how we do with groups (see (Tollefsen, 2002) and chapter 5 of (2015b)). Consider how belief attribution fits into our broader understanding of human behavior, e.g. my frantic croissant rituals. An outside observer might wonder a few things, such as, Why is he weighing cake flour? That observer might use their "mindreading" abilities to answer that question, thinking, "He must believe that cake flour is important for his dough; he will probably add that to the mixing bowl soon." The explanation of my behavior works through reference to my beliefs and desires. When we use mindreading in this way to explain something's behavior, we are taking what Dennett (1991) calls the "intentional stance" towards it.

To understand the intentional stance, it helps to contrast it with two other "stances" one might take: the physical stance, and the design stance. We take the physical stance towards something when we only use the laws of physics to understand what it does. To explain and predict the behavior of dough rising in the oven, we need only look to the basic physical facts of the matter: the chemical composition of the dough, how it has been layered, the temperature of the oven, and so on. For some explanations, the physical stance is too cumbersome. For example, you might want to know why, when certain areas on the front of the oven are pressed, the configuration of LEDs changes and the oven's temperature rises. An explanation in terms only of circuits and gases would leave out something significant: that the oven was designed to make its internal temperature correspond to the numbers displayed on its face. That explanation is only available from the design stance, where we can

avail ourselves of concepts of design and function. We do not merely use the design stance for artifacts, however; we take the design stance towards biological functions as well, like phototropic mechanisms in plants.

The intentional stance goes further than both the design and physical stances. When taking the intentional stance, we use mental verbs to make sense of something's behavior. We treat the thing in question as an agent, and do our best to rationalize its behavior by attributing mental states or processes to it. This is what we do when examining my flour-weighing. In taking the intentional stance towards me, you assume that I'm a rational agent, and attribute to me the beliefs and desires that would make my behavior rational. The intentional stance can be taken towards non-agents; for example, we might say that a plant growing in the direction of the sun is "trying" to get more sunlight, or that the croissants in the oven "refuse" to rise. These uses of the intentional stance are superfluous, however, since explanations from the design or physical stance suffice. We can explain the flower's phototropism from the design and physical stances, and the rising dough from the physical stance alone. There are no further details to be gleaned from the intentional stance, so mental state attributions here look metaphorical.

Suppose we have something, X, which behaves in certain ways that we wish to explain. Taking the physical stance only gets us so far; likewise for the design stance. When we take the intentional stance towards X, however, we get much further in predicting what X will do next. Given that the use of the intentional stance is vindicated by the success of explanations made from it, the use in explaining X's behavior here seems warranted. If X were an individual human, there would be

94

no issue with taking the intentional stance here. Suppose X is actually a group of humans, however. Would belief and desire attributions from the intentional stance be mere metaphor? It depends on the group and the behavior, but if taking the intentional stance towards the group really yields better explanations (unlike saying that the flower is "trying" to find sunlight), then why insist that belief attributions to groups are merely metaphorical?

A critic can respond here on Quinton's behalf in the following way. Belief attributions to groups are not mere metaphors; the summative metaphor thesis is false. However, group belief attributions are parasitic on belief attributions to the group members. Taking the intentional stance towards groups is a shortcut to taking it towards the members. If we knew everything that the members believed, we would know what the group believes. Group belief attributions are no more informative than attributions to the members.

The critic here is endorsing the summative reduction thesis, or what I will hereafter call "summativism." How does it fare?

3.3 Summativism

According to summativism, a group G believes p if and only if most of the members believe p. The sum of G's members' beliefs determines what G believes. If the DNC believes that Russia influenced the election, then the majority of DNC members believe that Russia influenced the election. The converse holds as well: if most of the DNC's members believe that Russia influenced the election, then that's what the

DNC believes. The sum of the members' beliefs that p is a necessary and sufficient condition for the group believing p.

Summativism has its advantages. It is parsimonious. To understand group belief, we only need to draw on the familiar resources of ordinary, individual belief. Nothing else is needed, save for some sort of summing function that yields the group belief. It also respects the intuition that group mental states are determined by individual mental states. Put another way (though not synonymously; see (Epstein, 2015), especially chapter 8), group states supervene on individual states. Relatedly, it respects the intuition that the contents of individual states restrict the contents of group states. For example, a group cannot believe p if the group's members have never even considered whether p.

Note, however, that summativism makes claims that are logically stronger than these intuitions. Not only do individual mental states provide the supervenience base for group states, their sum logically entails what the group believes. Further, not just any individual propositional attitude determines what the group believes. It is the beliefs of the individuals that determine what the group believes. This means that, if most of the group's members only tentatively accept p, then the group cannot have full belief that p. The group can, at most, tentatively accept p as well. For summativism, not only do member state contents restrict the contents of the group states, but the member propositional attitudes do as well. Only the sum of the group members' individual beliefs can determine what the group believes. No other behavior or state will do.

Gilbert (in (Gilbert, 1989), but also (Gilbert, 1987), among others) provides

criticisms of summativism's two theses: (1) that the sum of the members' beliefs that p are sufficient for the group believing p, and (2) that the sum is necessary for said group belief.

First, if a majority of the members of a group G believe p, does G itself believe p? Suppose that every member of the Philosophy Graduate Student Association believes that it's dangerous to go rock-climbing on sandstone. However, none of the PGSA members ever discuss this belief, nor do they consider whether anyone else believes it. Even so, if most members of the PGSA believe something, then the PGSA believes it, according to summativism, so that the PGSA would indeed believe that climbing sandstone cliffs is dangerous. Intuitively, though, the PGSA does not believe that, so a group's majority believing p cannot be sufficient for the group to believe p. To put additional pressure on the sufficiency claim, suppose that each member of the rock climbing club believes that it's dangerous to climb sandstone. If member belief is sufficient for group belief, then the club believes that sandstone climbing is dangerous. But if the rock climbing club and the PGSA have exactly the same members, then that entails (according to the sufficiency claim) that the PGSA also believes that sandstone climbing is dangerous, since each member believes it. This is a difficult consequence to accept, especially if there is never any discussion of rock-climbing in PGSA meetings; we would be more inclined to say that the PGSA has no opinion either way. So the majority of G's members believing p does not suffice for G believing p; the sufficiency claim is false.

Second, if G believes that p, must the majority of members believe p? Gilbert and Priest (Gilbert & Priest, 2013) imagine a University Event committee meeting to

select a speaker for the University's 50th anniversary. Phyllis, a committee member, makes a suggestion: "J.K. Rowling would be a great speaker." After a discussion where support for the suggestion outnumbers objections, another member chimes in: "So we agree that Rowling would make a great speaker?" There's some silence, some voiced agreement, and no objections. As a result, committee members could say things like, "We think Rowling would be a great speaker," "We decided that Rowling would be a great speaker," "Our view is that...", etc. Such self-ascriptions would be appropriate. Even so, it is possible for most committee members to disagree or be apathetic, while standing by the committee's position as the committee's position. This could happen in a number of ways, but here's one possibility. Imagine that Phyllis believes Rowling to be a poor public speaker, and wishes to see her fail in public. She makes the suggestion in the hope that the committee will decide she would be a great speaker. Phyllis makes her suggestion, and the other committee members defer to her, having no personal opinion on the issue. Still other members agree, wishing to avoid conflict. As a result, the Event committee comes to think that Rowling would be a great speaker, even though no member of the committee believes it. Thus, a majority opinion among G's members that p is not necessary for G to believe p.

3.4 Re-grouping and moving forward

Summativism is an attractive, but ultimately unsuccessful, theory of group belief. It only accounts for a very narrow range of cases, where consensus among members

seems to determine what the group believes, or where a lack of consensus seems to determine the group's lack of an opinion. Other cases, however, prove difficult to explain for summativism. To account for those cases, we need a non-summativist account of group belief. Before articulating such an account, it would help to collect the philosophical data which the non-summativist theory should explain.

For starters, the non-summativist theory should capture those features of summativism that made it an attractive theory. Summativism gives a parsimonious theory, in that explaining group belief requires no novel explanatory devices distinct from ordinary belief. It also accounts for related intuitions about how group beliefs are "built from" what the group members do. First, the activities and/or states of the group members fully determine the group's beliefs. (Alternatively, the group activities/states fully depend on what the members do; or, the group activities/states supervene on the members' activities/states.) Second, the content of any of G's contentful activities/states is restricted by the content of activities/states of G's members. This means that G cannot have an opinion about whether p if G's members have never discussed whether p, or discussed anything from which p might readily be inferred. This leaves open the possibility that a group could have implicit beliefs that have never been discussed openly in the group. Next, a non-summativist account should explain how the counterexamples to summativism are possible. It ought to be a virtue of the non-summativist view that it can explain why the rock climbing club and the PGSA have different views, despite having exactly the same members, and why the Event committee can believe something that none of its members seem to believe. If non-summativism can't explain these, it's no better off than

its summativist rivals.

There are a number of non-summativist positions in the literature (for instance, (Pettit, 2003; Tuomela, 1992)), but I will focus on the view articulated and defended across several papers and books by Margaret Gilbert. Unsurprisingly, the core of her non-summativism is joint commitment. To state Gilbert's view plainly, a group G believes that p if and only if G's members are jointly committed to emulating as far as possible a single body that believes p. In other places, she says that G believes p if and only if G's members are jointly committed to believing p as one body. There's a lot here to unpack; let's start with joint commitment.

I've developed my own view of joint commitment in the first chapter, and it will serve our purposes here without contradicting Gilbert. But why appeal to joint commitment here? What purpose does joint commitment serve in explaining cases of group belief?

Consider the Event committee case again, particularly the point at which we say that they have settled on the belief that J.K. Rowling would be a great speaker. Now suppose that later in the meeting, Dave, without preamble, offers up the following: "J.K. Rowling would actually be a bad speaker." Gilbert predicts, plausibly, that the Events committee would react to Dave's assertion with stony silence. Someone could reasonably respond, "But we thought she would be great!" They could say this, even if everyone personally believes otherwise. This would constitute a rebuke directed at Dave, not for contradicting the majority opinion (i.e. a personal belief held unanimously by the group members), but for contradicting the group's attitude. This is one of the features of group belief, that it has, as Gilbert puts it (following

Durkheim), "coercive force." In other words, when a plural subject G believes that p, that belief generates for G's membership the standing to rebuke one another for speaking against G's attitudes. Perhaps Dave could have gotten away with qualifying his statement by flagging it as merely his personal opinion, but without such a preface, he invites a reminder from the group to get in line.

This is where the appeal to joint commitment helps. Remember that joint commitments generate obligations among the jointly committed people. When Jourdan and I jointly commit to making croissants, I issue a directive to myself and to her (and Jourdan does likewise), and I accept the directive from myself and her (and Jourdan does likewise). Thus, we each owe compliance to ourselves and to each other. If I wander away from the kitchen to check my phone, Jourdan has the normative standing to get me back to work with a gentle reprimand. This is the case when we jointly commit to making croissants, to taking a walk, and even to forming a group attitude. The standing to rebuke generated by a group belief is the same normative standing formed in any joint commitment. It just so happens that what Dave fails to comply with is his obligation to act as a mouthpiece for the group opinion. So joint commitment plays an important role in understanding group belief.

The Event committee story portrays a tidy instance of a group coming to believe something. Somebody proposes p, it goes unchallenged, and the group thereby believes p, as if by a vote. This story might suggest that group beliefs are largely decided by overt tallying out in the open between group members. However, a group's belief can often "go without saying," as Gilbert puts it. To see how, think about a non-belief case, a standard joint commitment to a shared activity. Suppose Mary

is heading out to lunch with Georgina, and John sees them and says, "Hold on, I just need to pack up my things." Mary and Georgina wait, and when John comes to them, all three go out to lunch. When Georgina and Mary wait for John, they have formed an implicit agreement. Georgina and Mary didn't need to stop and ask John, "Shall we go to lunch?" His statement and their waiting is the joint commitment; it is implicit in how they behave. Further, Gilbert points out that joint commitments can build over time, and it might be that there's no determinate time at which the joint commitment comes about (Gilbert, 2000, pp.5-7).

Likewise, joint commitments to believe something "as one body" can be adopted implicitly. The content of the commitments can come through in the intra-group behavior of the group members. A particular kind of behavior is what Gilbert calls the "shocked surprise response" (Gilbert, 1994, p.236). Suppose we start discussing global poverty, and the best ways to remedy it. Someone suggests cheaper, more effective birth control; yet another suggests more direct relief efforts; Bob suggests building more schools. As we talk about the merits of each option, Bob makes an off-color joke about the poor. Stunned silence ensues, since we seemed to be on the same page with regards to people in poverty, and Bob has seemingly defected from that agreement. Gilbert explains the shocked surprise by appealing to joint acceptance. We had all jointly accepted certain things (e.g. that global poverty is a serious issue), and Bob has failed in his duties as someone who was part of that joint commitment.

It need not be the case that they have explicitly, openly stated that we should not make jokes at the expense of the poor. Our attitude towards that particular

102

content might come through in our tones of voice. We could be having this conversation after seeing a documentary about global poverty. The shocked surprise doesn't require any univocal agreement on some particular content. Bob can be met with the stunned silence by saying something contrary to some attitude implicit in the group's behavior, which can be unarticulated.

Suppose we follow Gilbert here, and grant that there is a joint commitment that underwrites the Event committee's standing to rebuke Dave What is it that they are all jointly committed to doing, such that Dave fails to do his share? When I wander out of the kitchen with my phone, I'm not doing what I ought to do in light of my joint commitment, specifically, the joint commitment to make croissants. The minutes I spend on my phone are minutes I'm not spending on dough lamination, which rightly earns a rebuke to get back to work. What joint commitment is Dave failing to comply with? When the group belief goes right, it permits the Event commitment to report to their superiors at the university by saying things like, "We think that Rowling would be a great speaker." If Dave is not doing his part, he must be acting contrary to whatever joint commitment permits statements like that. He is not failing by personally believing that Rowling would be a bad speaker; as has been noted, his personal feelings are irrelevant to what the group thinks. According to Gilbert, he is failing to emulate with his partners, as far as possible, a single body that believes that Rowling would be a great speaker. When he contradicts the group view, he is not doing his part. As Gilbert and Priest put it, the parties to a plural subject which believes that p will fulfill their joint commitment if "they emulate, by virtue of the several actions and utterances of all combined, a single body that

believes that p. In more familiar terms, they are to act as if they are 'of one mind' with respect to p" (Gilbert & Priest, 2013, p.11).

While cases like this make the case for how personal opinion can come apart from the group opinion, and how joint commitment might be relevant, they do not explain the mechanics of how the joint commitment to emulate a unified, believing body gets established.[1] In the Event committee case, Phyllis makes a proposal, nobody objects, and it becomes the group position that Rowling would make a great speaker. According to what Gilbert and Priest (ibid.) call the Negotation of Collective Belief thesis ("NCB thesis"), parties to paradigmatic conversations "are negotiating the establishment of one or another collective belief on the basis of proposals put forward by one or another interlocutor" (ibid., p.14).

What does this have to do with joint commitment? As the Event committee case shows, the presence of a joint commitment explains the various normative positions of the members of the Event committee. Dave's faux pas and the subsequent reactions from his fellow committee members make sense when understood as Dave flouting his obligations qua plural subject member. His remarks, however, are not made in isolation; they are made in the broader context of an Event committee tasked with scheduling a speaker for the university. The negotiations are embedded within a larger project, a meeting convened for the sake of finding a speaker. This particular

---

[1] It is unclear at times what role the phrase "as one body" is supposed to play in the account, since it seems redundant on its face. However, it seems to do two things. First, it makes clear that the joint commitment to believe p is not a joint commitment to believe p as an individual, but rather a joint commitment to believe p together. Second, since the belief serves to unify *their* actions as a group, so that what they do in light of that belief is done as if by a single body. To say that a group believed p "as one body" tells us how the belief is functioning in the group. Thanks to Austen Clark for discussion on this point.

plural subject, the Event committee, needs a position because it needs to act, just like individual agents do. However, consider one of the examples that Gilbert and Priest use, and the lesson they draw from it.

> Consider two individuals, call them Jenny and Tom, waiting together at a bus stop on a sunny day. The two fall into conversation. "Sure is a beautiful day!" says Tom cheerfully. Jenny, whose actual preference is for cool cloudy weather, replies, "I suppose it is." She speakers pleasantly, but without great conviction.

Two things happen here. First, Tom makes a collective belief proposal by saying that it's a beautiful day. Second, Jenny accepts his collective belief proposal by applying in the affirmative, even though she does not personally believe it's a beautiful day. Of course, Jenny might have done otherwise, as in the following variant:

Tom: "Sure is a beautiful day!"

Jenny (in a tone of doubt): "Mmm"

Tom: "But...it is so bright and sunny!"

Jenny: "You're right: it really is a beautiful day."

In this version, Jenny does not accept the collective belief proposal outright. Tom must first provide a reason for accepting the collective belief proposal; in this case, he suggests that the day's brightness and sunniness is a reason to believe that it's beautiful. The conversational dynamics get more complex from here, as Gilbert and Priest show later in their discussion of the ways of accepting or rejecting proposals,

the distinction between implicit and explicit collective beliefs, and the principles that seem to govern it all.

Setting these details aside for the moment, one might wonder whether chats between strangers at a bus stop as described by Gilbert and Priest could be underwritten by joint commitments, giving rise to the rich normative standings on display in the Event committee case. Two strangers chatting hardly seem to be the paradigm of plural subjects. Moreover, they do not face the pressure to take a position, or form a belief, in the way that the Event committee does. The Event committee needs a collective belief to act on, by virtue of their larger purpose. Tom and Jenny have no such concerns, inasmuch as they really are strangers.

To address these issues, remember that undertaking a joint commitment does not require explicit statements, e.g. "Would you like to have a conversation with me?" The exchange of self- and other-directed directives can be facilitated by pre-existing social conventions. For strangers to shake hands, they must jointly commit to shaking hands. Proposing the joint commitment to a handshake does not always require saying something like, "Would you like to shake hands as a greeting?" One stranger proposes the joint commitment by extending their hand; the other accepts it by shaking the extended hand with theirs. The proposal and acceptance are built into the joint behavior by convention. Likewise, strangers can jointly commit to a conversation when one makes a remark, asks a question, etc., and the other responds in kind. Having a conversation doesn't require an explicit invitation to a joint commitment; in some sense, the opening remark and its response, the "striking up" of the conversation, *is* the joint commitment to the conversation. So Tom and

106

Jenny can be a plural subject by virtue of their joint commitment to converse.

But why do *mere* conversation partners need collective beliefs? With Phyllis and the Event committee, the need to negotiate their position is clear, but less so with Tom and Jenny. There might be a general normative pressure for a conversation to maintain a minimal level of coherence, and this is best understood as coherence among collective beliefs. Consider an example elsewhere:

> Joe meets Karen and, wanting to say something pleasant, comes out with "Lovely day!" Karen, wanting to be agreeable says "Yes, indeed!" Joe and Karen then come across Fred, who grumbles about the day's weather. Karen confidently responds, on behalf of Joe and herself "We think it's a lovely day!" Karen's statement seems to be on target, as a statement of collective belief, irrespective of any personal beliefs of the parties regarding the weather. (Gilbert & Pilchman, 2014, p.195)

Suppose Joe interjected after Karen's report to Fred with, "Oh, is it now?" On the NCB thesis, this should surprise Karen, given that Joe had, as a party to the joint commitment to converse, put forward the proposition that it's a lovely day as a possible collective belief, which Karen accepted. Joe would be flouting his joint commitment with Karen by putting the quality of the weather back into question. This makes it difficult for Karen to report what they believe as a result of their conversation. The beliefs negotiated by the conversation partners need to be coherent, or stable, if reports about the conversation are to be coherent.

3.5 Objection: Is this an account of belief?

The non-summativist view of collective belief holds that a plural subject G believes *p* when they are jointly committed to acting like one body which believes *p*. One might wonder, first, whether this account is circular, since "belief" appears as both *analysans* and *analysandum*. This worry is unfounded, however, since the *analysandum* is not belief *simpliciter*, but *group* belief. Recall the analysis of joint commitment from the previous chapter. It did not analyze joint commitment in terms that presuppose joint commitment; instead, it drew upon the concepts used to define more general notions of commitment. An account of group belief can help itself to the same strategy, using the concept of generic belief to illuminate collective or group belief, without the worry of circularity.

But this presupposes that an account of generic belief, or belief *simpliciter*, will apply cleanly to whatever it is that happens in the Event committee case. It cannot be taken for granted, however, that "belief" is an apt descriptor for what results from the negotiation described by Gilbert and Priest. A full account of belief might show that those states of groups are not beliefs after all, but something else. What else could they be?

Beliefs are supposed to be mental states that aim at truth (see (Williams, 1973; Velleman, 2000)); a belief succeeds or fails depending on whether it is true. According to some who endorse doxastic involuntarism, beliefs only form automatically. For example, if I'm in the kitchen when the oven beeps, I will believe that the oven has beeped, and the formation of this belief is not up to me. I cannot revise or rescind this belief, simply by an act of will. By contrast, I could will myself to wonder

whether the belief is true, to imagine relevant alternatives that I have to rule out. Those mental acts are up to me. I cannot, however, simply make the decision to disbelieve, and have that decision suffice all on its own to make me disbelieve. That is, I might decide to take a pill that will cause me to change my mind, and take the pill, thereby changing my mind. Relatedly, I might "decide" to form an opinion on the weather by looking outside. But this is not changing my mind, or forming a view, merely through an act of will; these are not "direct" means to forming a belief (Alston, 1988a). Also characteristically, beliefs have an affective, emotional component. When I believe that the oven beeped, I'm inclined to think that it's true, and that it's false that the oven was silent.

Other mental states aim at truth too, but they do not have the involuntary, affective character of belief. I believe automatically that the oven beeped when I'm standing near the beeping oven. On the other hand, I can take a position willingly "for the sake of argument." I might be undecided about the truth of a controversial philosophical view, but suppose that it *is* true to reason through its consequences. If I'm a scientist, I might adopt the laws and theoretical posits of the best theory at the moment, in full knowledge that it might later be replaced by a better theory. These are instances of *acceptance*, rather than belief, according to Cohen (1989).

Critics of non-summativist views like Gilbert's maintain that such views are not about group *belief* at all, but group acceptance. Groups cannot believe because, whatever cognitive states they have, none of them are involuntary in the way that doxastic involuntarism says that beliefs are involuntary. Since those states are adopted voluntarily, and beliefs are necessarily not voluntary, those states of groups

are not beliefs.

Consider the Events committee again. Phyllis voluntarily throws out a suggestion for consideration. The other committee members voluntarily agree (even if half-heartedly), or they voluntarily abstain from offering support or dissent. All of the members act voluntarily in taking the position, as one body, that Rowling would make a great speaker. So the group forms its position voluntarily. As such, it cannot be a belief. This is the *argument from doxastic involuntarism.* Here is the basic form of the argument:

1. If all of the individual contributions by group G's members are voluntary, and those voluntary contributions result in G being in state F, then G's being in F is voluntary.

2. The group members' contributions in non-summativists' paradigm cases (e.g. Events committee) are voluntary.

3. So the groups in those cases voluntarily adopt their resultant states (e.g. the Events committee "taking the view" that Rowling would be a great speaker).

4. If a state F of an agent S is a belief, then being in F is not under S's direct control, i.e. not voluntary.

5. Therefore, whatever state groups are in at the end of paradigm cases, such states are not beliefs.

Premise 4 is the definition of doxastic involuntarism, the truth of which I'm assuming here. It is not an uncontroversial view (though I find it convincing), but

110

accepting it for the sake of argument is important for meeting the critic here on their own terms. Premise 2 is, I take it, a fair assessment of the cases typically presented on behalf of non-summativism. Premise 3 follows from premises 1 and 2 by *modus ponens*. For the sake of defended non-summativism as a theory of group belief, belief properly so-called, we need to look at premise 1.

Premise 1 says something about the relation between the contributions of group members towards group activity, and the group state or act that those contributions result in. Consider the Events committee case, and suppose that the Events committee really does take the position that Rowling would be a great speaker, because of all the transactions among the group members with respect to that proposition. It is *because* all of those transactions were under each member's control that the resulting state is under the plural subject's control, and hence, not a belief. Is that true?

In general, the fact that the parts of a whole are all F does not entail that the whole itself is F. A human being is composed of atoms, each one of which is inanimate. It doesn't follow that the whole human being is inanimate, simply because the parts are. Premise 1 would fail if it were an instance of the (false) generalization that if all the parts of a whole are F, the whole is also F. So if premise 1 is true, some other explanation must be sought.

Maybe premise 1 must be accepted because the alternative is unacceptable. Suppose, for the sake of *reductio*, that a bucket brigade is putting out a fire. Every time anyone in the brigade passes a full bucket up the chain, their doing so is up to them. Are we to believe that, even though each member acts voluntarily, the brigade itself is acting automatically, just "going through the motions"? Could a

111

pair of people just "have" a collective belief, automatically, after a vigorous (and voluntary) negotiation between them? Surely not.

This is too quick, however, since there do seem to be instances where voluntary behaviors initiate an automatic, or involuntary response. Suppose, after dining out, that I want to calculate the tip. I begin thinking through the problem. The act of calculating is up to me; I'm doing this voluntarily. When I reach the end of the calculation, the answer is revealed to me. It doesn't seem to me that it's up to me whether I believe the answer. The deliverances of my considerations tip me towards belief. It was up to me whether I considered the issue of the tip at all, but it's not up to me whether what seems true as a result of those considerations commands my belief.

Alston presents a similar case:

> ...I engage in prolonged study of the mind-body problem or of the existence of God. I carefully examine arguments for and against various positions. It seems to me that none of the positions have decisively proved their case, even though there are weighty considerations that can be urged in support of each. There are serious difficulties with all the competing positions, though, so far as I can see, more than one contender is left in the field in each case. So what am I to do? I could just abandon the quest. But alternatively I could, so it seems, simply decide to adopt one of the positions and/or decide to reject one or more of the contenders. Is that not what I must do if I am to make any judgment on the matter? And isn't that what typically happens? I decide to embrace theism or

112

epiphenomenalism, and forthwith it is embraced.

...I would suggest that in each case the situation is better construed in some way other than as initiating a belief at will. The most obvious suggestion is that although in these cases the supporting considerations are seen as less conclusive, here too the belief follows automatically, without intervention by the will, from the way things seem at the moment to the subject. (Alston, 1988a, pp. 264–266)

An Alstonian gloss can be given for cases of negotiated collective belief. Take the Event committee case. Phyllis suggests Rowling, as something for the group to negotiate openly. This negotiation is the group considering whether Rowling would make a great speaker, and the exchanges in the negotiation by the members are each up to them. But when the negotiation ends, it is not up to them whether they believe, as a group, that they adopt the view that Rowling would be a great speaker, anymore than it is up to me whether I believe the result of my tip calculations. So the non-summativist view and the NCB thesis are compatible with doxastic involuntarism.

3.6 Objection: Can non-summativism explain group lies or group bullshit?

As explained in an earlier section, belief plays a central role in explaining human behavior. It does not only feature in mere descriptions of human activity, however; belief is for more than answering questions like, "Why did he look in the cupboard just then?". It is also an important component in evaluating others, in making

moral or epistemic judgments and holding people responsible. In particular, it is central to identifying lies and bullshit; we cannot hold people accountable for lying or bullshitting without invoking beliefs they have (or do not have). Further, as Lackey (Lackey, n.d.) points out, we often hold groups accountable for lying or bullshitting, so if these practices are to be coherent, we need an account of group belief that can distinguish group beliefs from group lies or bullshit. Call this *Lackey's desideratum*. She argues that non-summativist views cannot do this, that the plural subject view rules (incorrectly) that paradigmatic instances of group lies are genuine beliefs. As a result, she claims, we should reject non-summativist views.

To look closely at Lackey's argument, some place setting is needed. Lackey works with the Augustinian defition of lying, which holds that "a lie requires both (1) S asserting that p where S herself believes that not-p, and (2) S asserting that p with the deliberate intention to deceive" (Lackey, n.d., p.19). Lackey uses Frankfurt's analysis of bullshit (ibid., p.24), where S bullshits that p when S asserts that p, though S is unconcerned with the truth of p. Unlike considering whether p, or hypothesizing that p, bullshitting that p involves presenting oneself as, perhaps, believing p, or putting p forward as true, though one is not particularly concerned with the truth of p. Lackey presents two examples to illustrate the group cases. First, her Tobacco Company case illustrates group lies:

> Philip Morris, one of the largest tobacco companies in the world, is aware of the massive amounts of scientific evidence revealing not only the addictiveness of smoking, but also the links it has with lung cancer and heart disease. While the members of the board of directors of the company

114

believe this conclusion, they all jointly agree that, because of what is at stake financially, the official position of Philip Morris is that smoking is neither highly addictive nor detrimental to one's health, which is then published in all of their advertising materials. (ibid., p.19)[2]

Let C stand for the proposition that cigarettes are safe. The intuitive response to Tobacco Company is that Philip Morris has lied. According to the minimal Augustinian definition, this means that Philip Morris meets two major conditions.

First, they have publicly put forward C as true, by producing promotional materials that carry the message that C. Such materials (commercials, print ads, etc.) seem to have some kind of assertoric force, so it is fair to say that Philip Morris puts C forward as true when they publish it as their official position.

Further, since Philip Morris is lying, they must either believe not-C, or disbelieve C. According to the case, Philip Morris has evidence that conclusively shows not-C, and they believe not-C on the basis of that evidence, though they publicly avow C. So Philip Morris meets the first criterion in the Augustinian definition. Next, they assert C with the intention to deceive. The intent to deceive in Tobacco Company is revealed by the fact that, though they have evidence for not-C, they assert that C in order to protect their bottom line. In the real-world version of the case that Lackey cites (Lackey, n.d., p.20), the intent to deceive is more obvious, since Philip Morris concealed the studies that showed that cigarettes are unsafe. So Philip Morris's

---

[2]Philip Morris, as a large corporation, is likely not held together by joint commitments alone, so using it to discuss joint commitment accounts of group belief might seem odd. I use them here because they're Lackey's examples (except for News Company). However, nothing hangs on the use of institutions here. Small plural subjects could presumably lie or bullshit just as well as institutions.

behavior in Tobacco Company (and real life) meets both Augustinian criteria for a lie.

Lackey's paradigm case of group bullshit, Oil Company, is another example of corporate malfeasance:

> After the oil spill in the Gulf of Mexico, BP began spraying dispersants in the clean-up process that have been widely criticized by environmental groups for their level of toxicity. In response to this outcry, the executive management team of BP convened and its members jointly accepted that the dispersants they are using are safe and pose no threat to the environment, a view that they then made public through all of the major media outlets. It turns out that BP's executive management team arrived at this view with an utter disregard for the truth—it simply served their purpose of financial and reputational preservation.

BP's assertion that the dispersants are safe (call the content of this assertion D) counts as bullshit on Frankfurt's analysis. BP did not assert D because they thought it was true; indeed, they might have no propositional attitude towards D at all. They only asserted D because it was expedient. In Oil Company, BP appears unconcerned with the truth. They are bullshitting that D.

Tobacco Company and Oil Company clearly depict instances of groups lying and bullshitting, and Lackey correctly points out how important it is that we are able to identify them as lying and bullshitting. According to Lackey, this is where non-summative accounts of group belief fail. Gilbert's account cannot distinguish between a group's genuinely believing that p, and deceptively asserting that p, so it

116

cannot distinguish between a group lie that p and a group's honest assertion that p, nor can it distinguish between an honest group assertion and group bullshit. They fail to meet Lackey's desideratum, and thus, they fail as accounts of group belief.

According to Lackey, Gilbert's account fails to do so because, when Philip Morris's executives jointly commit to accepting C as a body (as they do in Tobacco Company), they have thereby formed the group belief that C. However, for Philip Morris to lie that C, they cannot believe C, but it would seem that, if Gilbert's account is right, they really do believe C. Recall that the plural subject account of group belief holds that a group believes that p when they are jointly committed to acting together as one body that believes p. Since the executives of Philip Morris have jointly committed to accepting C as a body, they have satisfied Gilbert's account, so that Philip Morris counts as really believing C. They have gone through all the requisite motions for believing C as a group — likewise for BP and D in Oil Company.

What is worse, Gilbert, or any non-summativist, cannot identify the lie by saying that Philip Morris believes not-C because all the members believe not-C, since that would be a summativist explanation of the case. Indeed, these cases' structure mirrors the structure of the hard cases for summativists, where a group believes something that none of its members believe. So we have paradigm cases of group lies and bullshit that are indistinguishable from the non-summativist's paradigm cases of group belief. This is an unacceptable result.

We can generate further Lackey-style examples that should pose similar problems for non-summativist accounts. Being able to track lies and bullshit is important, and it is likewise important that we are able to track instances of sarcasm. If we

117

were unable to distinguish between someone asserting sincerely that p and asserting sarcastically, or ironically, or satirically that p, conversations could be confusing and difficult, and we might mistake instances of joking for sincerity. So if a theory of group belief could not make *that* distinction, it would seem to have the same problems that Lackey describes with respect to lies and bullshit. For example, consider the following case, News Company.

> The public has become aware of clear instances of corporate lies and bull-
> shit when investigations reveal the inner workings of how Philip Morris
> and BP reacted to facts (or a lack thereof) that threatened their bottom
> lines. The editorial staff of The Onion, a popular satirical news organi-
> zation, defends Philip Morris and BP in print by jointly accepting their
> "defense" (call this proposition F) as their official position. Nobody on
> the editorial staff or in the audience believes F — F is so outrageous that
> nobody could seriously believe it — but it makes for good satire when
> The Onion makes the case for F as if in earnest.

To understand sarcasm, satire, or any dramatic performance, the audience needs to recognize that what the speaker utters is not what the speaker really believes. The audience has to recognize that there is a mismatch between the apparent attitudes and the actual attitudes of the speaker (or, as in the case of News Company, speakers). Utterances of this kind do not count as lies or bullshit. Nevertheless, News Company seems to generate the same problems for non-summative accounts as Oil Company and Tobacco Company. Though nobody would attribute the belief that F to the staff of The Onion, they have nevertheless jointly accepted F as their official

118

position. Lackey would say that, according to non-summative accounts, they really do now accept F. The problem here does not come from the dishonesty involved in lies or bullshit, or further intentions to deceive. It comes from the mismatch between the attitude The Onion pretends to have, and the attitude they really have.

If News Company is just as troubling for Gilbert's theory as Tobacco Company and Oil Company, it shows that group dishonesty is not necessary to press Lackey's challenge on the non-summativist. News Company involves no intent to deceive, no wrongful handling of the truth, and yet it seems to generate the same problems. Therefore, a proper understanding of Lackey's desideratum should find the common thread running between lies, bullshit, and satire, and whatever that common thread is, accounts of group belief should show how groups can accommodate it. I propose that the common thread is the *disparity condition*, or "Disparity," which some person S meets when they present themselves as having an attitude they do not actually have, and they do this for the sake of some further end that is served by the disparity between the actual and apparent attitudes. When S satisfies Disparity, there is a mismatch between their apparent and actual attitudes, and the mismatch is deliberate. This rules out the possibility of satisfying Disparity accidentally, e.g. by having inconsistent beliefs. Below, I show how the problem cases involve Disparity.

When S lies that p, S presents herself as having an attitude towards p which she does not actually have, for some further purpose. This is not sufficient to be a lie, according to the Augustinian criteria, because S must intend to deceive, and her actual attitude must be something like disbelief that p or a belief that not-p. There is nevertheless a mismatch between her actual and apparent attitudes, and

119

this mismatch serves some goal of S's, but the fact that Disparity holds does not determine what motives S has for the mismatch. In a case of lying, S maintains the mismatch in order to deceive. So Disparity is at least a necessary but insufficient condition for lying.

Likewise, when S bullshits that p, S presents herself as having an attitude towards p that she does not actually have. Again, S's meeting Disparity is insufficient for S to bullshit; there is a particular kind of mismatch between her apparent and actual attitudes that distinguishes bullshitting from lying. The liar is very concerned with the truth, as Frankfurt points out, but bullshitters are not. The mismatch between actual and apparent attitudes in lying is something like belief and disbelief, or belief that p and belief that not-p (i.e. when S lies, she appears to believe p, but actually disbelieves p or believes not-p). When S bullshits that p, S presents herself as believing p, though S does not really believe p — nor believe not-p, nor disbelieve p. S satisfies Disparity when bullshitting by presenting herself as having an attitude towards p though she really has no attitude at all. Nevertheless, satisfying Disparity does not make one a bullshitter, since Disparity does not specify what kind of mismatch there is, or why the speaker brings the mismatch about. So Disparity is necessary but not sufficient for bullshitting.

Further, when S says satirically that p, S satisfies Disparity, because there is a mismatch between the attitude S presents herself as having, and the attitude she really has. The Onion does not really believe F, but they present themselves as believing it. This is not sufficient for satire, since there are further facts about why they have such mismatched apparent and actual attitudes (e.g. they have them

120

because it will be funny). The Onion satisfies Disparity, but this obviously does not make their actions satirical all on their own. Disparity is necessary but insufficient for satire.

If Disparity is what lies, bullshit, and satire have in common, then accounts of group belief need to show how groups can satisfy Disparity. Lies and bullshit do not present a special case of Disparity that non-summativists need to account for separately from satire. Indeed, for non-summativist theories to satisfy Lackey's desideratum, they only need to show that groups can satisfy Disparity, i.e. that they can purposefully maintain a mismatch between their apparent and actual attitudes.

It is worth noting how minimally the group activity is described in Lackey's cases. In Tobacco Company, we are told that Philip Morris jointly commits to C being their official position, due to what is at stake financially; in Oil Company, we are told that BP jointly accepts D without regard for D's truth. We do not know what these commitments consist in, however, or how they were arrived at. If the NCB thesis is true, one of the more familiar means for establishing a collective belief is through conversational give-and-take. What conversation, then, establishes the collective beliefs in these cases? The details are never provided, nor are details of other procedures that might establish the collective beliefs. The cases stipulate that BP's board jointly accepts D, or that Philip Morris jointly commits to C being their official position. It is one thing to stipulate that Gilbert's conditions for group belief are met, however, and another to show that the conditions are met.[3]

---

[3]As Ballantyne (in press, pp.10–11) puts it, "Thinking about the focal cases so abstractly is perhaps similar to observing an automobile from 10,000 or 15,000 feet above and trying to determine the type of automobile it is. ... Looking down from those heights, all cars would look pretty much the same to us. Likewise, our perspective in the focal cases is less than ideal."

What would one of these cases look like in action? For the sake of simplicity, suppose that Philip Morris's board consists of two members: Andre and Louis. In the Tobacco Company case, Philip Morris, the company, is said to be aware of the evidence that tobacco is harmful. What does this have to do with Andre and Louis? It could mean that they read the research together, discussing the reports as they go; it could mean that they each read the research alone, never discussing its implications. The latter possibility is ruled out by Lackey's description of the case, however, since the board jointly agrees to their official position being C as a means of protecting their financial stakes. Presumably, the official position could not be adopted in this way unless the board also agreed that the research's conclusions (which they believe) threatened their bottom line, so some discussion would have to occur between them about the bottom line being threatened. With apologies, here's a dramatic reenactment:

Andre: These results look bad for us.

Louis: Nobody will buy our products if this research becomes widely known.

Andre: Okay, so our official position will be that cigarettes are safe.

Louis: Right.

On the NCB thesis, in the first two lines, Andre and Louis form the collective belief that their company's profits are threatened, since Andre makes a proposal which his conversation partner corroborates. As a result, Andre proposes something else: that their official position should be C. Louis accepts the proposal, and a collective belief is established. But do they collectively believe C, or is their collective belief instead *that their official position shall be C*? Tobacco Company is ambiguous

on this point, but the dialogue above seems to show Andre and Louis committing to adopting an official position, not to forming a collective belief that C. After all, the dialogue can continue as follows:

Louis: It's too bad the research turned out that way.

Andre: Yeah.

The implication here is that they believe collectively that the research correctly indicates the dangers of tobacco use. This is not, however, an incoherent turn for the conversation to take; they are not contradicting their commitment to their official position in talking this way. On the NCB thesis, then, the results of the conversation are that (a) they believe not-C, (b) they decide to publicly state that C as their official position, and (c) they do so in order to protect their bottom line. They meet Lackey's preferred criteria for lying, they satisfy Disparity, and all without violating the NCB thesis.

Lackey might respond by claiming that Andre and Louis negotiate their collective belief through a summativist procedure. The conversation essentially boils down to a series of sums of their individual attitudes on different issues. This is an explanation that non-summativists should want to avoid, on Lackey's view. However, on the NCB thesis, what happens in the negotiation of a collective belief need not reflect what the participants personally believe. The conversation between Andre and Louis can unfold as it does even if both of them privately harbor doubts about the quality of the research, but do not wish to appear "anti-science" in front of their peers. This even allows for the possibility that Andre and Louis each privately agree with the official position, even though they are lying when they collectively adopt the official

position that tobacco is safe. This is something non-summativism can account for, which summativism cannot.

Why does it seem hard for a group to satisfy Disparity on Gilbert's account? It could be that the plural subject account of group belief is overly "behavioristic," in the pejorative sense of the term. Or, to put the same point a different way, Gilbert's account cannot make sense of beliefs as "inner episodes" which cause or explain outward behavior. If outward behavior is *all there is* to belief, then merely going through the motions of believing that p suffices for really believing p. This, however, stems from the same abstracted use of Gilbert's view as we saw before. Compare this with an objection levied against Ryle's account of mental states (or at least, against what is commonly perceived as Ryle's account). Ryle is thought to have held that any attribution of mental states to a person could be analyzed without remainder into attributions to that person of dispositions to behave in certain ways. To say that I believe there is flour in the cupboard is to say I am disposed to behave in a large variety of ways. For example, if I were asked where the flour is, I would reply, "It's in the cupboard"; when gathering ingredients for croissants, I would root around in the cupboard, muttering under my breath about flour. Talk about mental states is really just talk about such subjunctive conditionals; Cartesian inner episodes need not apply.

Critics of Ryle might respond that such conditionals are often true of actors as well. If Ryle is right, then when Olivier mounts the stage as Hamlet, he is truly possessed of Hamlet's melancholic genius. On Ryle's account (says the critic), Olivier believes what Hamlet believes, granted the right conditionals are true of Olivier. This

is an absurd conclusion, generated by the inability of Ryle's theory to find a place for Disparity, in that the apparent attitudes *just are* the actual attitudes. Satisfying the right overt behavioral criteria determines what mental states someone has. A similar worry is at play in Tobacco Company, Oil Company, and News Company. In every case, they seem to jointly commit to believing something as a body, and that act makes it the case that they really do, as a group, believe what they say. Since they are jointly committed to acting as one body that believes some proposition, they do so believe. The satisfaction of overt behavioral criteria (in this case, going through the joint commitment motions) determines their states. As in the problem cases for Ryle, the apparent and actual attitudes collapse into the same.

The counterfactuals supported by genuine beliefs are not as shallow as the Olivier-Hamlet case supposes, however. Olivier does not believe that his father was murdered by his uncle like Hamlet does. He might be disposed to say the right words and make the right faces at the right time, but he does not feel conflicted blood-lust towards the actor playing Claudius. A liar or bullshitter might go through the motions in some circumstances with some people, but they do not always behave like a true believer. As Ryle (1979) notes, someone mimicking a sore loser is exercising a skill that the sore loser does not, even if the mimic's reproduction of the sore loser's grouching is perfect. The mimic doesn't wish that the game had gone differently, or obsess over how their game went wrong. They only pretend to. Liars, bullshitters, and satirists pretend, too, for different purposes, and even if they act just like believers, they are manifesting a skill in a way that the believer does not.

The critic's argument only works against a superficial version of Ryle's view; the

suite of Rylean resources affords a sympathetic reading of the problem cases. In a motto: there is jointly committing to pretending like a believer, and there is joint commitment to behaving like a believer. The non-summativist view is given different formulations which fudge this distinction. Consider, for example, the following clarification of "A and B collectively believe that p if and only if A and B are jointly committed to believing p as a body":

> What is it for people jointly to commit one another to believe some proposition *as a body* – the particular form of joint commitment at issue in the plural subject account of collective belief? One way of explaining this is roughly as follows: the parties will fulfill their joint commitment if – at least in the context of the present interaction, and to the extent that this is possible, they emulate, by virtue of the several actions and utterances of all combined, a single body that believes that p. In more familiar terms, they are to act as if they are "of one mind" with respect to p.

Lackey's examples can only work on the assumption that G's acting as if they have jointly committed to believing p as a body *makes it the case* that they have jointly committed to believing. Lackey examples do not support this, however. In Tobacco Company, Philip Morris jointly commits to making C their official position; this might amount to nothing more than jointly adopting a policy of confidently saying that C, not believing C as a group. In Oil Company, BP is presented as "jointly accepting" D, but what group behavior this joint acceptance consists in is not described. It is not said whether BP arrived at D through negotiation of collective

belief, like the Event committee did, or if, instead, the CEO said, "Shall we say that D?," to which everybody replied affirmatively. These distinctions matter, because they will inform the dispositional profile we expect these groups to have. If Philip Morris established by negotiation that the best science says that not-C, and then decided that their official position would be C, they will behave differently

To respond on Gilbert's behalf, suppose we recast Tobacco Company as Tobacco Company 2:

> Philip Morris, one of the largest tobacco companies in the world, is aware of the massive amounts of scientific evidence revealing not only the addictiveness of smoking, but also the links it has with lung cancer and heart disease. The members of the board of directors of the company believe this conclusion, and they all jointly commit, because of what is at stake financially, to lying by saying that the official position of Philip Morris is that smoking is neither highly addictive nor detrimental to one's health, which is then published in all of their advertising materials.

In Tobacco Company 2, we take it for granted that Philip Morris lied. This entails, on the Augustinian view, that they will deceptively present their company as believing C, though they actually reject C, and they do this out of concern for their bottom line. This only differs from the original case in how they are described. In the first case, the board "jointly agrees" to an official position; in the second, they jointly commit to telling a lie. Lackey argues that we rightly judge Philip Morris to be lying in the first Tobacco Company case. If so, then rephrasing the story as in Tobacco Company 2 should be innocent. However, Tobacco Company 2 doesn't seem

to undermine non-summativism, since it entails that Philip Morris believes not-C. According to Lackey, the only way to account for that would be by pointing to what each board member believes, in which case one is building a defense by assuming that the sum of board members' belief that not-C suffices for the board's belief that not-C, which Gilbert herself denies.

3.7   Conclusion

We have seen that we need a concept of group belief to explain group behavior, just as we need individual belief to explain individual behavior. Put a different way, groups need beliefs if they are to act on their commitments, just as individuals need beliefs to act on theirs. Clearly, these group beliefs depend on what the individuals think, broadly construed. However, summativism – the view that, for any p, the group believes p if and only if the right members believe p – cannot account for some cases of apparent group belief. Thus, we adopted a non-summativist view, following Gilbert's lead, where a group believes that p when they are jointly committed to believing p as one body. They often adopt their views through a process of negotiation, exchanging proposals and counter-proposals for what they might believe as a group.

The non-summativist view encountered difficulties, namely that group belief seems untethered from individual belief, and that the concept of group belief outlined by non-summativists cannot carry out a job that we expect the concept to do, namely, distinguishing genuine beliefs from lies. However, a full understanding of the mechanics of negotiation and how it actually plays out showed that these worries

were not well-founded. What we find is that the non-summativist concept of group belief is up to the tasks that we set before it. And as we'll see in the next chapter, it is also up to the task of being the center of a joint epistemology.

Chapter 4

Joint Epistemology

4.1 Introduction

Group belief plays a central role in the explanation of group behavior. When some people constitute a plural subject, we explain their behavior, *qua* plural subject, in the same way we explain individuals' behavior: by positing inner states like beliefs and desires. When we want to explain what groups are up to, we have to appeal to their joint commitments and, downstream from that, what they are jointly committed to believing. But to what end? The descriptive project would be helpful for social psychology or sociology; this has largely been Gilbert's agenda since *On Social Facts* (1989). Is there anything else?

Tollefsen (2003) argues that we are committed to group beliefs through our normative practices of holding groups responsible. She notes that it's not unusual to hold Strawsonian reactive attitudes towards groups *per se*. Just as we can resent

an individual (say, someone who shoves past in order to board a train first), we can resent a group. For example, a traumatized public can be aggrieved with a space program for negligently allowing a space shuttle to launch and critically fail. A parent can be angry with Andrew Wakefield and his co-authors for distributing fraudulent research. We want to explain groups' behavior because we want to hold them accountable, and belief attribution is part and parcel of that.

Note what the two groups are blamed for. The space program is blamed for doxastic negligence, or laziness. It rushed to judgment and didn't put in the effort demanded by the risks of space travel. To put it another way, they were culpably ignorant; they *should have known* that the launch was unsafe. Tobacco companies, by contrast, *did* know the dangers of their products. They saw the relevant reports and disregarded and suppressed them. We resent the tobacco company not because they should have known, but because they knew and did nothing. So if our reactive attitudes are justified, then there must be such a thing as doxastic responsibility for groups. They can be irresponsible in forming their beliefs, but can be responsible as well, even to the extent that their beliefs amount to knowledge.[1] Group behavior must be epistemically evaluable in order to make sense of our normative practices of holding them responsible.

But is it possible for groups to be epistemically responsible? To see whether they are, we need to look at the recent responsibilist views in epistemology, and see how they fit with the joint commitment framework. Epistemological treatments of group beliefs have faced objections, and these challenges must be met. Mathiesen

---

[1]I am assuming here that belief is a component of knowledge, though not everyone accepts that assumption, notably Timothy Williamson (2000).

has argued that, whatever group beliefs are, the Gilbertian sort cannot be epistemically evaluated, because they cannot reliably carry out the normative work of belief (Mathiesen, 2006). Carter has argued that group beliefs cannot be subject to defeaters, and so are, in an objectionable sense, undefeatable (Carter, 2015). After addressing these objections, we will turn to an application of the group responsibilist approach to a topic in contemporary epistemology: peer disagreement.

4.2   Epistemic responsibility

The familiar questions in 20th century epistemology include:

- What distinguishes knowledge from true belief? Relatedly, why is knowledge better than true belief?

- What is the nature of evidence? Justification? Warrant? Must it be "accessible," in some sense, to the would-be knower?

- What kinds of knowledge do we have, and what are their sources? Can we answer the skeptic?

The list is hardly exhaustive, but it is a fair sample. Questions like these concern when and how positive epistemic states obtain. They describe the dimensions along which epistemic norms get satisfied. In this sense, they resemble questions in normative ethics, such as, "What makes a given action right or wrong?" or, "Why are actions done 'from duty' better than actions that happen to conform with duty?" While these questions matter for ethical theory, they seem detached from the moral

concerns that guide us in our everyday lives. We want more from moral philosophy than theories that just outline the formal structure of right action; we want practical prescriptions and guidelines. Bare theorizing only takes us halfway, and might even alienate us from our moral practices (e.g. Stocker, 1976; Railton, 1984). Moral philosophy ought to do more, and be rich enough to guide us and inform our casuistry. Likewise, we might want more from epistemology than criteria that distinguish good beliefs from bad ones, or knowledge from non-knowledge. Instead, we might want some guidance; we might instead want a *regulative* epistemology, as Roberts and Wood (2007) have put it.

Enter virtue epistemology. The first appearances of virtues in contemporary epistemology saw traditional epistemology as incapable of providing satisfactory theories of knowledge and justification. Sosa (1980), for instance, argued that the two main theories of justification, foundationalism and coherentism, could not show how epistemic properties supervene on non-normative, non-epistemic properties.[2] Where, then, to ground our beliefs' justification? Sosa suggests that the normative status of our beliefs is derivative of the status of our reliable, stable intellectual dispositions that get us to the truth – as Sosa described them, intellectual virtues. This parallels the areteic turn in ethics, where the moral evaluation of acts became secondary to the evaluation of character traits.[3]

---

[2]In a related vein, Churchland (1979) complained that traditional epistemology's idealized epistemic agent bore no resemblance to actual human beings, and could not in any event describe the "cognitive virtues" (his term) of language-learning infants, who are incapable of forming the propositional attitudes that traditional epistemology traffics in. Churchland is not often recognized as a virtue reliabilist, however.

[3]It is not unusual for virtue epistemologists to draw the parallel with virtue ethics in this way. Virtue ethicists reoriented moral theory towards the evaluation of agents, not acts, which consequentialists and deontologists focus on. Likewise, virtue epistemologists reject the "belief-

In some ways, the intellectual virtues originally discussed by Aristotle, like theoretical wisdom or craft knowledge (Kraut, 2008), are not dissimilar from the virtues discussed by Sosa. But there seem to be other character traits, like open-mindedness, intellectual courage, or intellectual humility, which look less like craft knowledge, and more like the Aristotelian excellences of *moral* character (e.g. courage, liberality), yet are distinctively intellectual. These virtues do not seem to fit into Aristotle's original taxonomy of intellectual virtues, and yet would seem to fit well into an agent-based epistemology that draws an explicit parallel between ethical and epistemic evaluation.[4] If we reorient epistemology towards an evaluation of epistemic and intellectual character, then we should take traits like intellectual humility or conscientiousness seriously. While virtue epistemologists agree that character comes first in epistemic evaluation, they disagree over what counts as a virtue.

There are two senses in which something can be a virtue, or an excellence (Battaly, 2008). In the first sense of "virtue," sharpness, balance, and the capacity to keep a sharp edge are virtues of chef's knives. They are features which contribute to a knife's excellence, to its performing its function well. A dull knife is bad in part because of its dullness – a dull blade cannot slice tomatoes as a knife ought to – while a sharp knife is good in part because of its sharpness. The virtues make the knife good, and the vices detract from its goodness. A virtue in this sense would be those cognitive faculties that make for a well-functioning mind, like accurate memory or perception (Churchland, 1979; Sosa, 1980, 2010; Greco, 2010, 2007). In another,

based" framing of epistemological questions, and argue for the primacy of the intellectual virtues in epistemology.

[4]See Zagzebski (1996) who, like Sosa, distinguishes between act-based and agent-based theories in ethics, and imports that distinction to epistemology.

more evaluative sense, "virtues" involve deep aspects of one's intellectual character. Virtues in this sense are more like the moral virtues discussed by Aristotle, like courage or justice. In contrast with the aforementioned perceptual faculties, character virtues must be learned and acquired. Being virtuous involves not just stable dispositions, but proper motivations as well, like a love of or desire for truth (Zagzebski, 1996; Montmarquet, 1993). We are not born with open-mindedness, conscientiousness, or perseverance; we have to cultivate those parts of our intellectual character. Even so, like cognitive virtues, character virtues are supposed to contribute to the excellence of the virtuous thinker (Code, 1987; Zagzebski, 1996; Baehr, 2011; Montmarquet, 1993; Roberts & Wood, 2007).

Virtue epistemologists disagree over which sense of virtue has theoretical priority. Virtue reliabilists like Sosa and Greco stress the importance of cognitive virtues like perception in understanding knowledge, particularly in providing a Gettier-proof analysis of knowledge. Sosa's AAA account, for instance, analyzes knowledge as apt belief: S knows that p if and only if S's belief is *accurate* (true), *adroit* (the belief's formation manifests S's cognitive competence), and *apt* (the belief is accurate because adroit) (Sosa, 2010). Likewise, Greco's agent reliabilism emphasizes the role of competency and ability in explaining how the agent acquired their true belief (Greco, 2007, 2010). Virtue responsibilists focus instead on the role of character virtues in understanding epistemic phenomena. Code, for instance, highlights the importance of epistemic responsibility (Code, 1987), while Montmarquet (Montmarquet, 1993) makes conscientiousness, the drive to believe truths and avoid error, central.[5] For

---

[5]Virtue epistemologists disagree over whether virtue theories do or should connect with traditional debates over justification or the analysis of knowledge. Reliabilists like Sosa and Greco bring

all their disagreements, however, the two kinds of virtue epistemologists overlap in their commitments in a variety of ways, sometimes yielding theories which are not purely reliabilist or responsibilist. Battaly (2008) endorses a mixed theory: reliabilist virtues for low-level knowledge, responsibilist for high-level. And while Greco's virtue theory is mainly reliabilist, he maintains that responsibilist virtues will sometimes be necessary for keeping an agent reliable (Greco, 2010, p.10).

The overarching commitment they all share, however, is to making epistemic agency the primary target of epistemic evaluation. Epistemic normativity governs people first, beliefs second. A person's epistemic achievements, such as they are, are credited to their intellectual virtue, however "virtue" is conceived.This is the *credit thesis* (Greco & Turri, 2011) in virtue epistemology. It is usually formulated in part to explain what is distinctive about knowledge. Again, consider Sosa's AAA account or Greco's agent reliabilism. Knowledge, properly so-called, is distinguished from true belief not only by being justified, but by the role that the agent's competence plays in the acquisition of the belief. In Sosa's central metaphor for cognitive agency, what makes an archer's shot excellent is not only that it hits the mark, but that hitting the mark is due to the archer's skill. The archer is responsible for the accuracy; it wasn't just a "lucky shot." Likewise, the virtuous thinker earns credit for their true beliefs because their believing the truth is due to their cognitive virtue.

We noted before that we often hold groups responsible for epistemic blunders. We accuse corporations of culpable ignorance, and to an extent, we codify this into law,

---

their virtue epistemologies to bear on traditional debates, as do Montmarquet and Zagzebski, who are responsibilists. However, Code and Roberts and Wood deny that any such connection obtains. This marks an in-house disagreement between *anti − theory* in virtue epistemology, and theories which seek to fit the virtues into traditional epistemology (Battaly, 2008).

as in the Federal Corrupt Practices Act of 1977 (U.S. Department of Justice, 2017). Ignorance might sometimes excuse mistakes, but not when we cultivate that ignorance by sticking our heads in the sand to avoid any knowledge of a foreign business partner's corruption. As before with group belief, this gives us good reason to take attributions of epistemic responsibility to groups seriously, as more than loose talk or metaphor. The sheer ubiquity of epistemic responsibility attributions to groups gives us *prima facie* reason to make sense of them, so that our normative practices don't prove to be incoherent, by resting on a mistake. Indeed, precisely these concerns (but with attributions of responsibility to individuals) drive Montmarquet's work on doxastic responsibility (Montmarquet, 1993). If one is to be responsible for what one does, one has to be responsible for the beliefs that informed one's actions. For instance, if we hold Hitler responsible for his crimes, we should hold him accountable for the hateful beliefs which led him there. If we say that Hitler could not have believed otherwise, or that his beliefs were out of his control, we run the risk of excusing what he did. In order to make sense of our normative practices of praising and blaming peoples' behavior, we need to make sense of doxastic responsibility.

In the previous chapter, I argued that there is a sense in which group beliefs are involuntary. A group is made to have the belief it does, not by deciding to believe p, but due to the negotiation between members of whether p. This is to say that groups do not have direct control, or what Feldman has called "response control" (Feldman, 2008), over their beliefs. We have a relevantly similar control over certain behaviors, like holding one's breath for five seconds, or winking on command, simply by deciding and willing in the right way. Doxastic involuntarists claim that we lack

such control over our beliefs. Believing that p is not under our control in the way that raising one's hand is. I agree, and I argued for such a dimension to group belief in the previous chapter. However, doxastic involuntarists like Alston use this feature of belief against deontological conceptions of justification. Alston argues as follows (Alston, 1988b):

1. If S's belief that p is justified, S ought to believe that p.

2. "Ought" implies "can."

3. So, if S ought to believe that p, S has control over whether to believe that p.

4. S rarely (if ever) has that control.

5. S's belief that p is not justified.

For Alston, the deontological conception of justification, together with a plausible ought-implies-can principle and doxastic involuntarism, entails that few if any of our beliefs are justified – if positive epistemic status indeed rests on a deontological notion of justification. If epistemic responsibility rests on direct response control, then there is no such thing as epistemic responsibility.

Again, I grant that a reading of premises (3) and (4), where "control" is Feldmanian response control, yields a valid argument. There is more to epistemic evaluation, however, than the criticism of unreflective beliefs like "It is raining" or "The grass is green," which someone might form automatically with a glance out the window. Montmarquet, after all, agrees that we do not directly decide what to believe, that we do not have the response control described by Feldman. We do, however, have

responsibility for how we believe, or the way we believe (Montmarquet, 1993, pp.61–62). Suppose I see on the news that a man has been arrested in connection with a murder. I cannot directly control my spontaneous belief that the man is guilty, but I am all the same responsible for whether I am conscientious in doing so. One might be tempted to read Montmarquet's point here like Alston might: I directly control whether I am conscientious on a given occasion, so I have indirect control over what I believe on that occasion. This is not quite right, however. Montmarquet's point recalls what Ryle (2009, p.93) said about a person enjoying digging:

> To say that a person has been enjoying digging is not to say that he has been both digging and doing or experiencing something else as a concomitant or effect of the digging; it is to say that he has been dug with his whole heart in the task, i.e. that he dug, wanting to dig and not wanting to do anything else (or nothing) instead. His digging was a propensity-fulfilment. His digging was his pleasure, and not a vehicle of his pleasure.

The gardener is not doing two things – working and enjoying – when they enjoy their gardening. The gardener's enjoyment is the manner in which they garden. To put it another way, if the gardener did two things (enjoying and working), it might be intelligible to say of the gardener that they are enjoying, full stop. What they are doing is enjoying. But again, as Ryle points out (in "Adverbial Verbs and Verbs of Thinking" in (Ryle, 1979)), this is nonsense, or at least elliptical, because to enjoy is to enjoy doing *something*. Reports of plain enjoyment, or of simply taking pleasure, are incomplete. Likewise, if I conscientiously believe that p, it's

139

not the case that I am doing both of the following: (1) being conscientious, and (2) believing that p. Rather, the conscientiousness is a particular manner of my believing, like the gardener's enjoyment is a manner of their gardening. One cannot simply be conscientious; one must be conscientious in the course of doing something. To conscientiously believe that p is to believe that p in a conscientious way. So one might grant Alston the point that we do not have direct response control over our belief, but we do have control over the manner in which I believe.

To illustrate this point, imagine a camera that uses film; when the shutter opens and closes, the light from outside makes an image on the film. When I take a picture, I have no control over what happens after the shutter opens, where the light enters the lens and makes an image on the film. I do, however, have control over the cleanliness of the lens. After the photos develop, I could curse the bad luck that produced a poor picture, like sunlight glinting off a pond in the background in just the right way to produce a lens flare, or a moth fluttering an inch in front of the lens. However, if the pictures had defects caused by a smudge on the lens, or leaving the lens cap on, I am at fault. So some parts of the picture-taking process are out of my hands (like the quality of the film, the light filtering in, or the world I photograph), but some are up to me (the condition of the camera). Those parts that are up to me make the difference between taking a picture expertly and taking a picture carelessly. Similar considerations apply to belief. There are aspects of belief that are not up to us, over which we have no direct response control, but it is up to us whether we are fit to believe well. There are aspects of the manner in which we believe that we can control, for which we are responsible.

Montmarquet's analysis of doxastic voluntarism explains (plausibly, in my view) how we are responsible for how we *form* beliefs. However, so as not to hang the whole case on his view, we will consider other ways in which we are responsible for our beliefs. There are other aspects of our believing that involuntarists criticisms do not affect. Even supposing that Montmarquet's voluntarism is false, and that we are not responsible for how we form beliefs, we still hold each other accountable for how we *maintain* our beliefs (Aikin & Talisse, 2013, pp.13-14). Suppose I grow up in a liberal household, and as a result, hold many of the same political beliefs as my parents. I subsequently guide my political activities (like voting, activism, and discussion with others) around these beliefs. I would be intellectually lazy if I had no interest in what "the other side" had to say in response to my beliefs, or on behalf of their own. If, on the other hand, I sought out my political opponents' best arguments, considered them at length, and continued to think like my parents, it's less clear that I've believed negligently. In both cases, we might say that, initially, my political beliefs are formed irresponsibly, but excusably so, since I would not have thought other than what my parents taught me. The cases diverge in how I maintain those beliefs, which I do control, and am responsible for. In summation, even if we are skeptical of doxastic voluntarism, and as a result, skeptical of deontological conceptions of epistemic normativity, there could still be room for deontological norms that govern how we maintain beliefs.

4.2.1 Group virtues?

For virtue epistemology, the intellectual character of epistemic agents has explanatory priority over the epistemic status of a particular belief. The virtues constitute the normative core of epistemic assessment; an epistemic achievement like knowledge is creditable to the agent's virtue, however conceived. So if we can assess groups from the virtue epistemological perspective, groups must be capable of being virtuous in the first place. As Lahroodi notes, we do commonly talk as if groups are virtuous or vicious (Lahroodi, 2007). Investigative teams can get tunnel vision, falling prey to confirmation bias by focusing on evidence that supports their initial hunches and disregarding disconfirming evidence. Contrariwise, they can also be open-minded, scrutinizing all the alternative hypotheses with equal diligence. Just as with group belief attributions, we should not rush to judge these as instances of loose metaphorical talk, since virtue and vice attribution to groups plays a central role in our practices of praising and blaming groups. The negligence that led to the *Columbia*'s disintegration in the atmosphere in 2003 was epistemic in nature; NASA failed to pay attention to the evidence they had of damaged heat shielding. To put it in Code's (1987) or Goldberg's (in press) words, NASA should have known that the shuttle could not survive re-entry. Our evaluation of NASA here mirrors how we evaluate individuals. A doctor can be intellectually lazy and thereby fail to know something they ought to know, say, because they own an issue of the *Journal of the American Medical Association* they have not read (an example from (Goldberg, in press)). Our practices of holding groups accountable for epistemic failures mirrors our practices of holding individuals accountable. Deflating these practices by

142

appealing to metaphorical loose talk won't do. We have to take them seriously.

Attributions of virtues and vices to a group presuppose epistemic responsibility on the part of the group. A pair of co-authors cannot be collectively conscientious unless they are capable, as a group, of regulating their intellectual activity. If, for example, an exegesis on "Empiricism and the Philosophy of Mind" counts as an epistemic achievement creditable to both co-authors jointly, then their joint activity must be something for which they can be jointly responsible. If their cognitive activity was out of their hands, then holding *them* accountable for successes or failures would be misguided or incoherent. To borrow Strawson's example, one would rightly resent someone who deliberately shoves past on a subway platform, but not someone who stumbles because they lost their footing or were shoved themselves. Similar examples hold in the intellectual domain. We generally don't criticize people for not attending to evidence they could not have had. We blame NASA for the *Columbia* disaster because they had video of falling debris damaging the heat shielding on the shuttle, but did not sufficiently review their evidence. We would excuse their ignorance as a tragic accident, however, if the video had been deleted by a freak server glitch. So the reactive attitudes not only extend to groups, but to group epistemic activity.

Strawson (1962) argued that the practices of holding people responsible are prior to any metaphysical disquisitions on responsibility, control, or determinism, so that the free will debate mistakenly places too much importance on whether free action is possible. He may have been right in the individual case, but in the group case, distinct questions arise. For groups, the question is not whether group responsibility is threatened by determinism, but whether groups can be responsible for anything at

all, "over and above" their members being responsible. That is, can *we* be responsible for X-ing, or is *our* responsibility nothing other than *your* being responsible and *my* being responsible?

Suppose responsibility is summative, so that our responsibility for X-ing is just a matter of my responsibility for doing my part in X-ing and your responsibility for doing your part in X-ing. If group responsibility is summative responsibility, then, according to Pettit (2007), responsibility deficits are ubiquitous. He describes a case of gross negligence which resulted in mass casualties (ibid., p. 171):

> The *Herald of Free Enterprise*, a ferry operating in the English Channel, sank on March 6, 1987, drowning nearly two hundred people. The official inquiry found that the company running the ferry was extremely sloppy, with poor routines of checking and management... But the courts did not penalize anyone in what might seem to be an appropriate measure, failing to identify individuals in the company or on the ship itself who were seriously enough at fault. As one commentator put it, "The primary requirement of finding an individual who was liable...stood in the way of attaching any significance to the organizational sloppiness that had been found in the official inquiry."

Summative responsibility holds that a group cannot be faulted for X-ing unless someone in the group can be faulted for X-ing. The official inquiry looked at the accident and could not identify anyone so sloppy that their sloppiness sufficed to cause the accident. Any given individual's bad behavior was excusable, and since everyone was excusable, nobody was responsible. And yet someone *must* be responsible; the

144

organization cannot be excused. So group responsibility is not summative. That does not entail that group responsibility is non-summative, though. We must figure out how it could be non-summative, lest we let the practice of holding groups responsible lapse into incoherence. The idea that groups could be epistemically responsible will be a non-starter without an explanation of group responsibility in general. Without an explanation of group responsibility, it could turn out that, contrary to appearances, reactive attitudes towards groups are incoherent.

Typically, holding someone responsible for X-ing presupposes or requires that they X-ed deliberately or on purpose. If I resent someone for shoving past me on the subway platform, I typically assume that they intended to do so. The person who stumbles into me did not intend to stumble into me. If a group is to be held responsible *qua* group, then they must, *qua* group, be capable of acting on group intentions, or some other state that initiates their behavior in an intention-like way. Joint commitments perform precisely this function, enjoining each member to action so that they are accountable to each other and themselves. If a group X-es because they jointly committed to X-ing, then it makes sense to say they "meant" to X or that they X-ed "on purpose" (Gilbert, 2002). In order for a group to be collectively responsible for X-ing, on Gilbert's account, for instance, the group must at least have jointly commmitted to X-ing. Another condition for collective responsibility might be something like a control condition: one's responsibility for X-ing presupposes that X-ing is up to one.

Is there a sense, then, in which we might hold jointly committed people epistemically responsible? For simplicity, suppose that NASA's *Columbia* team had a

145

subgroup whose purpose was to ensure the integrity of the shuttle's exterior, including its heat shielding. This team might have a number of routines they commit to carrying out, like reviewing footage of the shuttle for anything suspicious, and criteria for what counts as a suspicious phenomenon worth investigating. This team is, in virtue of their joint commitment, responsible for monitoring the shuttle's exterior, and each member of the group is responsible, to themselves and each other, for doing their part in carrying out the group's task. To the extent that they jointly committed to an epistemic task (namely, gather data about the hull's integrity and interpret it accordingly), they were collectively responsible for its success or failure. Their responsibility for their failure makes them an appropriate target for blame.

The idea that someone could bear responsibility for an epistemic failing, like believing something false or failing to believe something true, strikes some epistemologists as a non-starter. Alston (1988b), we saw before, argues against deontological conceptions of epistemic justification on the grounds on the basis of two premises: (1) ought implies can, and (2) doxastic involuntarism. Doxastic involuntarism is the view that beliefs, properly so-called, are formed automatically, even when they are the end products of deliberate and careful inquiry. It is impossible for someone to form a belief as a direct act of will. However, if our beliefs were something we could be held responsible for, then they would have to be acts which we had control over. But if doxastic involuntarism is true, we lack the requisite control. Since doxastic involuntarism is true, we cannot control our beliefs, and thus cannot be held responsible for them. As a result, epistemic appraisal cannot rest on responsibility, praise, or blame.

There is, as I have argued, a defensible notion of group belief which preserves the involuntary aspect of individual belief. Anyone who demands a concept of belief as a state which results from the triggering of a passive, unreflective intellectual capacity should be satisfied by the conception of group belief already defended (see section 5 of the previous chapter). Even if the view I've defended is misguided, or is not fully persuasive, it is unclear why epistemic appraisal should be restricted to suitably involuntary doxastic states, let alone "states" at all.

Again, grant for the sake of argument that beliefs – properly so-called objects of epistemic appraisal – are involuntary, so that Alston is right in saying that beliefs cannot be governed by deontological norms. Even if we grant Alston's point here, there is still room for non-doxastic states and processes which can be objects of epistemic appraisal. That is, belief is just one type of state among many that we appraise from the epistemic point of view. Consider the Stoic conception of epistemic agency, for example, which held that, while one receives appearances from the senses involuntarily, the *acceptance* of those appearances was active (Wright, 2014). If Stoic acceptances seem psychologically unrealistic, then consider instead suspension of judgment, a cognitive state that must be directly willed.[6] Stoic acceptance, and the more familiar suspension of judgment, do seem subject to epistemic appraisal, yet they are not the output of passive capacities, but are instead real cognitive actions. Sosa (2010) compares suspending judgment to forbearing from taking a particular shot when hunting, because the shot is judged to be too risky. If

---

[6]What exactly suspensions of judgment are is a matter of controversy. Jane Friedman (2013) argues convincingly that to suspend judgment about p is not simply to have a credence of .5 about p, or an imprecise range of credences.

someone can responsibly suspend judgment and thereby earn epistemic praise (i.e. their suspension of judgment is creditable to them), then epistemic praise accrues to more than mere beliefs, the outputs of passive cognitive capacities. Since individuals can voluntarily adopt stances and views in a way suitable for epistemic appraisal (like Stoic acceptances or suspensions of judgment), we should expect that groups can as well. The voluntary character of certain group states need not be a barrier to groups being governed by epistemic norms, since we find voluntary individual states which are governed by deontological epistemic norms as well.

Moreover, there are other reasons why we should extend epistemic appraisal beyond belief. If nothing but correct belief matters from the epistemic point of view, then other epistemic goods, like understanding or wisdom (Roberts & Wood, 2007, chapter 2), do not have a place in epistemic appraisal. That is to say, that if all epistemic appraisal (i.e. saying of someone that they have done well or badly as a thinker) depended upon creditable production of true belief, then other values could not figure into epistemic appraisal unless such values could be shown to depend entirely upon having the right kinds of true beliefs. Suppose, for example, that we want to credit Melissa for her deep understanding of scientific modeling, and that, if we epistemically praise or blame someone's epistemic activities, our grounds for doing so are exhausted by their beliefs and how they are formed. So if we praise Melissa's understanding, our grounds for doing so are exhausted by Melissa's beliefs and how she formed them. However, Melissa's understanding does not seem reducible or identical to some set of beliefs she has about scientific modeling. While it might involve belief in some fashion, understanding is a high-level cognitive achieve-

ment more sophisticated than belief.[7] If such is the nature of understanding, we cannot praise or blame Melissa for it, which seems at odds with the standard practice of valuing understanding for its own sake. To preserve that practice, and the praise/blame of activities that deepen one's understanding, we ought to broaden our horizons and include more than the deliverances of passive cognitive capacities in our epistemic assessments. This will include voluntary epistemic activities aimed at hard-won goods.

There is a path open, then, for holding groups responsible for their views, and for how they come by (or fail to come by) and maintain their epistemic goods. We have seen that there are good reasons to think that group belief has an involuntary aspect, and that, even if those arguments fail, there are other voluntary states which are clearly subject to epistemic appraisal. But if crediting groups for their epistemic successes and failures involves the groups having or lacking virtues, we need an account of how groups can have virtues at all. Suppose, to illustrate, we hold that a virtue is like an acquired skill (as proposed by Annas (2011)). In order to maintain that groups can have virtues, groups must be capable of having acquired skills, in a manner consistent with what has been established about how groups form (people make joint commitments), what groups are (jointly-committed people), and how they believe (many distinct processes, including negotiation). Some groups on a plural subject model can be said to acquire skills as a group (Palermos & Tollefsen, in press). A band can jointly commit to practice a particular song together every Monday, for example; a basketball team can practice a particularly complex play. Not all groups

---

[7]The skeptical can substitute wisdom in place of understanding here; it is close to true by definition that wisdom is a hard-earned good.

have the benefits of long-standing commitments, though, and as such, no time to acquire a skill or develop whatever traits constitute virtue. Yet it seems that a newly-convened committee composed of strangers could be epistemically virtuous or vicious after a short period of the members getting acquainted. So, at least on some understandings of virtue, there seems to be a problem for talking about virtuous plural subjects. While it makes sense to attribute virtue to some recently-formed groups, virtues should take time to develop. It doesn't come easily, so how can some groups have it so suddenly?

One quick answer, noted by Lahroodi (2007, pp.285–286), is to go summativist: a group G has a virtue V if and only if some of G's members have V. Lahroodi calls this "correlativism." For correlativism, it is no surprise how a group gets to be open-minded once they come together. The group simply inherits its open-mindedness from the open-minded member(s). There is no puzzle about how a group becomes virtuous, because it is wholly determined by the members' virtues. But is this right? It helps to recall a plausible, logically weaker thesis than summativism, whereon the attitudes of a group supervene on what that group's members do, such that the group-level attitudes preserve the relevant content from the individual level. Suppose, for instance, that a group G believes that p. On the supervenience thesis, it must be the case that some member(s) of G take some stance towards p – not necessarily belief. Summativists go one step further, and say that group attitudes just are the member attitudes. Non-summativists satisfy the supervenience thesis by claiming that a group's belief that p consists in the members' joint commitment to believe that p. At the individual level, the members are taking a stance with respect to p,

150

namely, committing with each other to believe that p as a group, even though this particular attitude towards p is not the same kind of attitude the group itself takes. To wit, the members are jointly committing to believe p, while the group believes p, full stop. All the same, non-summativism satisfies the supervenience thesis, since the theory of joint commitment shows us how the members' action together yields the group-level belief.

When we go from belief to virtue, we see that the difficulties for summativism about belief reappear in a new guise for summativism ("correlativism") about group virtue, but again, rather than abandoning summativism wholesale, it helps to take account of what makes summativism plausible in the first place. Where the attitude supervenience thesis yielded a plausible constraint on a theory of group belief, a similar view about group virtues might appear here as well. According to the virtue supervenience thesis, when a group G manifests a group-level virtue V, G's manifestation of V depends on G's members acting individually or together in ways that are constitutive of V. For instance, according to correlativism, a group member manifesting perseverance suffices for the group-level manifestation of perseverance. Anti-correlativism denies this, but agrees with the trait supervenience thesis. What the group members are like determines what the group is like; groups can't differ with respect to their traits without a corresponding difference in what their members are like.

An example can help here. Suppose that two people, Eric and Finn, want to improve their attendance at the gym, but have motivation problems. They hate working out, feel insecure about exercising in front of others, and would rather do

something easy with their time, like play video games. All the same, they realize that they need to get healthy, so they jointly commit to going to the gym together every other day. While it's hard for Eric to feel excited about the gym, he finds it easier to get moving if he can hold Finn to their commitment. Neither Eric nor Finn ever shake the feeling of dragging themselves to exercise, but each one's fidelity to his commitments makes them jointly overcome the obstacles to getting healthy.

What traits do Eric and Finn have, and what, if anything, do they jointly manifest? Neither one seems to possess perseverance or grit with respect to their physical health goals. Adapting from King (King, 2014, p.3507), one has the virtue of perseverance if and only if one is disposed to continue in one's endeavors for an appropriate amount of time, with serious effort, in the pursuit of goods, and despite the presence of obstacles to one's acquiring, maintaining, or disseminating these goods. Neither Eric nor Finn have this disposition individually, since their track record shows that they will often, for example, excuse themselves from working towards their health goals for bad reasons. On their own, they're quitters. When they jointly commit to going to the gym, they get new practical reasons to go. They don't want to let each other down, nor do they particularly want to lose face. When they act like this together, they overcome the obstacles to improving their physical health, and they go to the gym more often, in spite of each one's protestations. Even though neither one individually looks like an exemplar of grit, together they persevere.

It would be too quick to say this example shows that groups can have perseverance even if its members do not. The fact that Eric and Finn persevere together does not entail that they together possess the character trait of perseverance. As King

152

points out, there is a difference between *exercising* a virtue, and acting in a way *characteristic* of the virtue. To exercise the virtue, one must possess the virtue in question, while acting in a way characteristic of it does not. Someone can act bravely on a single occasion without having courage, or spontaneously tell a joke without having much wit. For all this example shows us, Eric and Finn can act together in a way characteristic of perseverance; more needs to be shown if they have genuinely exercised the virtue. Suppose, then, that Eric and Finn's persevering is not something that happens once or twice, but repeatedly over many months. Their persevering even manifests modal stability. They normally do not allow each other to skip a trip to the gym, and only a few circumstances would permit it, like illness. When their behavior exhibits this kind of stability and regularity, we have more reason to say that it flows from virtue, rather than being a happy accident. But what is it about Eric and Finn together that regularly spurs them to exercise, which they lack individually? It is their joint commitment to be healthy, and each one's fidelity to their commitments, which accounts for their actual and possible successes.

Their perseverance does not simply consist in their joint commitment or their individual characters. Suppose Egon and Fred also jointly commit to exercising, but their commitment is non-specific and unenthusiastic, so that they decide to go to the gym, but without any particular schedule or guiding ethos. The non-committal character of their joint commitment makes it easy to satisfy. On Monday, Egon can ask Fred to go to the gym, to which Fred can respond, "Maybe later this week." On Friday, Fred can then ask Egon the same, to which Egon can respond, "Next week instead?" When their joint attitude does not involve taking their goals seriously,

153

there is no shared standard to hold themselves to. Without such, the normative structure afforded by joint commitment cannot help the jointly committed parties regulate themselves.

We can now see why correlativism is false. First, it claims that a virtue V cannot be had by a group G unless a member of G has V. This is false, as the case of Eric and Finn shows. But could correlativism's sufficiency claim be true, i.e. that if a member of G has V, G has V as well? Lahroodi notes that two groups can have the same members, but have different virtues. Suppose Cooper and Truman have intellectual perseverance as partners on a research project, but they also regularly play bridge as partners. If the character of the members sufficed for a group to have V, then we would expect groups with exactly the same members to have exactly the same virtues, and yet Truman and Cooper as bridge partners need not have the intellectual grit of Truman and Cooper on the project. Indeed, Truman and Cooper at the card table might be positively lazy and negligent. So having the same membership does not make groups the same in their character; correlativism's sufficiency claim is likewise false (Lahroodi, 2007, pp.285–286).

4.2.2 The nature of group virtues.

If Eric and Finn have a virtue together than Egon and Fred lack, we would want to know how what they have could count as a virtue, and whether it (assuming it *is* indeed a virtue) counts as something they have *together*, rather than a mere sum of personal characteristics.

To answer the first problem, we need some account of what virtues consist in.

154

Some, like Annas (2011) and Bloomfield (2000), see a deep connection between virtues and skills, as both involve acquired excellences of character and ways of assessing the normative features of particular circumstances (at least according to ancient sources). Virtue epistemologists, but mainly responsibilists, also conceive of virtue as acquired traits. Zagzebski (1996), for instance, thinks that to possess an intellectual virtue involves having a characteristic motivation – a love of truth and a desire for cognitive contact with reality – which makes the agent reliable in obtaining true beliefs. This does not necessarily involve a skill-like component, but the approaches do share a commitment to virtue being a deep, hard-won aspect of one's character.[8]

This commitment to the psychological reality of stable character traits has attracted criticisms that its empirical content has no grounding in the psychological literature. Situationists cite experiments which show that normatively irrelevant details of one's circumstances can change how one responds to normatively salient aspects of the environment (Hursthouse & Pettigrove, 2011). According to critics, these experiments expose the pernicious instability of character traits, so that such traits could not play the foundational role in a branch of moral theory (let alone normative epistemology). Without the requisite psychological structures for virtue theorists to hang their foundational concepts on, their theories fall apart. Moreover, even if such traits were stable enough and could be acquired, it is difficult to see how jointly committed people could exercise them, unless they had long-term com-

---

[8]Whether virtue requires a success component is a minor controversy in the virtue literature. Montmarquet (1993, pp.20-21) argues that, even if the intellectual virtues only produced false beliefs in, say, a Cartesian evil demon world, such traits would still be virtuous. See also Wright (2009, 2013), who argues for a Stoic conception of virtue which eschews success conditions.

mitments which could afford the time for acquisition. Only long-lived groups could have such traits; Eric and Finn would not, unless time allowed.

Recent virtue theorists have located the possibility for virtues in stable *attitudes* rather than character traits. Attitudes, as studied in the empirical literature on personality, are psychological complexes involving connections of varying strength between cognitive and affective states (Tanesini, 2016a, 2016b; Webber, 2013). Attitudes, properly so-called, are not propositional attitudes, like belief or desire. Rather, they are complexes of psychological states which associate a summary evaluation[9] (with a positive or negative valence) with an object, broadly construed. I could have a negative attitude towards fennel, so that when I am disposed to react negatively to fennel when I encounter it. By contrast, I could have a positive attitude towards the philosophy of George Berkeley, such that opportunities to discuss or learn about Berkeley excite me and grab my attention. Attitudes can also differ in strength. "Strength" here does not refer to the felt intensity of the affective states associated with the evaluation; for instance, my aversion to fennel is not strong simply because I respond to fennel by gagging. The strength or weakness of the attitude is in the *association* between the psychological complex and the object of the attitude. My negatively-valenced attitude towards fennel is strong because fennel nearly always provokes the same response in me. If fennel only occasionally, or rarely ever, provoked such a response, it would be a weak association. For instance, if, on a single occasion, the scent of fennel turns my stomach, but never does again, then the attitude, however negative, is weak.

---

[9]These evaluations can serve as many as six different functions, but the explanation will not be needed here.

Virtues, on this account, are kinds of attitudes directed at particular objects. For instance, Tanesini argues that intellectual modesty consists in caring for one's epistemic successes, not because such goods reflect well on oneself, but because they promote epistemic goods (Tanesini, 2016a, p.17). This association between the summary evaluative state and one's epistemic success has the same features typically attributed to virtues: "...their motivational structure, their unreflectiveness, their cross-situational stability and their ability to direct attention" (ibid.).

What Eric and Finn have, which Egon and Fred lack, is a positive group attitude that structures their joint activity. On the attitude approach, they do not have a collective *character trait* which makes them persevere, but rather a strong attitude that they strengthen through shared practice. The strength of *their* grit does not come from Eric and Finn's individual perseverance, but from their fidelity to their commitments. It is their readiness to hold each other to their shared standards which enables their perseverance, not their individual grittiness or personal attitudes. This example shows that a group can act from virtue even if the members do not have themselves have the virtue in question. In particular, their virtuous activity is a manifestation of the attitude to which they are jointly committed.

It is a short step from Eric and Finn's perseverance to intellectual perseverance. Suppose Cooper and Truman are collaborating on a research project, and when they work alone, they find themselves easily distracted. Encountered difficulties tend to discourage them and get their minds wandering. So they set up regular times to work together, to talk about the project and its various hurdles. While these meetings do not necessarily make the work easier, or smooth away its problems, they find them-

selves more engaged and focused. The work stays challenging, but becomes fun; they establish a flow. They stick to their project consistently, in a way that they could not had they worked more independently. Their joint activity exhibits intellectual perseverance as King understands it. Cooper and Truman are disposed, through their shared attitude, to continue in their endeavors for an appropriate amount of time, with serious effort, in the pursuit of goods, and despite the presence of obstacles to their acquiring, maintaining, or disseminating these goods. They do this despite falling short of these standards as individuals. *They* manifest perseverance together; their behavior comes from their joint commitment. Like Eric and Finn, Cooper and Truman are not just jointly committed to working together, but also have a joint commitment to have a strong attitude as a group towards their work and its challenges. And again like Eric and Finn, it is attitude to which they are jointly committed which does the motivating work in making them persevere through hardship. Each one's individual fidelity to their commitments is crucial, but without the shared attitude, there would be little for them to be faithful to.

Given that Cooper and Truman manifest intellectual perseverance as a group, it follows that there are group virtues, which can be possessed independently of a group member having the virtue in question. Those virtues, when exercised, make their possessors responsible and creditable for the achievements they bring about. Moreover, when the virtue is possessed by a group *via* its joint commitments, the group actions which manifest the virtue are thereby creditable to *them*. They get credit for what *they* did together, not merely for what each individual did, because it was their joint commitments which brought about the valuable outcomes. This

158

account of group virtue holds that V is a *group* virtue when a group is jointly committed to holding the attitudes constitutive of V; a group manifests a virtue V when the virtue-constitutive attitudes to which they are jointly committed motivate the group's actions.

At this point, however, a skeptic might respond that the examples considered above do not show that there are *group* virtues at all, but rather individual virtues that find their expression in group contexts. Eric and Finn's joint achievements are not creditable to their group perseverance, but to virtues possessed by Eric and Finn as individuals, namely, their fidelity to their joint commitments. Their joint attitudes are not necessary for explaining their success as gym buddies. We need only reference how Eric and Finn each take their commitments seriously, and that they have a joint commitment. The group is, in one sense, ineliminable from the explanation, but only because it features in how Eric and Finn's virtues are described. The group $itself$, however, need not have any virtues for the example to work (Cordell, 2017), and the same goes for Cooper and Truman. Their individual virtues do the heavy lifting (granting that these virtues make irreducible reference to groups), but they are not virtues of the group, properly so-called.

Consider, for instance, a kind of intellectual liberality, exemplified by Socrates's dialectical midwifery or Parfit's legendary generosity. Such a virtue could no doubt help a group flourish, especially a group that requires a lot of conversation. If everyone in the group took a selfless, giving attitude towards their fellow group members, so much the better for the group. Suppose that Cooper and Truman each have such a virtue, and we claim further that their group-directed generosity is

159

constitutive of the group's open-mindedness. The skeptic would claim that, however the group behaves *qua* group, its "open-mindedness," such as it is, is epiphenomenal. Cooper's and Truman's group-directed generosity suffices to explain the relevant details – no "group virtues" necessary.

While this objection correctly identifies the importance of fidelity and how joint commitments figure in its content, it overestimates how much personal fidelity (or group-directed virtues generally) can explain, and underestimates the role of joint commitment in establishing a shared attitude. Egon and Fred, the failed gym buddies, could possess fidelity to their commitments, yet fail to motivate themselves to go to the gym. If the above objection were right, this would not be possible, and yet it seems to be anyway. Egon and Fred could be faithful to their commitments, but fail to jointly commit to the right kind of shared attitude. On the account proposed here, joint activity counts as an exercise of virtue when it flows from the shared attitude that constitutes the virtue. Without the joint commitment to the attitude, a group will only act virtuously by accident. So the members' fidelity to their commitments is not enough; they need to have the right kind of commitments, namely, commitments to the kinds of attitudes that are constitutive of virtue.

4.3   Is group epistemology possible?

Virtue epistemologists hold that epistemic success is creditable to epistemic agents exercising their intellectual virtues, however construed. Groups can exercise intellectual virtue in the formation of their beliefs, so that their epistemic achievements are

creditable to their virtue. Moreover, their virtue is not simply a matter of their members' virtues; it is instead the virtues of the *group*, rightly so called. However, some concerns about the viability of group epistemology remain. Mathiesen (2006), for instance, has claimed that group belief, understood as a joint commitment to believe as a group, cannot aim at truth in a way fit for epistemic assessment. Carter (2015) has also cast doubt on the possibility that group beliefs can amount to knowledge, given that they seem insensitive to defeaters.

4.3.1 Mathiesen: The aim of belief

Mathiesen (2006) argues that group beliefs, as understood by Gilbert[10], cannot have the substantive epistemic features beliefs require for playing their epistemic role. She argues:

1. For belief to play a distinctively epistemic role, it must aim at truth.

2. The states that are beliefs according to plural subject theory do not reliably aim at truth.

3. Therefore, plural subject beliefs cannot play an epistemic role.

The argument above is valid, so to reject its conclusion, at least one premise must be false. Premise (1) makes a claim about what it takes for a belief to be assessed from the epistemic point of view, and premise (2) claims that plural subject theory's beliefs lack this essential feature. We will consider each premise in turn.

---

[10]Mathiesen focuses specifically on Gilbert's view here; she does not intend for her argument to generalize to all non-summative theories of group belief.

Premise (1) appeals to a popular thesis about one of belief's essential features: it aims at truth. To outline this view, Mathiesen turns to two of its main defenders: Williams and Velleman. Mathiesen (2006, p.163) attributes to Williams (1973) this description of the aim of belief:

- a belief succeeds or fails by the truth norm – if I come to think my belief is false, I stop believing it;

- to believe that p is to believe that p is true;

- saying "My view is that p" generally implies a claim that p is true.

This list fills in the view that the *norm* of belief is truth, but Mathiesen appeals to Velleman (2000) to further flesh out the notion that belief *aims* at truth (Mathiesen, 2006, p.164):

> According to Velleman, a view "has the aim of being the acceptance of a truth when it is regulated, either by the subject's intentions or by some other mechanism in ways designed to ensure that it is the truth" ... This does not mean that the practices that the agent employs must actually be reliable or "ensure" that what is accepted is true. Rather, they must be such that they are used by the agent, *because* the agent has the goal of accepting the truth and only the truth.

Williams tells us how truth is the norm of belief, but Velleman, according to Mathiesen, shows us how belief *aims* at truth: the practices and mechanisms that regulate belief do so because they are taken to ensure the truth of their outputs.

162

Premise (1) of Mathiesen's argument takes this view for granted; I will grant it for the sake of argument.

It is the above sense of "aiming" at truth that helps us to define two senses in which a group can be epistemically rational. She draws on Foley's (1992) account thus (Mathiesen, 2006, pp.165-166):

> A group G is **subjectively** epistemically rational iff G has the goal of believing* truths and avoiding falsehoods and abides by epistemic practices that G *believes** are effective in achieving that goal.

> A group G is **objectively** epistemically rational iff G has the goal of believing* truths and avoiding falsehoods and abides by epistemic practices that G *are* effective in achieving that goal.

Subjective rationality poses no problems for groups *per se.* A group might even jointly commit to procedures that are effective in getting to the truth. The problem, according to Mathiesen, is not that groups cannot adopt these practices, but that Gilbert's plural subjects cannot, because of the nature of their beliefs (as described by plural subject theory). Plural subject theory holds that a group believes that p if and only if the members are jointly committed to believing that p as a group. Thus, members of a group that believes p will have obligations to take p for granted in group contexts; they must not gainsay the group's view that p; they will rightly earn rebukes for flouting the group view in their capacity as group members; and so on.

Here, then, is Mathiesen's argument for premise (2). Suppose we think that groups can be conscientious inquirers, and do so by adopting the practice of open

inquiry and exchange of reasons in group deliberation. Consider such a group, G, which, through deliberation, adopts the view that p. This now means that G's members have obligations with respect to p; as members of G, they have to carry on with the other members of G on the assumption that p. If a member of G comes to think there are good reasons for thinking not-p, we might think that conscientious inquiry requires them to introduce those reasons into the group deliberative context. However, according to Mathiesen, the member's obligations to G to take p for granted would make it inappropriate to raise those concerns to the group. They would face social sanction for doing what they epistemically ought to do. Thus, a Gilbertian plural subject is unlikely to be subjectively or objectively epistemically rational; the normative structure of joint commitment which works to regulate practical rationality for groups backfires undermines group epistemic rationality.

This is a subtle argument, and a serious problem for plural subject theory, if premise (2) is true. One might first respond to Mathiesen by noting that Gilbertian group beliefs are no worse off than individual beliefs that aim at truth in the way she proposes. Indeed, the normative structure of group belief makes it hard to change, a risk that Gilbert herself notes (Mathiesen, 2006, p.169), but then again, the norms governing individual belief also will make it hard to change one's mind. If I reflect on my belief that p, I will infer from it that p is true, and that I would not believe it if I thought p was false. That is in part constitutive of belief: taking something to be true. Further, if I take p to be true, I will have no use for evidence against p; that evidence is probably misleading, given that p is true. So, *contra* Mathiesen, one might conclude that group belief and individual belief suffer from

164

the same shortcomings. That's regrettable, but it makes group belief no worse off than individual belief. While initially promising, this line of response misconstrues Mathiesen's point. She does not argue merely that the norms of group belief make it difficult for members to do their epistemic duty; it is not merely the case that they will face uncomfortable social sanctions for doing so. Rather, the problem is that such sanctions are justified by the norms of joint commitment; the norms of group belief seem to *require* that nobody object to the group beliefs. After all, if we believe that p, then our joint commitment obligates me to do my part in our acting as if p is true, and I am obligated not do anything which would make us act as if p is false. Individual belief does not face this issue. There might be contingent features of our psychology that give our beliefs a sometimes-counterproductive sort of inertia, but nothing like norms that make rethinking one's beliefs irrational. So Gilbertian group beliefs and individual beliefs are not partners in crime after all, since they are guilty of very different things.

Mathiesen's argument does, however, rely on some obscurity in how Gilbert's theory is presented (even by Gilbert herself). For premise (2) to work, when G believes p, then G's members must take the position that p "as group members in group contexts". They can, of course, speak in group contexts on their own behalf, so long as they qualify what they say (e.g. "speaking for myself..."), but doing so will not make the group epistemically rational. So, any member of G is required by their joint commitment to act *as if* they believe p. This, Mathiesen assumes, is what it takes for G to believe p as one body. For us to maintain our belief that p, none of us can contradict the belief that p, and this is, *ipso facto*, acting as if one believes p.

165

However, there is room for a difference between contributing to the group's being a single body that believes p, and acting, as a group member, as if one believes p. An obligation to the former does not entail an obligation to the latter. Suppose Donna and Laura jointly commit to believing p as a group. However, Laura comes to think there are some reasons to doubt p. Intuitively, Laura ought to bring these up in her discussion with Donna; she cannot, however, contradict the group belief that p. If Mathiesen is right, then they cannot be epistemically rational, because the members cannot revise their views in light of good reasons. Suppose, though, that Laura says to Donna, "I know we think that p, but maybe we should revisit why we do. What if q, which entails not-p?" Laura has not obviously contradicted the group belief that p; she has not obviously flouted her obligations to herself and Donna to maintain their acting as if p together. To see why, consider the individual case. It is possible for Donna to believe p, but wonder whether her belief that p is as well-founded as she thought. Merely having that suspicion does not dissolve her belief. Williams pointed out that, if Donna regards her belief as false, she would not continue believing it, but merely reconsidering the epistemic status of her belief does not do this. Likewise, in the group case, if Donna merely introduces preliminary reasons to reconsider their belief that p, she does not violate her obligation to maintain the group's believing p as one body. To make a proposal in the negotiation of group belief merely for the *reconsideration* of an established view is not to presuppose irrationally that the group's view is false. The group would still believe p if it started reconsidering the epistemic basis for p. So the normative structure of group belief for plural subject does not conflict with good epistemic practice, unless a much stronger aim of belief

166

thesis is adopted which rules out the Laura and Donna case. In conclusion, while Mathiesen's argument is valid, her case for (2) falters on this point.

4.3.2   Carter: Defeating group belief

Carter (2015) has argued that group knowledge based on Gilbert's account of group belief cannot make sense of defeaters, making well-formed group beliefs oddly un-defeatable. If group knowledge could not be subject to epistemic defeat, this would make group knowledge very different in kind from individual knowledge. As a result, he argues that we need a different conception of group belief, which allows group beliefs to be sensitive to defeat like individual beliefs are.

Defeaters, he points out, come in many varieties. First, defeaters can either be psychological or normative; second, they can be either rebutting or undercutting. Psychological defeaters behave in the following way. Suppose I have a belief that p at time t_1, and I later acquire the belief that q at t_2. However, q indicates that not-p. My new belief that q acts as a defeater for my belief that p. To be more specific, it acts as a rebutting defeater, by showing that p is false. It instead might have acted as an undercutting defeater, by showing that I either acquired my belief that p in an unreliable way, or that my belief that p has been sustained in an unreliable way since t_1. Note that psychological defeaters only act as defeaters *once they are believed*. Normative defeaters, on the other hand, need not actually be believed by a subject; instead, they are beliefs that the subject in question ought to have (say, because the subject's evidence supports it). Suppose again that I believe that p at t_1, and there is a belief I ought to have, q, which would indicate that not-p (or would

167

undercut belief that p). The belief that q, which I ought to have, would then be a normative defeater for my belief that p. The following table sums up the categories:

	rebutting	undercutting
psychological	believing q defeats belief that p by indicating not-p	belief that q defeats belief that p by indicating p was formed/sustained unreliably
normative	subject ought to believe q, which would defeat belief that p by indicating not-p	subject ought to believe q, which would defeat belief that p by indicating p was formed/sustained unreliably

According to Carter, the plural subject theory of group belief can accommodate neither psychological nor normative defeaters. They cannot accommodate psychological defeaters, because the nature of group beliefs rules out the possibility of a group believing both p and its defeaters. Further, the plural subject theory cannot accommodate normative defeaters, because the categorical reasons generated by epistemic oughts find no footing in joint commitment contexts, which (according to Carter) can only accommodate hypothetical reasons.

First, consider Carter's argument against psychological defeat for groups. He argues that, when proponents of the plural subject theory also accept a plausible principle about how groups act through their members, they lose the resources to

make sense of defeat for group beliefs. Consider the view of group belief Carter credits to Gilbert (ibid., p. 8):

> *Joint-Acceptance Belief* (JAB): (i) A group, G, believes that p iff the members of G jointly accept p; (ii) the members of G jointly accept that p when the members conditionally commit to accept that p; (iii) members of G conditionally commit to accept that p when each is committing to acting as if p provided the others do.

While somewhat more baroque than the account of group belief defended here, JAB more or less captures the core aspects of the view.[11] Next, consider a plausible principle about how groups act by means of their members acting (p. 9):

> *Group/Member Action Principle* (GMAP): For every group, G, and act, ϕ, G performs ϕ only if at least one member of G performs some act or other that causally contributes to ϕ.

According to GMAP, a group cannot do anything without the members causally contributing. It is impossible for a group to act without the members doing their parts. Groups have no causal power of their own; it depends entirely on the powers of their members.[12]

---

[11]Conditions (i) and (ii) involve an epicycle through "joint acceptance" present in some, but not all, of Gilbert's writings on group belief, where she uses "belief" and "acceptance" mostly interchangeably. It won't matter for our purposes.

[12]Compare this with supervenience theses about group attitude content. For instance, a group cannot believe that p if all of its members are wholly ignorant of p. A group cannot take an attitude of any kind at all towards a proposition p unless the members have some attitude, any attitude, towards p.

Carter's argument against psychological defeaters for plural subject beliefs proceeds as follows (pp. 9–10):

1. Group A believes p only if A's members are committed to acting as if p. [from JAB (iii)]

2. A can acquire the belief that q, where q indicates that not-p (whether by rebutting or undercutting), only if at least one of A's members causally contributes to A's acquiring the belief that q. [from GMAP]

3. Any such action in (2) contributes to violating one's commitment to acting as if p.

4. Therefore, if A acquires the belief that q, where q indicates that not-p, it's not the case that A retains its group belief that p. [from (1)-(3)]

5. Group A's belief that q, where q indicates not-p, can function as a psychological defeater for A's belief that p only if A acquires the belief that q while retaining the belief that p.

6. Therefore, A's belief that q cannot be a psychological defeater for A's (original) belief that p. [from (4) and (5)]

The argument is valid. Given that (4) is a logical consequence of (1), (2), and (3), and since (5) simply states what a psychological defeater for a group belief would be, the premises to scrutinize are (1), (2), and (3). Since the first premise follows from JAB (and, it should be said, from the view of group belief defended here), and the second follows from GMAP, those principles are also vulnerable.

170

What about (3)? Suppose we are working on a cutting-edge project at a particle collider. Carrying out our work requires a lot of group beliefs, including beliefs in the viability of our field's models. This means that we are jointly committed to believing that our models are sound, and thus, we are all committed to taking the soundness of our models for granted in group contexts. A psychological rebutting defeater that belief would be, say, the belief that the models are bogus. If, in the middle of our work, one of us says, "The models are bogus!", they would violate their joint commitment, which involves not contradicting the group belief. What would have contributed to a rebutting defeater for the belief in the soundness of the models cannot function as such here, since the violated commitment dissolves the original belief, and with the original belief gone, there is nothing to defeat.

The problem, then, is not that group belief makes it hard for groups to revise their views, as Mathiesen argued; rather, the plural subject theory makes defeaters impossible. We saw above that, according to Mathiesen, the normative structure of group belief forbids group members from doing what they epistemically ought to do; Carter, by contrast, argues that the normative structure of group belief precludes members from believing a psychological defeater at all, period. Individuals sometimes change their beliefs after acquiring a defeater; according to Carter, groups do not change their beliefs in response to defeaters they possess, because they *cannot acquire* psychological defeaters. So, for all Carter says, groups can rationally change their minds, but not in response to defeaters, if he has given a sound argument against Gilbert. But has he?

Note that, as an example of participation in joint epistemic projects, the exas-

171

perated exclamation ("the models are bogus!") is unusual. Instead, consider the following type of case. Suppose a strike team, T, is planning a military assault, beginning with a landing on a beach. After some deliberation, where they consult their considerations (call these "C"), they decide that North Beach is the preferred landing zone. Thus, T believes that North Beach is best suited as a landing zone (as opposed to, say, South Beach). Call this belief "N". Later, after some discussion about further details of the mission, one of the members of T, Hicks, says, "North Beach would be a good landing zone, but maybe we should reconsider South Beach, and some of C?" He brings up some of the considerations in C considered before, but puts them in a new light. He puts them forward as reasons why South Beach would be better suited for their purposes. After some negotiation, T comes to believe that South Beach is best suited for the assault, and thus believe not-N.

Do JAB and GMAP suggest that Hicks has violated his joint commitment to believe N? If Carter is right, then when Hicks proposes that they re-open the case for South Beach as a landing zone, he must be violating his commitment to taking N for granted with everyone else in S during their deliberations about the mission. While the exasperated physicist might seem to gainsay the group in an inappropriate way, it's less clear that Hicks does. After all, he does not say that N is false during deliberation with T, nor does he say that it is merely his personal view that not-N. Instead, he puts forward a proposal for something that the group might consider and renegotiate: that North Beach is their best option as a landing zone, all things considered. In doing so, he does not stop acting as if N is true, with respect to his S activities. He merely *proposes* that they review their previous decisions; it is

only an opening step in a negotiation with T. On its own, Hicks's proposal does not gainsay T's belief that N. To see why, consider a case involving an individual person, Vasquez.

Vasquez has the same task as T: plan a military assault, starting with finding a suitable landing zone. She reflects on all the considerations, C, for each of her options, and comes to believe N on the basis of C. As more details of the mission fall into place, she begins to wonder about the grounds for believing N. Even so, she still believes N; she just wants to revisit how she reached that conclusion. She goes back to C and reconsiders her reasoning. She comes to believe South Beach is better suited than North Beach. That belief defeats her belief that N.

In general, believing that p and reconsidering your evidence for p are not incompatible. One can believe p while reviewing the reasons against p; wondering whether one's evidence might actually defeat your belief that p does not automatically lower one's confidence in or defeat the belief that p. To see why, consider cases where one's belief survives, or is even strengthened by, a review of the evidence. Vasquez judges that N given C, but later wonders how much C really supports N. She reconsiders C, and comes to think that C supports N more than she initially realized. She thus comes to believe N more firmly than before. The reconsideration of reasons for believing N enriches and deepens her judgment, rather than shaking it. So in general, mere doubts about the grounds of one's beliefs do not on their own psychologically defeat belief.

Hicks's proposal to review the case for N, and the negotiation that follows, does for T what Vasquez's doubts do for her reasoning. Just as Vasquez does not lose

173

her belief that N just by reconsidering N's support, T also does not lose its belief that N when it reconsiders N's support. However, as Carter understands his premise (3), Hicks would be acting just like the exasperated particle physicist, contravening his commitments and dissolving the group belief that N. Yet he does not seem to be doing that at all. T responds to Hicks in a way consistent with psychological defeat. So Carter has not shown that plural subject accounts cannot make sense of psychological defeat, since the Hicks case seems to provide an instance of it. His argument against normative defeaters for groups still remains, however.

A normative defeater is a belief that someone *ought* to have, which, were it believed, would defeat a different belief they have. Suppose, for instance, that I believe p, but I possess evidence for q, which I negligently fail to believe. I ought to believe that q, and if I was believing as I should, q would defeat my belief that p (either by rebutting or undercutting). Thus, q is a normative defeater for p. So, Carter proposes the following (p. 13):

> *Group Normative Defeater* (GND): A (group) normative defeater for group A's belief that p is some belief that A ought to have, that q, where q indicates that A's belief that p is *either*: (i) false (i.e., rebutting) or (ii) unreliably formed or sustained (i.e., undercutting).

What kind of reasons do *epistemic* oughts generate? According to Carter (citing (Feldman, 2000; Railton, 1997; Cuneo, 2007b)), epistemic oughts are categorical in nature, so that their normative force applies independently of one's prior commitments or practical goals. Categorical reasons are distinct from hypothetical reasons,

174

reasons which one has because of one's goals, desires, or commitments. Carter now presents a straightforward *reductio*.

1. Suppose that a group G's belief that p is normatively defeated; there is a belief that G ought to have, that q, which would defeat their belief that p were it believed.

2. The ought in question is *either* non-categorical *or* categorical.

3. If the ought is non-categorical, it is not a genuine epistemic ought.

4. If categorical, it is both categorical and non-categorical.

5. Both horns of the dilemma are absurd.

6. Therefore, G's beliefs cannot be normatively defeated.

The argument is valid. Premise (1) is assumed for the sake of *reductio*. Premise (2) states an uncontroversial thesis[13]: an ought is either categorical or not. Premise (3) says that hypothetical oughts are not epistemic oughts. According to Railton, for instance, two agents with exactly the same background beliefs and evidence would have the same reasons for believing any given proposition, despite whatever differences they may have in their goals and desires. So the oughts must be categorical, if they are genuinely epistemic. But, if premise (4) is to be believed, this generates an absurdity. Carter defends (4) by claiming that the absurdity arises from GMAP and JAB. Suppose q is a normative defeater for G's belief that p. G ought to believe that

---

[13]Ignoring intuitionists for the sake of argument.

q, and if we assume GMAP, someone in the group ought to causally contribute to bringing about G's believing q. Suppose, then, that the ought is categorical. This, Carter says, gives rise to a new problem (p. 15):

> Given (JAB), any commitment M would have, qua member of the group G, to acting as if q (e.g. any commitment to ϕ) will always be *conditional*, in the sense, that M will, qua member of G, be only ever be committed to acting as if q provided others do. But then it seems that the claim that M ought to ϕ is such that it both (i) must be categorical, and (ii) must not be categorical.

Carter illustrates this point in a footnote by using a familiar example from Gilbert (ibid.):

> Compare here Gilbert's discussion of the obligation Jack has when jointly committing to walking with Jill. When Jack does something that's not compatible with walking together - for instance, walking faster than will be reasonably possible for Jill to walk along side him - then Jill can rightly rebuke him. Plausibly, Jill then is released from any obligation she had to walk at a pace accommodating Jack.

According to Carter, then, since a member's obligation depends on one's joint commitment with others, the ought is not categorical. A member's obligation here is like Jack and Jill's obligations to each other. Their normativity depends on their commitments, and reasons generated by things like commitments and desires are not categorical. So when we assume that the ought of "G ought to believe q" is

176

categorical, it generates obligations for G's members that are both categorical and non-categorical.

Carter's argument works on the assumptions (a) that epistemic defeat must come from categorical oughts, and (b) that reasons or oughts are not categorical unless they are entirely independent of any commitments. In response to (a), we might note that some epistemic projects might be optional. One must voluntarily adopt some projects with epistemic ends, so that any epistemic oughts are hypothetical or conditional. For instance, the epistemic oughts that scientists face in the laboratory have very different contents from the epistemic oughts faced by ordinary people. The conscientiousness demanded by the rigors of, say, cancer research, looks very different from the epistemic demands faced by non-scientists. But these demands depend on the goals of the cancer researchers; they only have those epistemic reasons by virtue of their scientific goals. Nevertheless, we might think these are genuine epistemic oughts.

Carter assumes that (b) categorical reasons must be independent of a person's commitments. However, the sense of dependence here is somewhat obscure. Suppose I promise my mother that I will send my grandmother a card for Mother's Day. I have made a commitment to my mother; I have committed myself to sending my grandmother a card. Moreover, I have now incurred a moral obligation to send a card because of my commitment. Promissory obligations provide categorical reasons, despite the fact that they depend for their existence on voluntary commitments taken on by promise-makers, yet because they are generated by one's commitments, they seem to be non-categorical. Here is how Shafer-Landau (Shafer-Landau, 2009) marks

the difference:

> Categorical reasons, as I will define them here, are reasons that obtain independently of their relation to an agent's commitments. Such reasons do not depend for their existence on their being instrumental to the achievement of any of an agent's desires, goals or cares...

> If there are categorical reasons for action, then practical instrumentalism is false. Practical instrumentalism ... is the view that the only reasons there can be are so-called *hypothetical reasons*, i.e. reasons to do things that are in some way ancillary to the achievement of one's commitments (cares, desires, wants, goals, etc.).

If we understand categorical reasons this way, then it is hard to see how promissory obligations provide categorical reasons. Likewise, if we think that we have certain special obligations generated by various commitments (like obligations to one's children or spouse), then this understanding of categorical reasons will not rule such obligations as providing anything but instrumental reasons. If we adopt the more minimal characterization of hypothetical reasons that Carter provides, where obligations are *conditional* in some sense, then the problem is even more acute. All promissory obligations would be non-categorical, since the reasons only obtain *provided that* one has made a promise. Likewise, if I have an obligation to my partner (e.g. to prioritize her needs over those of a stranger), it cannot be categorical, since the obligation only obtains *provided that* I am in a relationship. Some of the categorical reasons we have seem to depend on commitments we make, which, by the lights of some theories of categorical reasons, ought to be a contradiction in terms.

Suppose I have a reason to ϕ. The normativity of that reason, if it is an instrumental/hypothetical/non-categorical reason, is exhaustively explained by (i) ϕ-ing's being a means to bringing it about that ψ, and (ii) my desiring that ψ. Now suppose I promise my mother that I will send my grandmother a card, because making and keeping that promise will make my mother happy. I subsequently have a reason to send my grandmother a card. What explains the normativity of this reason? I made the promise in order to make my mother happy, so among the conditions that explain why I ought to keep my promise are the facts that keeping my promise would make my mother happy, and thus, would serve my goal to make my mother happy. Moreover, though, I have a reason to keep my promise because anyone who makes a promise ought to keep it, regardless of the ends they wish to bring about by promising. After all, if I stopped caring about making my mother happy, I would no longer have an instrumental reason to send my grandmother a card, but that would not dissolve the *moral*, categorical reason to keep my promise. Given the promise I made, I must still send the card, even though doing so would no longer serve to fulfill any goals I have. I have a reason other than the instrumental ones to send the card and fulfill my promise. So explaining the normative force of my reason to send my grandmother a card will not be exhaustively explained by invoking my ends; we also have to invoke the general obligation we have to keep our promises. But, importantly, we would *not* have those reasons *unless* we actually made the promises. We only have those reasons *provided that* we make the promises in question. So the mere fact that a reason would not obtain unless certain conditions obtained as well (like the act of making a promise) does not make that reason non-categorical or

179

instrumental. Promises are just such a case. They generate categorical oughts [14]

Consider again a group member's obligation to contribute to their group believing q, the normative defeater for their belief that p. Suppose that T believes N, that North Beach is the best landing zone. However, there is a normative rebutting defeater for N – S, the proposition that South Beach is best. It follows that T ought to believe S, and thus, by GMAP, someone ought to do their part in bringing it about that T believes S. If Hicks *epistemically* ought to bring it about that T believes S, then his reason to do so should be categorical. According to Carter, though, it will also be non-categorical, since Hicks would not have the reason unless he was in the group. We can see now that this is not quite right. What explains Hicks's reason is, in part, the fact that he has made a commitment with the other members of T, and they have jointly committed to planning a military mission. However, T also faces the same cognitive demands that individual agents do, to respond diligently to their evidence and properly respond to all their salient considerations, regardless of their ends. This no doubt factors into how we explain the normativity of Hicks's reason to contribute to T believing S. Hicks ought to influence T's believing S, not only because he is jointly committed with others to their mission, but also because T ought to respond appropriately to all epistemic considerations. Thus, the plural subject theory can accommodate normative defeaters, in addition to psychological defeaters.

---

[14]Thanks to Nathan Kellen and Michael Lynch for discussing this section on categorical reasons with me.

4.4 Intra-group disagreement

What can a responsibilist approach to joint epistemology tell us about more familiar problems in social epistemology? One such problem, that of peer disagreement, has engendered protracted debate. In one guise, the problem is a practical one: when I encounter an epistemic peer (that is, someone I recognize as my equal with respect to the relevant epistemic qualifications) who disagrees with me about p, what impact, if any, should this have on my attitude towards p? More specifically, what should I do in disagreements after full disclosure?

Feldman (2006) says that a disagreement after full disclosure occurs when epistemic peers have thoroughly discussed the issues, know each other's reasons and arguments, and know that they've each come to incompatible conclusions on the basis of those considerations. These disagreements contrast with, for example, disagreements where no evidence has been shared, or where the peers do not know about their disagreement. Feldman argued that such disagreements force a sort of skepticism; such disagreements demand an attitude adjustment, if not suspension of judgment, then a lowering of credence. The skeptical conclusion has attracted criticism (for instance, it relies on a view about responding to evidence that is much stronger than standard evidentialist views; see (Ballantyne, in press)). The responsibilist joint epistemology can give it a rationale when the peers in question are not isolated individuals, but are members of the same group, and the disagreement occurs in the context of their shared activity.

First, I will introduce the conceptual machinery of joint commitment, based on the work of Margaret Gilbert. This will tell us the difference between peers who

happen to find themselves in a disagreement and peers who disagree in the course of their shared activity. Then I introduce the joint commitment view of group belief, where the normative structure of joint commitment explains the normative structure of group belief. This will then inform cases of disagreement between peers in jointly committed groups, and explain what would be rational for the group to do in light of their disagreement, as well as what would be rational for the group members to do. This, I argue, helps us understand conciliatory responses to disagreement after full disclosure. One advantage to this approach, I hope, is that it does not rely on the debate between the uniqueness thesis and permissivism. The contrast between these approaches will come out towards the end.

4.4.1 Joint commitment again.

What is it for some people to do something together? Margaret Gilbert has argued that it can't be simply a matter of their forming personal commitments to do something. For us to walk together, for example, it can't be simply that I commit to the walk and you commit to the walk, because such a conjunction of commitments fails to account for the norms that govern joint activities. For instance, when we take a walk together, I cannot simply break off from our walk by taking off in a different direction. That would not simply be odd, but incorrect. This does not mean we're both locked in to the walk until it ends, of course. If I ask to end the walk prematurely, I can leave only if you grant permission. The wrongness of leaving without permission is part of the normative structure of joint activity, but it cannot be explained by appealing to no more than our personal commitments. If I con-

182

sciously rescind my personal commitment to the walk in the middle of it, then how could my conscious decision to do something else be irrational? When I commit by myself to take a walk, there's nothing *per se* irrational in deciding to do something else instead, so long as I rescind my initial commitment to walking. That wouldn't change just because someone else formed a parallel personal commitment to walking as well. What explains the norms of joint activity has to be something more than mere personal commitments.

Gilbert argues that joint commitments make joint activity possible. Our walk is something we do together because *we* commit to it. We make a joint commitment to walking, and this joint commitment isn't reducible to a conjunction of personal commitments, even if they are interrelated in complex ways. Some, like Michael Bratman, have worried that irreducibly joint commitments are primitive and, hence, theoretically unsatisfying, or explanatory dead ends. This need not be the case, however. Joint commitments share basic structural features with personal commitments, by virtue of their both being kinds of commitment. To see how, consider what joint and personal commitments have in common.

Commitments of all kinds are normatively robust. Commitments provide reasons to those who make them; making a commitment changes one's normative position. If I commit to getting groceries tonight, I thereby give myself a reason to get groceries tonight. That reason won't go away if I happen to forget about my commitment. Consider how I might chastise myself if I forgot to go shopping. Again, though, this doesn't mean commitments lock me in. It can be revoked, so long as I do so properly; ignoring a commitment won't do. I have to rescind it. Commitments get

such normative powers due to their directive-like structure. Commitments are like commands: they have authors and recipients. The author sets the agenda, which the recipient is responsible for carrying out. In personal commitments, it turns out that the author and recipient are the same person. Joint commitments are more complicated. Like personal commitments, individuals in a joint commitment issue directive-like representations to themselves, which they take up. However, they simultaneously issue those representations to each other, and take them up from each other. So, when A and B jointly commit to X-ing, A issues an "X" directive both to A and B, B issues an "X" directive both to B and A, and both A and B take up these "X" directives. The normative structure of commitment *per se* explains how personal commitments make you responsible to yourself, and how joint commitments make you responsible both to yourself and to someone else.

To see how this works in context, think about walking together again. If walking together were simply a matter of interrelated, complex personal commitments, there would be nothing incorrect about my suddenly breaking off from our walk. I might even say, "After some reflection, I'd rather walk alone. See ya!" If all I've done is make a personal commitment to walking with you, I have played by the rules. I made a commitment, meaning I issued something like a directive to myself. This makes me the sole agenda-setter and recipient, and as such, I am only responsible to myself for carrying out the commitment. I rescinded that commitment, made a new one, and acted on that instead. But if you think that I did do something wrong in defecting from our walk together, personal commitment is insufficient for explaining why. In a joint commitment, by contrast, I am not only answerable to myself, but

184

to you as well. My unilateral decision to leave the walk is not enough to dissolve the commitment; I need to get your permission, too.

4.4.2 Group deliberation and belief.

As has been argued already, groups face the same sorts of demands that individuals do, with respect to carrying out their commitments. I cannot fulfill a commitment to make croissants this evening without knowing the answers to questions like: do I have the right kind of flour? Do I have enough of it? Do I have enough time to make croissants? Groups have similar needs. If we are walking together, where are we going, and for how long? To carry out their commitments, groups need beliefs to guide their actions, just like individuals. They often form these beliefs by means of conversation. Consider a small group, such as two people, Allison and Brian, taking a walk together. Allison says to Brian, "I was thinking Horsebarn Hill would be good for a walk." Brian replies, "Horsebarn Hill it is." He could also reply, "Would it?", or, "Wouldn't the campus be better?" Depending on how the conversation develops, Allison and Brian will determine where they will go walking. This is not, of course, the only means by which groups form their beliefs. Some employ more straightforward voting procedures, perhaps using the kind of premise-driven aggregation proposed by Pettit and List (2011). For simplicity, we can focus on group beliefs grounded in conversation.

According to Gilbert and Priest (2013), what happens in conversation is something like this. When Allison asserts that Horsebarn Hill would be good, she is proposing that she and Brian collectively believe that Horsebarn Hill would be good

for a walk. She is putting that proposition out there as something for them to believe, disbelieve, or question together; she is making a move in a negotiation. When Brian responds by saying that Horsebarn Hill would indeed be good, he accepts Allison's proposal, and it becomes their collective belief that Horsebarn Hill would be good for a walk. Had he responded with less confidence or outright rebuttal, the negotiation for the collective belief would have been tougher. Allison would have to make a case for her proposal by giving reasons for them to believe, or by rebutting counter-proposals. Hence, according to Gilbert and Priest's *negotiation of collective belief thesis* (ibid., p. 14), groups can establish their beliefs by means of negotiation in conversation.

Allison and Brian might establish their group belief through negotiation, but what is the nature of such collective beliefs? Gilbert has long defended the joint commitment account of group belief, according to which a group believes that p if and only if they are jointly committed to believing that p as one body, or to emulating as much as possible a body that believes p. It may be unclear at first what work joint commitment does here, but consider how group beliefs function in contexts like Allison and Brian's. They leave for their walk, they are on their way, when Brian says, "Horsebarn Hill is ugly and unpleasant." This would likely evince shocked surprise from Allison; hadn't they just decided that Horsebarn Hill would be good? Brian's gainsaying of the group belief is just like suddenly wandering away from Allison without a word during the walk. Without qualification ("On second thought..."), it makes no sense. Brian would be violating a norm in contradicting their group belief.

186

This is where joint commitment lends some explanatory power. The norms that are violated by contradicting the group are the norms provided by joint commitment. Taking on a joint commitment means being answerable to all parties involved, not just yourself, hence the norms against unilateral defection from what the group has decided to do. If Brian gainsays the group, he not only breaks against the commitment he has made to himself, but that which he has made to Allison as well. His participation in the group is answerable to her just as much as to himself. When the group decides that Horsebarn Hill is good, it's incumbent upon both Brian and Allison to maintain that attitude together, *even if* neither Brian nor Allison personally believe that Horsebarn Hill is good. They might both be indifferent; maybe Allison suggested it for want of some place, any place, to go. But as far as the group is concerned, they do jointly believe that Horsebarn Hill would be good.

Small one-off groups like Allison and Brian provide good models for the mechanics of the negotiation of collective belief, but are not especially compelling portraits of epistemologically-interesting group reasoning. For that, consider instead paradigmatic deliberative groups, such as juries or advisory boards. These groups involve the kinds of high stakes reason-exchanging capable of pumping our intuitions.

4.4.3 Peer disagreement in groups.

Arguably, negotiations for collective belief occur in all sorts of groups, even very short-lived ones with indeterminate commitments. I will focus here on negotiations in jury deliberations. This model has its advantages and disadvantages. On the one hand, jury deliberations are familiar and have clear-cut goals with high stakes. On

the other, jury deliberation sometimes requires unanimity among the jurors, or at least, a majority opinion. It is unclear whether the kinds of group attitudes possible on our view will satisfy the court's requirement for unanimity. Even so, the example will serve as a model, even if imperfect. Other groups could serve just as well, like the EPA's various advisory boards. The special reason for looking at juries, however, is that they sometimes face deadlock.

Consider the 2016 trial in Montana of Joseph Campbell, who shot and killed Tim Newman (Brandt, March 6, 2016). Campbell told police initially that he shot Newman in self-defense. In a later interview, he maintained that Newman posed a threat to him, but changed the details somewhat, contradicting his earlier testimony. The prosecution called Dr. Werner Spitz as an expert witness, who argued that Campbell's telling of Newman's death did not match the forensic evidence. Campbell's lawyers responded by calling their own forensic expert as a witness, who concluded that Dr. Spitz's reconstruction was not the only one compatible with the evidence. With these factors to consider, and including testimony from others (including Campbell himself, and his wife), the jury had to decide whether Campbell was guilty of intentional homicide (murder in Montana law), or whether the shooting was justified by Campbell fearing for his life. The jury deliberated for 12 hours before declaring themselves hopelessly deadlocked, at which point the judge declared a mistrial.

Feldman has named disagreements with this character as "disagreements after full disclosure". Epistemic peers have such a disagreement when they form conflicting beliefs on the basis of shared evidence, and this fact is common knowledge

188

between them. That the evidence is shared is built into jury deliberations (as is their peerhood), since the jurors are participating in the trial together. The fact of their disagreement comes out through deliberation, the weighing of the evidence, and the inferences supposedly licensed by the evidence. All of these things – the significance of the evidence, the credibility of the witnesses – are common knowledge between the jurors.

Disagreements like this involve a lot of negotiation, in the sense indicated by the negotiation of collective belief thesis. For instance, one juror might propose that Dr. Spitz's preeminence makes him a more credible expert witness than the defense's, while another can reply that, however credible Spitz is, the defense made an excellent point. Some juror might argue that Campbell's story was credible, only to be met with another juror responding that nobody could reasonably take Newman to be a threat. According to the negotiation of collective belief thesis, exchanges like this constitute negotiations, involving proposals, counterproposals, and rejections. In the Campbell case, not enough people were swayed for the jury to make a verdict; they were hopelessly deadlocked. In other words, the negotiation stalled. They could not decide Campbell's guilt one way or the other, and they reported as much to the judge. Whatever the individual jurors' attitudes were, the jury seems to have suspended judgment as a group.[15]

---

[15]I have not focused much on what peerhood amounts to in situations like this. I take what I say here to be neutral with respect to views on peerhood.

4.4.4 Peer disagreement.

In the peer disagreement literature, epistemologists ask whether disagreement with one's epistemic peers on the basis of shared evidence rationally requires a change in attitude. Conciliationists say "yes": in the face of disagreement, suspend judgment or lower your credence. Steadfasters deny this, claiming that some disagreements do not require conciliation.[16] What these views have in common, however, is that they frame peer disagreement as face-offs between individual epistemic agents (or sometimes as face-offs between group agents, as in Carter's (2014) paper on disagreement between groups). But what if you and your peers are jointly committed to answering the same question? If you are a part of a group like a jury or advisory board, what, if anything, does your group's attitude have to do with your own?

The conciliationist's motivating insight is that there seems to be some pressure to change one's mind, if only a little bit, in the face of disagreement. The joint commitment view of group belief can explain this pressure. Joint commitments have a normative structure that makes jointly committed parties accountable to each other and themselves. If we are jointly committed to X-ing, then I'm responsible to you and myself for doing my part in our X-ing, and you are likewise. If we are jointly committed to suspending judgment about p, then I cannot gainsay the group attitude, at least not in group deliberative contexts. I could speak for myself, but not without preamble. Since the group view about p has not been settled, I cannot rationally take p for granted in our discussions about p. It is irrational, not unlike

---

[16]There are other mixed views, like those offered by Lackey's (2008) justificationist view or the recent virtue-theoretic treatment by Church and Samuelson (2017).

jointly committing to taking a walk together, only to wander off without warning 10 minutes later. Hence, group beliefs, or attitudes generally, provide some normative pressure on the behavior of the group's members. This on its own does not tell us much about what to do when you disagree with your group. It would only explain the social pressure you feel not to speak out, to keep quiet if you know what's good for you – hardly an epistemic consideration.

The machinery of joint commitment can do more than that, though; it can tell us how a disagreement with one's group member peers can give you reason to change your mind. Suppose I'm a juror, entering into deliberations with my fellow jurors. Having thought hard about what I've heard, I am convinced, especially by Spitz's testimony, that Campbell murdered Newman. When we start deliberating, I say as much to the group of jurors, arguing as best as I can that Spitz's testimony all but proves Campbell's guilt. Other members of the group agree, but some do not. "What about the other expert's testimony? Even if Spitz is more credible, we can't ignore the points he made." And so on. Suppose that we all do our best in our deliberation, are all epistemic peers, yet still argue to a standstill.

Now consider what this kind of disagreement means for me, one of the parties to this joint commitment. Prior to group deliberation, I did my best to form my view; I am accountable to myself for the conclusions I draw from my epistemic activities. Suppose I have formed my judgment responsibly[17]. But now, entering into deliberation with my group, I am accountable to more than myself in my epistemic

---

[17]Bracket the issue here of whether this amounts to knowledge. Instead, we will take it to amount to something like justified belief as understood by Montmarquet: S is justified in believing p insofar as S is epistemically virtuous in believing p (Montmarquet, 1993, p.99).

activity within the deliberative body; I am accountable to others, as they are to me. Let's suppose that *we* form *our* judgment responsibly (again bracketing whether this amounts to group knowledge). On the one hand, I've done my best to form my personal judgment about Campbell's guilt, concluding that he's guilty, but on the other, *we* have done *our* best to form a judgment about Campbell's guilt, and we've found ourselves unable to do so.[18] I have good reasons for thinking Campbell is guilty, and the group has its good reasons for suspending judgment. [19]

On the joint commitment account of group belief, there is no irrationality yet. Members of groups disagree with their groups all the time; indeed, it is a feature, not a bug, of joint commitment accounts that a group can disagree with all of its members. However, consider my normative situation on the Campbell jury. I am solely, personally responsible for following the evidence to the conclusion that Campbell is guilty, *and* fully responsible (with others) for following that same evidence to suspension of judgment. This falls out of our joint commitment to deliberate as a jury. Just as I am personally responsible for X-ing when I act on my personal commitment to X-ing, I am responsible still for X-ing when I act on my joint commitment to X-ing. Suppose further that, upon reflecting on what I've done in both cases, I find no mistakes in what I've done myself or what I've done with others. I even endorse both actions as instances of reasoning done well. In the case of deciding Campbell's guilt, after I find that I still disagree with the jury after full disclosure, I also find that I

---

[18]I'm not considering cases where other jurors have checked out of deliberation. Perhaps all that matters is that the main drivers in deliberation are doing their epistemic best.

[19]Note the difference between the negotiating disagreement and the standard disagreement between peers. When you and I disagree outside of joint commitment, each of us is only responsible for our own opinions. In a joint commitment, however, we will end up responsible not only for our own individual views, but for *our* view together.

cannot identify any mistakes in what I did alone or what we did together, and that I even positively endorse both lines of reasoning. So I credit myself with reaching my own conclusions prior to deliberation, but I also credit the jury for reaching its contrary conclusions after deliberation – which includes my own efforts.

Compare my circumstances on the jury with the graduate student discussed by White (2005, p.452). A graduate student who goes to MIT finds himself with epistemic standards S, which, given evidence E, lead them to believe p. However, the student realizes that, had they gone to Berkeley instead, they would have had standards S' instead, and thereby believed not-p on the basis of E. They reason further that, had this information been available to them before graduate school, and if all they cared about was knowing whether p, then the decision to go to one school or another would seem arbitrary. Cases like this, where the permissivists seem to license arbitrary choices, differ from my circumstances on the jury in a crucial detail. Through my own reasoning and my reasoning together with the other jurors, I actually come to be responsible for contrary attitudes, the forming of which I retroactively endorse. Apparently permissive cases with arbitrary choices, like White's, do not have this feature. What the student might see is that, after learning either S or S' and concluding on the basis of E either p or not-p, no mistake is made by the lights of the epistemic standards adopted. However, in this case, the student is not responsible for forming both judgments, nor are they endorsing the reasoning that brought them there. For me on the jury, I have done both. So while one might be skeptical about whether White's student needs to lose confidence about p, one might see why I'd be of two minds as a juror on the Campbell trial. I disagree with

the jury I'm on, but I also cannot disavow either path *I took* from the evidence to the judgments, especially if I take myself to be creditable for following each line of reasoning (in one case by myself, in the other case with others). This is apt to produce *some* cognitive dissonance, which can be resolved by reconciling the two views. This might only amount to my concluding that I still sort of think that Campbell is guilty, but everyone made good points in deliberation. Even so, it is a downgrade in confidence from my initial judgment, a rational response to dissonance, and the normative structure of joint commitment makes it possible.

It might still be an open question how we ought to respond to disagreement in other cases, or in general. Thus, for all I've argued, joint commitment has little to do with how to respond to disagreements after full disclosure *per se*. However, such disagreements in a joint commitment context can give us reason to be conciliatory, when we are party to the disagreement.

4.5 Conclusion

When we achieve epistemic successes as individuals, those successes reflect positively on us as inquirers. We get credit for believing responsibly. Virtue epistemologists have taken these theses seriously, developing theories of epistemic virtue to account for the underlying features of agents that enable epistemic success. One approach we have seen identifies virtues with attitudes, evaluative states directed at objects. Virtue epistemology thus explains epistemic achievements of different kinds (for instance, true belief) as resulting from an agent manifesting the right attitudes.

While this picture applies easily to individual inquirers, more must be done to show whether it applies to group inquiry. We should want to do this, since we often credit groups with epistemic successes (or failures, as the case may be), implying that groups can be responsible for these achievements. Thus, we might explain group success by appealing to epistemic virtues. Just as we credit individual successes to the agent's manifestation of virtue, we might credit group successes to *their* virtue. But where do we find these virtues at the group level?

Group responsibility depends on the joint commitments which hold the group together. What we are responsible for depends, in part, on what we have committed ourselves to. If we credit a group with success, then we hold them responsible for it, and if we hold them responsible for it as a group, there has to be a joint commitment which holds them together, which makes them jointly responsible. Among the commitments that matter here will be commitments to attitudes, the kinds which constitute virtue. We credit groups for epistemic successes when they have and manifest the right attitudes; epistemic failures accrue to failures to have the right attitudes.

Applying this understanding of group epistemic virtue is fruitful in contexts where group members disagree with their group's judgments. This kind of intra-group disagreement motivates a kind of conciliationism in particular contexts. As an individual, we might credit ourselves with reasoning well to one position, but as a group member, credit ourselves with reasoning well (with everyone else) to a contrary view. Thus, we credit ourselves with virtuous reasoning to positions which cannot be jointly believed. We find ourselves pulled in different directions, where we

195

might not in cases of one-off disagreement with a peer. In a group context, we are the locus of full responsibility for contradictory views. One rational, virtuous response to this unusual predicament is, of course, to lower our confidence in our views. So the normative pressure of disagreement in a group context is not mysterious, because it comes from ourselves. This is not a one-size-fits-all theory of the epistemological significance of disagreement, but it does capture the importance of disagreement in one of its natural habitats.

References

(1991). (1).

Aikin, S. F., & Talisse, R. B. (2013). *Why we argue (and how we should): A guide to political disagreement.* Routledge.

Alston, W. P. (1988a). The deontological conception of epistemic justification. *Philosophical Perspectives*, *2*, 257–299.

Alston, W. P. (1988b). The deontological conception of epistemic justification. *Philosophical Perspectives*, *2*(n/a), 257–299.

Annas, J. (2011). *Intelligent virtue.* Oxford University Press.

Anscombe, G. E. M. (1957). *Intention.* Harvard University Press.

Baehr, J. S. (2011). *The inquiring mind: On intellectual virtues and virtue epistemology.* Oxford University Press.

Ballantyne, N. (in press). Is epistemic permissivism intuitive? *American Philosophical Quarterly*.

Battaly, H. (2008). Virtue epistemology. *Philosophy Compass*, *3*(4), 639–663.

Baxter, D. L. M. (1988). Identity in the loose and popular sense. *Mind*, *97*(388), 575–582.

Bennett, K. (2011). Construction area (no hard hat required). *Philosophical Studies*, *154*(1), 79–104.

Ben-Yami, H. (2004). *Logic and natural language*. Ashgate Publishing Limited.

Bickle, J. (1997). *Psychoneural reductionism: The new wave*. MIT Press.

Bloomfield, P. (2000). Virtue epistemology and the epistemology of virtue. *Philosophy and Phenomenological Research*, *60*(1), 23–43.

Boolos, G. (1984). To be is to be a value of a variable (or to be some values of some variables). *Journal of Philosophy*, *81*(8), 430–449.

Brandt, A. (March 6, 2016). Mistrial declared in campbell homicide case; retrial possible. *Independent Record*. Retrieved from `http://helenair.com/news/crime-and-courts/updated-mistrial-declared-in-campbell-` $6b08 - 5002 - 9d1e - e8802db61e2a.html$ (Accessed: 05-01-2017)

Bratman, M. E. (2014). *Shared agency: A planning theory of acting together*. Oup Usa.

Carter, J. A. (2014). Group peer disagreement. *Ratio*, *27*(3), 11–28.

Carter, J. A. (2015). Group knowledge and epistemic defeat. *Ergo, an Open Access Journal of Philosophy*.

Chang, R. (2013). Commitment, reasons, and the will. In R. Shafer-Landau (Ed.), *Oxford studies in metaethics, vol. 8* (pp. 74–113). Oxford University Press.

Church, I. M., & Samuelson, P. L. (2017). *Intellectual humility: An introduction to the philosophy and science*. Bloomsbury Academic.

Churchland, P. M. (1979). *Scientific realism and the plasticity of mind*. Cambridge University Press.

Clark, A. (1994). Beliefs and desires incorporated. *Journal of Philosophy*, *91*(8), 404–25.

Code, L. (1987). *Epistemic responsibility*. Published for Brown University Press by University Press of New England.

Cohen, L. J. (1989). Belief and acceptance. *Mind*, *98*(391), 367–389.

Cordell, S. (2017). Group virtues: No great leap forward with collectivism. *Res Publica*, *23*(1), 43–59.

Cuneo, T. (2007a). *The normative web: An argument for moral realism.* Oxford University Press.

Cuneo, T. (2007b). *The normative web: An argument for moral realism.* Oxford University Press.

Effingham, N. (2010). The metaphysics of groups. *Philosophical Studies*, *149*(2), 251–267.

Epstein, B. (2015). *The ant trap: Rebuilding the foundations of the social sciences.* Oxford University Press.

Feldman, R. (2000). The ethics of belief. *Philosophy and Phenomenological Research*, *60*(3), 667–695.

Feldman, R. (2006). Epistemological puzzles about disagreement. In S. Hetherington (Ed.), *Epistemology futures* (pp. 216–236). Oxford University Press.

Feldman, R. (2008). Modest deontologism in epistemology. *Synthese*, *161*(3), 339–355.

Fitzpatrick, R., & Heiberg, J. (2007). *Euclid's elements.* Richard Fitzpatrick. Retrieved from `https://books.google.com/books?id=7HDWIOoBZUAC`

Foley, R. (1992). *Working without a net: A study of egocentric epistemology.* Oxford University Press.

Frege, G. (1979). Logic in mathematics (P. Long & R. White, Trans.). In F. K. Hans Hermes & F. Kalbach (Eds.), *Posthumous writings* (pp. 203–250). University of Chicago Press.

Friedman, J. (2013). Suspended judgment. *Philosophical Studies*, *162*(2), 165–181.

Gilbert, M. (1987). Modelling collective belief. *Synthese*, *73*(1), 185–204.

Gilbert, M. (1989). *On social facts.* Routledge.

Gilbert, M. (1990). Walking together: A paradigmatic social phenomenon. *Midwest Studies in Philosophy*, *15*(1), 1–14.

Gilbert, M. (1994). Remarks on collective belief. In F. F. Schmitt (Ed.), *Socializing epistemology: The social dimensions of knowledge* (pp. 235–56). Rowman Littlefield Publishers.

Gilbert, M. (2000). *Sociality and responsibility: New essays in plural subject theory.* Rowman Littlefield Publishers.

Gilbert, M. (2002). Collective guilt and collective guilt feelings. *Journal of Ethics*, *6*(2), 115–143.

Gilbert, M. (2006). *A theory of political obligation: Membership, commitment, and the bonds of society.* Oxford University Press.

Gilbert, M. (2007). Collective intentions, commitment and collective action problems. In F. Peter & H. B. Schmid (Eds.), *Rationality and commitment.* Oxford University Press.

Gilbert, M. (2013a). Commitment. In H. LaFollette (Ed.), *The international encyclopedia of ethics.* Wiley-Blackwell.

Gilbert, M. (2013b). *Joint commitment: How we make the social world.* Oxford University Press Usa.

Gilbert, M., & Pilchman, D. (2014). Belief, acceptance, and what happens in groups. In J. Lackey (Ed.), *Essays in collective epistemology.* Oxford University Press.

Gilbert, M., & Priest, M. (2013). Conversation and collective belief. In A. Capone, F. LoPiparo, & M. Carapezza (Eds.), *Perspectives on pragmatics and philosophy* (pp. 1–34). Springer International Publishing. Retrieved from `http://dx.doi.org/10.1007/978-3-319-01011-3`$_1$ doi: 10.1007/978-3-319-01011-3$_1$

Goldberg, S. C. (in press). Should have known. *Synthese*, 1–32.

Greco, J. (2007). Knowledge as credit for true belief. In M. DePaul & L. Zagzebski (Eds.), *Intellectual virtue: Perspectives from ethics and epistemology.* Clarendon Press.

Greco, J. (2010). *Achieving knowledge: A virtue-theoretic account of epistemic normativity.* Cambridge University Press.

Greco, J., & Turri, J. (2011). Virtue epistemology. *Stanford Encyclopedia of Philosophy.*

Heim, I., & Kratzer, A. (1998). *Semantics in generative grammar.* Blackwell.

Hindriks, F. (2013). The location problem in social ontology. *Synthese*, *190*(3), 413–437.

Hursthouse, R., & Pettigrove, G. (2011). Virtue ethics. *Stanford Encyclopedia of Philosophy*.

King, N. L. (2014). Perseverance as an intellectual virtue. *Synthese*, *191*(15), 3501–3523.

Kraut, R. (2008). Aristotle's ethics. In *Stanford encyclopedia of philosophy.*

Lackey, J. (n.d.). *Group belief: Lessons from lies and bullshit.*

Lackey, J. (2008). A justificationist view of disagreement's epistemic significance. In A. M. A. Haddock & D. Pritchard (Eds.), *Proceedings of the xxii world congress of philosophy* (pp. 145–154). Oxford University Press.

Lahroodi, R. (2007). Collective epistemic virtues. *Social Epistemology*, *21*(3), 281–297.

Linnebo, Ø. (2012). *Plural quantification.* Stanford Encyclopedia of Philosophy.

List, C., & Pettit, P. (2011). *Group agency: The possibilty, design and status of corporate agents.* Oxford University Press UK.

Lynch, M. (2013). Epistemic commitments, epistemic agency and practical reasons. *Philosophical Issues*, *23*(1), 343–362.

Mathiesen, K. (2006). The epistemic features of group belief. *Episteme*, *2*(3), 161–175.

McKay, T. (2006). *Plural predication.* Oxford University Press.

Millikan, R. G. (1984). *Language, thought and other biological categories.* MIT Press.

Millikan, R. G. (1995). Pushmi-pullyu representations. *Philosophical Perspectives*, *9*, 185–200.

Montmarquet, J. (1993). *Epistemic virtue and doxastic responsibility.* Rowman Littlefield Publishers.

Nesse, R. (Ed.). (2001). *Evolution and the capacity for commitment.* Russell Sage Foundation.

Oliver, A., & Smiley, T. (2013). *Plural logic.* Oxford University Press UK.

Olson, E. T. (1997). *The human animal: Personal identity without psychology.* Oxford University Press.

Olson, E. T. (2007). *What are we?: A study in personal ontology.* Oxford University Press.

Olson, E. T. (2014). Personal identity. In E. Zalta (Ed.), *Stanford encyclopedia of philosophy.*

Palermos, S. O., & Tollefsen, D. (in press). Group know-how. In J. A. Carter, A. Clark, J. Kallestrup, S. O. Palermos, & D. Pritchard (Eds.), *Socially extended epistemology.* Oxford University Press.

Parfit, D. (1984). *Reasons and persons.* Oxford University Press.

Pettit, P. (2003). Groups with minds of their own. In F. F. Schmitt (Ed.), *Socializing metaphysics.*

Pettit, P. (2005). Construing Sen on commitment. *Economics and Philosophy*, *21*(1), 15–32.

Pettit, P. (2007). Responsibility incorporated. *Ethics*, *117*(2), 171–201.

Quinton, A. (1976). Social objects. *Proceedings of the Aristotelian Society*, *76*, 1–27.

Railton, P. (1984). Alienation, consequentialism, and the demands of morality. *Philosophy and Public Affairs*, *13*(2), 134–171.

Railton, P. (1997). On the hypothetical and non-hypothetical in reasoning about belief and action. In G. Cullity & B. N. Gaut (Eds.), *Ethics and practical reason* (pp. 53–79). Oxford University Press.

Ramsey, W. (2008). Eliminative materialism. In *Stanford encyclopedia of philosophy*.

Ritchie, K. (2013). What are groups? *Philosophical Studies*, *166*(2), 257–272.

Roberts, R. C., & Wood, W. J. (2007). *Intellectual virtues: An essay in regulative epistemology*. Oxford University Press UK.

Ruben, D. (1983). Social wholes and parts. *Mind*, *92*(366), 219–238.

Ruben, D. (1985). *The metaphysics of the social world*. Routledge Kegan Paul.

Ryle, G. (1979). *On thinking*. Blackwell.

Ryle, G. (2009). *The concept of mind: 60th anniversary edition*. Routledge.

Searle, J. (1990). Collective intentions and actions. In J. M. Philip R. Cohen & M. Pollack (Eds.), *Intentions in communication* (pp. 401–415). MIT Press.

Sen, A. (2005). Why exactly is commitment important for rationality? *Economics and Philosophy*, *21*(1), 5–14.

Shafer-Landau, R. (2009). A defence of categorical reasons. *Proceedings of the Aristotelian Society*, *109*(1pt2), 189–206.

Sheehy, P. (2006). *The reality of social groups*. Ashgate.

Sidelle, A. (1998). A sweater unraveled: Following one thread of thought for avoiding coincident entities. *Noûs*, *32*(4), 423–448.

Sosa, E. (1980). The raft and the pyramid: Coherence versus foundations in the theory of knowledge. *Midwest Studies in Philosophy*, *5*(1), 3–26.

Sosa, E. (2010). *Knowing full well*. Princeton University Press.

Stocker, M. (1976). The schizophrenia of modern ethical theories. *Journal of Philosophy*, *73*(14), 453–466.

Strawson, P. F. (1962). Freedom and resentment. *Proceedings of the British Academy*, *48*, 1–25.

Tanesini, A. (2016a). Intellectual humility as attitude. *Philosophy and Phenomenological Research*, *94*(3).

Tanesini, A. (2016b). Teaching virtue: Changing attitudes. *Logos and Episteme*, *7*(4), 503–527.

Thomson, J. J. (1997). People and their bodies. In J. Dancy (Ed.), *Reading parfit*. Blackwell.

Thomson, J. J. (1998). The statue and the clay. *Noûs*, *32*(2), 149–173.

Tollefsen, D. (2002). Organizations as true believers. *Journal of Social Philosophy*, *33*(3), 395–410.

Tollefsen, D. (2003). Participant reactive attitudes and collective responsibility. *Philosophical Explorations*, *6*(3), 218–234.

Tollefsen, D. (2015a). *Groups as agents*. Polity.

Tollefsen, D. (2015b). *Groups as agents*. Polity.

Tuomela, R. (1992). Group beliefs. *Synthese*, *91*(3), 285–318.

U.S. Department of Justice. (2017). *Foreign corrupt practices act*. https://www.justice.gov/criminal-fraud/foreign-corrupt-practices-act.

(Accessed: 2017-04-29)

Uzquiano, G. (2004). The supreme court and the supreme court justices: A meta-physical puzzle. *Noûs*, *38*(1), 135–153.

Velleman, D. (2000). On the aim of belief. In D. Velleman (Ed.), *The possibility of practical reason* (pp. 244–81). Oxford University Press.

Webber, J. (2013). Character, attitude and disposition. *European Journal of Philosophy*, *21*(1), 1082–1096.

White, R. (2005). Epistemic permissiveness. *Philosophical Perspectives*, *19*(1), 445–459.

Wiggins, D. (1980). *Sameness and substance.* Harvard University Press.

Williams, B. (1973). Deciding to believe. In B. Williams (Ed.), *Problems of the self* (pp. 136–51). Cambridge University Press.

Williamson, T. (2000). *Knowledge and its limits.* Oxford University Press.

Wright, S. (2009). The proper structure of the intellectual virtues. *Southern Journal of Philosophy*, *47*(1), 91–112.

Wright, S. (2013). A neo-stoic approach to epistemic agency. *Philosophical Issues*, *23*(1), 262–275.

Wright, S. (2014). The stoic epistemic virtues of groups. In J. Lackey (Ed.), *Essays in collective epistemology.* Oxford University Press.

Yi, B.-U. (2014, October). Is there a plural object? In D. Baxter & A. Cotnoir (Eds.), *Composition as identity.*

Zagzebski, L. T. (1996). *Virtues of the mind: An inquiry into the nature of virtue and the ethical foundations of knowledge.* Cambridge University Press.