HOW WE DO WHAT WE DO

Joint Action and Spontaneity

ALEXANDER EMIL LEFERMAN

A DISSERTATION SUBMITTED TO

THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN PHILOSOPHY

YORK UNIVERSITY

TORONTO, ONTARIO

AUGUST 2021

ABSTRACT

The dissertation defends a novel account of joint agency, one that accommodates the neglected phenomenon of spontaneous joint action. The goals of the dissertation are to reveal the importance of spontaneous joint action, to show why these actions are problematic for many accounts of joint agency, and to produce a satisfactory theory of them. Chapter 1 argues that being capable of explaining spontaneous joint actions is in fact a requirement on a satisfactory theory of joint agency and this poses a challenge to meeting the other requirement on such a theory, the togetherness requirement. Spontaneous joint actions are those performed by co-agents who have not interacted in ways that bind them together. The challenge, then, is to adequately explain how co-agents are joined together without binding interaction. Chapter 2 reviews the literature and argues that extant theories do not meet the challenge of spontaneity because they cannot satisfy both requirements together. Chapter 3 develops the *reasons account of joint action*, which appeals to normative group reasons, in order to meet the challenge. Grasping a group reason forces agents to occupy the co-agential point of view because the group reason indicates who is to act and what they ought to do together. If two agents grasp their group reason, they are already bound together such that were they to act on the reason, they would act together spontaneously. Chapter 4 investigates which theory of normative reasons is consistent with spontaneity. Both motivating reasons and internalism about normative reasons are found lacking. Instead, it is argued that realism about normative reasons provides the best account of normative group reasons because the objective nature of real reasons eliminates the need for binding interaction, and it can more easily accommodate the inherent publicity of group reasons. Finally, Chapter 5 argues for realism about normative reasons, the existence of group reasons, and the unrestricted publicity of normative reasons. It does this by showing how these are all consequences of Davidson's triangulation argument.

For

Lori, John, Debra, and Bryan

ACKNOWLEDGEMENTS

To my mind, a good philosopher has a voice. She is able to express herself in a way that makes others see things in a new light. I certainly don't have one of those. However, I think that a condition for possessing one is to be afforded the space to pursue one's own philosophical eccentricities within boundaries that protect one from going off the rails. I am most grateful to my supervisor, Robert Myers, for guiding me by his voice while affording me the space to develop my philosophical point of view. Having the responsibility to develop in this space was a struggle, especially when I was all too often deaf to guidance. Yet with Bob's supreme patience, kindness, and commitment, the struggle was satisfying and occasionally joyful. Of course, Bob has done a great deal more for me than help me through the dissertation process. Since taking the first-year seminar at York with him, Bob has had unparalleled influence on how I think about philosophy and the questions that matter. I cannot thank him enough for being an exemplary philosopher, supervisor, and human being. I can only hope that we might speak together one day.

If there is one person who could give Bob a run for his money with respect to my philosophical development, it is Claudine Verheggen. With Bob, she co-taught the first-year seminar that set me on my way, and, through many more interventions, she has left an indelible mark on my thinking. Though I have surely fallen short in this dissertation, I am indebted to her for showing me the importance of truth and philosophical rigour.

I am thankful to A.J. Julius, Louis-Philippe Hodgson, and Alice MacLachlan for serving on my committee and being committed to this project. Without your engagement, the dissertation would pale in comparison to what it is today. I also wish to thank Paul Hurley, my external examiner, Brian Singer, my internal examiner, and Michael Giudice, the chair of the examination. I am especially appreciative of Paul's discerning comments and questions. And I must reserve special thanks for A.J. Beyond serving on my committee, his generosity and demandingness during my visit

to UCLA were invaluable to the development of this project. Because of our discussions during that time, I came to see the practical potential of not only this project but of philosophy in general.

I will always be grateful for being able to spend these years in the congenial environment that has been cultivated by the philosophy department at York University. I feel very lucky to have had teachers, who were always available and encouraging, and to have studied with my fellow graduate students, who were committed to fostering a philosophical community of equals. In particular, I have learned so much from Olivia Sultanescu, Sam Steadman, Ben Winokur, Ali Karbalaei Mahdi, David Rocheleau-Houle, and Bogdan Florea-Alexandru, but, most importantly, I am grateful for their friendship.

I cannot thank my wife, Debra, enough for having the strength to bear the burden of being married to a philosopher and for enduring a philosopher like me, in particular. And I thank my family for their constant support despite the fact that I refused to listen to them and become a tradesman who would always have work. Though I now possess a PhD, there is time yet for me to become a plumber.

TABLE OF CONTENTS

INTRODUCTION

There is no question that philosophy, at least since Descartes, has treated the individual's perspective as basic. How do *I* think? What can *I* know? How should *I* act? Throughout the 20th-century, things did not really change despite much more consideration of the social bases for various phenomena. Action theory is one example: philosophers were, and still are, primarily concerned with what made an *individual's* action intentional as opposed to unintentional. Indeed, in spite of the rise of the study of social phenomena, philosophers were rarely concerned with the social unless it helped to shed light on the individual's perspective. And even then, it was the last recourse. In the philosophy of language, Wittgenstein taught us that an individual's meaningful utterance depends, in some way, on social practice. The purpose of appealing to the social, after many failed individualist attempts, was to get at the individual perspective. However, more recently, there has been pushback against this individualist presumption, or what I'll call the primacy of the individual. There appear to be irreducibly social phenomena, joint actions, such as walking together, social objects, such as money, and, more generally, social practices themselves which may ground these other social phenomena. That these social phenomena appear to be irreducibly social is reason enough to at least unsettle the primacy of the individual in philosophy. Philosophical inquiry about them will get nowhere if these phenomena are indeed irreducibly social and it refuses to consider some kind of social perspective. This dissertation is concerned with one type of social phenomenon, joint action, and it aims to show that we fail to do justice to it if we do not take on a social perspective, what I call the *co-agential perspective*. The presumption of the primacy of the individual will cause us to fail to meet our duty in philosophical inquiry: to do justice to the phenomenon of joint action.

Consider a simple case of joint action, that of walking together. The first thing to notice is that this action cannot be done on one's own. I cannot go for walk together by myself—I must walk together with another. The second thing to notice is that walking together is not the mere

aggregation of individuals walking on their own. The fact that two people are walking side by side and at the same pace does not add up to their walking together. In such a case, neither of the individuals intend to walk with the other, it is simply a coincidence. Even though an outsider might judge them to be walking together, that is, jointly, they each know that they are walking alone. Something more is needed than the aggregation of two or more individuals' perspectives and actions to get joint action. Perhaps the individual perspective can be maintained so long as the individuals intend to act jointly. This is a real possibility that moves away from carving up the world in strictly individualist terms: the contents of their intentions are not for individual actions but *joint* action. The social makes an appearance, though I will show this tack to be insufficient. Once we take the social seriously, we start to see joint actions all over the place: we move couches together, ride the subway together, cook dinner together, and, despite often speaking passed each other, do philosophy together. We certainly perform many actions as individuals, but we also do many things intentionally together.

For the past thirty years or so, philosophers have been offering theories attempting to do justice to the phenomenon of joint action. The exemplars of the phenomenon on which philosophers focus, walking together, painting a house together, executing a passing play together, all involve some degree of planning or agreement. To execute the passing play, we draw up the play and assign roles. To go for a walk together, we agree to head out in a half hour. Indeed, it seems that in order to do something together, we must *already* have done something together. Though this seems suspiciously circular, it is difficult to see how agents are in a position to think that they are acting with another without having already done something together. If we haven't agreed to go for a walk together, why would I think that I am walking *with* you rather than merely side by side, coincidentally? However, there seem to be many joint actions that do not fit the exemplary mold. Consider these examples. First, I come upon someone pushing her car on her own and I start

pushing the car with her—we pushed the car together. Second, while we are having a picnic in the park, it suddenly begins to pour cats and dogs. Immediately, you pack up the food, I grab the blanket, and we rush to the car. We packed up the picnic together.[1] Third, I am walking on the sidewalk heading towards the subway station. Another person is heading in the opposite direction. There is also a very large man, Shaq, who, all of a sudden, trips and needs help getting to his feet. As he is very large, we both rush over and help him up together. Lastly, some younger boys in the schoolyard who are playing notice that their bully is walking towards them. Typically, they each think that it is easier to accept his abuse because they cannot as individuals do anything to stop him. However, today, one of the boys stands up to him, in spite of his individual inability. All of a sudden, the others join in and, together, they put an end to their being bullied. In each of these cases, the agents who acted together did not plan, agree, or do anything relevant together before acting together. These joint actions were *spontaneous*. They acted together immediately, or without having done something to bind each other together. They did not act as individuals. They sprang into joint action—they acted together spontaneously.

One aim of this dissertation is to argue that theories of joint action that are constructed for the exemplary cases cannot explain these special spontaneous cases. Moreover, the main aim of the dissertation is to show that these spontaneous cases of joint action are not simply special cases, but that spontaneity is a central to joint agency in general. If a theory cannot explain spontaneous joint actions, then something has gone seriously wrong—the phenomenon of joint action has been done an injustice. In short, the reason for this is that spontaneity holds a central place amongst the web of essential concepts involved in our thinking about joint action. I'll argue in Chapter 1 that the essential conceptions that constitute the web are togetherness, spontaneity, rational agency, and

[1] Both of these examples come from John Searle's insightful paper (1990). One of its most important insights, in my mind, is recognising this kind of joint action.

viewing another as a partner or co-agent. To omit one of these skews our understanding of intentional joint action and of the essential concepts themselves. This may seem odd given that, as I claim, philosophers have not noticed how important spontaneity is to joint action.[2] My diagnosis of this blindness is that philosophers, whether consciously or not, are guided by the individualist presumption. Starting from the individual's perspective, a theory must show how agents bridge the gap between them in order to be joined together, and, in the spontaneous cases, nothing happens that would constitute the necessary bridging. The primacy of the individual blinds us from seeing the possibility of spontaneous joint action because it assumes that we must build a bridge.

This might seem to imply that the solution is to make the collective, as opposed to the individual, primary or basic. If agents are somehow *already* bound together, then there is no burden of bridging the gap between agents. But if this is true, we might wonder how individuals can think and act as *individuals*. Abandoning the primacy of the individual and putting in its place a totalising collective does not help us because it must be possible for us to act both as individuals and jointly. Not only does our experience weigh heavily in favour of this consideration, the very idea of joint action makes no sense without the idea of individual action. In order to individuate joint actions, we must be able to differentiate them from individual actions. Joint action does not take up the entire category of action, so we must be careful to make space for both individual and joint action. This poses a particular problem: how can we think of collectivity in a way that permits an explanation of spontaneous joint action without entirely vitiating the individual perspective? What I contend is that we, as rational agents, have the capacity to occupy different perspectives, some of which are individual and at least one of which is collective, the co-agential perspective. It is a fact about us that we can think and act individually as well as jointly. The task is to show how this is possible.

[2] Of course, some have noticed spontaneous cases and thought of them as an extensional bonus of the theories. In particular, (Searle 1990; Tuomela 2005; 2009).

Before describing the method and outline of the dissertation, I want to put two distinctions on the table to hopefully provide clarity and to focus in on the precise phenomenon under investigation here. I have already introduced a portion of the first distinction under the banner of the primacy of the individual. The distinction is that between individualism and holism.[3] In the social sciences generally, there is a principle called methodological individualism, which roughly says that a social theory should only consider the actions and motivating states of individuals in the explanation of social phenomena.[4] There are a number of consequences incurred by adopting this principle. Ontologically, for instance, there is no overarching entity constituted by the joined-together agents, such as a party, group, or collective. In the two-person case of going for a walk together, individualism would say that there are two agents who bear a special relation, that of being joined together. It is the relation between them that allows us to credit their action as a joint action. But there is never more than the individuals, and simply because agents act together does not mean that they constitute a greater ontological entity. Or, semantically, a theory only appeals to the distributive 'we', which refers to an aggregate of individuals.[5] On the other hand, there is holism or nonindividualism, which includes irreducibly social entities into its theorising. Ontologically, a holistic theory may include super-agents or group minds that stand over and above the individuals who constitute them. Or, semantically, it may use the collective 'we', which refers to a unified group of agents, as opposed to the distributive 'we'. Individualists often complain about parsimony when holists invoke irreducibly social entities, but holists complain that an adequate explanation requires these holistic notions.[6] I raise this distinction because it is helpful to see that some theoretical consequences are incurred by the approach that one adopts. But, more importantly, I raise it

[3] Gilbert uses these terms to discuss different approaches to social theory (1989).
[4] For a helpful summary of methodological individualism, see (Heath 2020). To see some of the nuances and consequences of the principle, see (Stoutland 1997; Schmitt 2003b; 2003a)
[5] Michael Bratman is the main proponent of this semantic consequence (2014).
[6] For instance, Searle and Bratman are very worried about group minds and unnecessarily inflating our ontology (Searle 1990; Bratman 2014).

because Annette Baier has challenged holists in that they still approach joint action, sometimes seemingly unwittingly, from the perspective of individualism.[7] Despite some theories including holistic entities and language, the prominent theories of joint action all begin with individuals and attempt to build up from their perspective and attitudes. Consequently, jointness is a construction. Baier, I think rightly, complains that an explanation of jointness is never really in view when we theorise in this way.[8] Jointness cannot be constructed, and especially not spontaneously. Looking ahead, the theory that I offer attempts to do justice to Baier's complaint. One of the upshots of the co-agential perspective is that it provides a means of accounting for joint agency without construction. The benefit of avoiding construction is that spontaneity seems much more attainable.

The second distinction that I wish to flag is related to the first, in particular, with respect to the ontological consequences. The second distinction is that between joint agency and group agency. It is not uncommon to run these two ideas together. Certainly, when we act together, we might refer to ourselves as a group, we might call our action a group action. But groups or institutions are ontological entities that are distinct from the particular members who constitute them.[9] In joint agency, two or more particular people are joined together. There are a number of differences between joint and group agency. For one, a group's members can change over time, but the group remains the same. The supreme court loses and gains justices every so often and yet *the* supreme court does not change—the court that decided 100 years ago is the same court as today. On the other hand, if I am walking with *you*, I cannot be engaged in *that* joint action if someone else takes your place. Furthermore, it is not implausible for an institution to survive without any members at all. This is simply not a possibility for joint agency. Another important difference is

[7] (Baier 1997).
[8] I should note that I believe that Baier conflates two distinct notions, the social and jointness. Not all sociality is joint sociality. Nevertheless, I think that the concern stands.
[9] (List and Pettit 2011). Group agency, what the group believes or does, is often considered a function of its members.

that a group may act when only *one* of its members acts. For instance, the philosophy department might award an essay prize to a student and yet that decision was made by the chair alone. The whole idea of joint agency is trying to understand how two or more agents are bound together in a way that they can act together. It is impossible for a joint action to occur as the result of only one person's behaviour. Though it is common to run these two ideas together, it is important to keep group, corporate, or institutional actions distinct from joint actions. At the very least, it would require a significant argument to show that they are the same. The phenomenon at issue in this dissertation is *joint agency* and not group agency.[10]

The dissertation is divided into five chapters. The first poses a challenge for philosophical inquiry about joint agency. I call it the challenge of spontaneity, which says that an adequate theory of joint agency must meet two requirements, the togetherness and spontaneity requirements. The togetherness requirement says that a theory must explain how agents are bound together so that they count as acting together. The spontaneity requirement says that a theory must explain how agents can act together immediately or without binding interaction. The spontaneity requirement is justified by the web of essential concepts discussed above. What is challenging about these two requirements is that any seemingly available answer to one or the other requirement seems to make meeting the other impossible. For instance, if to be bound together, we must come to an agreement, then a form of binding interaction has bound us together and, so, spontaneity is impossible. The challenge of spontaneity threatens the very idea of joint agency. In Chapter 2, I survey the joint action literature and show how prominent theories fail to meet the two requirements. In Chapter 3, I offer an account of joint action that meets the challenge of

[10] Sometimes in the literature, 'collective action' is used for 'joint action'. I will keep these ideas separate as well, except when discussing the first-person plural pronoun because, at this point, the distributive-collective distinction is a convention. The reason I keep these two notions separate otherwise is because collective action sometimes refers to parallel action, two coordinated actions that are not joint. Indeed, when we talk about collective action problems, we often talk about parallel action and not joint action.

spontaneity: the reasons account of joint action. This theory relies on normative reasons for joint action. The reasons account meets the challenge because if agents are bound together in virtue of sharing a normative reason for joint action and responding to it, then they are joined together and can act jointly without binding interaction. What is required is for agents to be able to respond to their reasons, in this case, their reasons for joint action. The reasons account raises two serious questions that I take up in Chapters 4 and 5, respectively. First, what must the nature of normative reasons for joint action be like for them to serve in the explanation of spontaneous joint action? I show how relying on motivating reasons or adopting internalism about normative reasons would require the violation of the spontaneity requirement. Instead, it is argued that realism about normative reasons is the best theory because realism is committed to the mind-independence of normative reasons and can easily accommodate the inherent publicity of reasons for joint action. Second, do normative reasons for joint action actually exist and, if they do, how are they public? I provide an argument that argues not only for realism about normative reasons but for the existence of *reasons for joint action* and for their unrestricted publicity, that is, for the claim that their normative force extends to everyone in some way. The argument is a development of Donald Davidson's triangulation argument. I argue that the conclusions I draw are implications of the triangulation argument. Roughly, in order to make sense of our normative thought, normative reasons must exist, and we acquire normative thought by triangulating on normative reasons. Moreover, the act of triangulation only makes sense if there are reasons for joint action and they are unrestrictedly public. In the end, joint action is at the very heart of our cognitive and practical lives.

I want to conclude with two points, one, a brief warning and, two, a longer methodological point. First, Chapter 5 shows how Davidson is at the center of this dissertation. However, Davidson is an influence throughout, from the conception of agency in Chapter 1 to the discussion of desire and its relation to normative reasons in Chapter 3 and, obviously, the triangulation

argument in Chapter 5. The warning is that this Davidson is not the Davidson with whom you may be accustomed. On the theoretical side, it is commonplace to think of Davidson as an 'interpretationist' because of his work on radical interpretation. Interpretationism very roughly says that an agent's meaning is determined by being interpreted by another who is guided by her own language and standards. However, Davidson is not an interpretationist: for him, the world plays a very important role in the determination of meaning—indeed, it is the third apex of the triangle. This is what makes Davidson a realist about normativity. On the practical side, it is extremely common to think of Davidson as a champion of reductionism, of the great causal mechanism, of the disappearing agent. One could be excused for thinking this had he only published "Actions, Reasons, and Causes". However, his later work shows how woefully incorrect is the attribution of reductionism to any area of his philosophy. Davidson's philosophy, as the triangulation argument shows, is systematic and thoroughly non-reductionist—its purpose is to make sense of meaning and normativity in a non-revisionary fashion. So, I implore the reader to not imbue how they view what I say about Davidson with any of these incorrect ideologies about Davidson that unfortunately continue to float around contemporary philosophy. To see the implications of his later thought and how the argument helps to think about joint agency, the light must be raised so that the shadow figure of 'Davidson' vanishes. A subterranean aim of this dissertation is to show how illuminating Davidson's thought, especially his later thought, still is.

Finally, I wish to mention a methodological point in the hope that it helps to clarify how I proceed in Chapter 1. One might complain that I am not sensitive enough to some of the major disagreements in the literature on joint agency as I formulate the challenge of spontaneity and, thus, that the conceptions of the central concepts in joint agency, as I describe them, need not be accepted from the get-go since they fall on a side of those disagreements. After all, I describe these central concepts as they strike me for their intuitiveness and for their conceptual adequacy when

limited to the phenomenon of joint action. However, this dissertation is an investigation into the very foundation of joint action and, by the end, of normativity. The reason why I believe that it is important to start in this manner is to see where these intuitive conceptions lead and, as I argue, they appear to lead to a tension at the heart of idea of joint agency. But a tension that has been left unacknowledged. The task of the dissertation is to problematise the tension and formulate an adequate solution to it. Nevertheless, because this is a foundational inquiry, we should look at these concepts anew, and we should not begin by presupposing conceptions of these ideas that are already tempered as responses to other problems in joint agency. How could we determine what is adequate to explain a phenomenon if we do not have some idea of what is to be explained? A simple example of the kind of disagreement that I am talking about here is the choice between utilising the collective or distributive 'we'. The collective 'we' is the intuitive notion; however, it seems to come along with objectionable ontological commitments. Consequently, philosophers who are convinced of these ontological worries choose to employ the distributive 'we'. Decisions like this, though, are responses to problems and tensions in our conceptions of the matter at hand. That is, when we come to such tensions and problems, we have a few theoretical options. First, we may decide to *revise* our intuitive conceptions as a response to the relevant problem that arises because it is theoretically profitable to do so. The intuitive conception is conceptually inadequate due to certain of its entailments. But it must be acknowledged that a revision of a fundamental and intuitive conception bears a cost, namely, the cost of having to think of this phenomenon in different terms and losing a grip on what appeared to be an essential feature. For example, with respect to abandoning the collective 'we' in favour of the distributive 'we', the unity that is implied by the collective 'we' is lost. The best revisions try to account for this loss by trying to capture that essential feature in some other way. Nevertheless, the cost of revising may be too steep and the theory ultimately loses sight of or never reclaims a grip on the original phenomenon.

Second, a different response to tensions and problems may be warranted if we exhaust all the revisionary options. We may respond as sceptics when no adequate account can be provided of a phenomenon. If we are convinced that intuitive conceptions involved in the phenomenon result in an irresolvable tension and none of the revisionary conceptions get an adequate grip on the phenomenon, then the responsible response to it is to withhold judgment. For instance, if no good solution can be found for the challenge of spontaneity, then it is warranted to be sceptical about joint action in general. The sceptic does not need to deny the appearance of a particular sort of phenomenon. She does not even deny the phenomenon's existence—she is not a nihilist. What she objects to is the affirmation of its existence. Without an adequate theory, we are not in a position to claim that the appearance of the phenomenon *is* indeed the phenomenon. The sceptical response is a thoroughly philosophical response to a tension. The sincere sceptic is still someone who is rapt in wonder because she is still very much engaged with the phenomenon. With the most fundamental phenomena, she sees that which others claim to exist. She cannot deny that things *seem* to be the way that others claim them to be. She simply believes that there is no good reason to believe things to stand in that way. The sceptical attitude, though, requires that she engage with every argument that arises and, because of this, she must maintain her wonder in order to be committed to engage with these new arguments. The question posed by the phenomenon is always alive for the sceptic.

However, there is a third response to a philosophical problem: pluralism. Unlike scepticism, pluralism, in my view, is rooted in an anti-philosophical temperament—it buries wonder alive and refuses to acknowledge its very audible screams. When our intuitive and revisionary conceptions of a phenomenon fail to give an adequate, unified account, the pluralist proclaims that there are two or more kinds of some thing. For instance, in this dissertation, a possibility is that individual and joint agency just are different kinds and are yet both expressions of agency. It is not that these kinds are species of some genus that unifies them. It is more that we call these kinds by the same name, but

they are not the same kind of thing and there is no way to unify them. But this isn't quite right either because pluralism, as a response to a philosophical problem, is not merely a matter of polysemy. Rather, the kinds are conceptually distinct and yet, for some reason, they fall under the same umbrella term. Unlike scepticism, which defers a positive verdict while keeping wonder intact, pluralism aims to squash wonder by giving a positive verdict that pleases nobody. The pluralist insists on an answer without offering any explanation of the phenomenon at all. This is why the pluralist has an anti-philosophical temperament: carrying the weight of wonder and its attendant task of explanation tires him out and makes him cynical—what is an explanation, in his eyes, is *in fact* no explanation at all. Unlike the sceptic, the pluralist is a cynic about the task of philosophy.

Provided that my assessment of pluralism is correct, it must be avoided at all costs. In a way, it does not provide an answer to the question it allegedly answers. An account that invokes pluralism is inadequate. I wish to make two final remarks about this methodological point, which was that we must begin the inquiry from the intuitive conceptions of our phenomenon and see where they lead. First, pluralism should be kept distinct from non-reductionism, which simply claims that one thing cannot be reduced to another. For instance, the fact that the normative cannot be reduced to the non-normative does not necessarily imply pluralism. Though there is a disunity in non-reductivism, it is unclear that the two or more things that cannot be reduced ought to form a unit. Showing how two things are related is different from showing that they form a unity. With pluralism, on the other hand, the assumption of unity is always in the background. Second, and more importantly for what comes, it is very much possible that revisionary conceptions of some phenomenon entail or unwittingly espouse pluralism. With an eye to what will come, being revisionary and adopting the distributive 'we' may entail two conceptions of rational agency that cannot be squared with each other. There is the rational agency that I engage with as an individual and there is the rational agency that I engage with when I act with another. Of course, it is not

necessary that revision leads to pluralism, and there can be adequate revisions. But if a revision does lead to pluralism, then the revision is inadequate. From my point of view, one must allow wonder to inspire for as long as one can feasibly manage. Trying to hold on to the intuitive conceptions in the phenomenon of joint agency is a commitment to wonder and explanation. Perhaps, in one's time, one is able to produce one piece of the immense puzzle which philosophy aims to solve. One places that piece on the table, awaiting others to adjoin their pieces. In time, we might solve the puzzle together but if we do not, the joint endeavor of philosophy will continue to be spurred on by wondering about how what appears to us to be the case makes any sense at all.

Aristotle famously tells us that philosophy begins in wonder—the state of being ignorant of why things appear the way that they do. It is wonder and its attendant feeling of perplexity that compels us to find an answer, to quelch wonder by understanding. But sometimes our response is too eager and fails to do justice to the human experience of wonder. In Greek, the word for *wonder* is related to *fear*. Lacking an account and being ignorant can be a terrible thing. And that fear may cause us to answer our problem too hastily in order to escape it and feel momentarily secure in the world. One of the existential problems of philosophical inquiry is how to continue to bear the weight of wonder, to continue acknowledging one's ignorance and do justice to the phenomenon at hand, without succumbing to the terror of simply not knowing and by providing an inadequate account. Pluralism is a way of succumbing to the terror of a seemingly unintelligible world: like an outworn veil that offers the illusion of solace, it shelters us in our ignorance. However, in my view, a philosophical orientation is a commitment to do justice to the experience of wonder and resolutely answer its call by giving an account of the world. Admitting that there just is no single, unified account is no real shelter from our ignorant condition.

CHAPTER 1

Joint Action and the Challenge of Spontaneity

The final scenes of the political masterpiece *The Battle of Algiers* depict a special kind of joint action:

> One day in French occupied Algeria, for no apparent reason, Algerians gather in the streets chanting "Long live Algeria!" and carrying flags made of sheets and rags, representing the independent nation that they demand. The immense crowd, composed of seemingly every native Algerian, bears down on the colonists. As they rush past guards, they are shot at. In spite of meticulous French intelligence, the narrator tells us that there was no known reason for the demonstration that day, and even the exiled head of the FLN was clueless. By all accounts, Algerians woke up one morning to demonstrate together for their independence, at all costs.[11]

The Algerians' demonstration is not only an instance of a joint action. What makes it special is its *spontaneity*. Seemingly from nothing, a large number of people began demonstrating together for their sovereignty and freedom. Their action is unified by a singular goal and yet it is not the product of some plan or prior coordination. As the movie presents it, the idea of their independence moved them into the street in order to achieve their goal together. It might seem implausible that such a large joint action could take place without planning or the like, but implausibility does not rule out the possibility of spontaneous joint action.[12] What makes their protest spontaneous is that they are joined together in action without some prior binding interaction of the sort that allegedly and typically joins agents together, such as mutual promises, expressions of willingness, uttering conditional intentions, or planning. Indeed, what we must wonder is how the protesters are, in a sense, already bound together when the first person steps out into the street to protest with others such that, when the others join her, they protest together spontaneously. How could agents be joined together in this spontaneous manner?[13] As I use it, 'spontaneous' does not signify speed,

---

[11] This is my description of the final scenes of Pontecorvo's film.

[12] A.J. Julius points out that even on accounts involving something like planning, such as Bratman's or Velleman's, such large-scale actions are too cognitively demanding (Julius 2023, chap. 9; Bratman 2014; Velleman 1997). One attempt to address large-scale actions of this sort is (Shapiro 2014).

[13] Of course, we need not begin with such a political or massive example. But to ask, just as in this case, how two agents end up spontaneously pushing a car together or spontaneously helping up a fallen individual together, *how* they are spontaneously joined together is the focus.

naturalness, or thoughtlessness, such as when a driver spontaneously turns the wheel to avoid an unseen pedestrian or the heart's spontaneous beating. Instead, what I mean by 'spontaneous' corresponds to a sense of *immediacy*. The Algerians were not joined together quickly or thoughtlessly and then demonstrated in the street together. Their demonstrating together was immediate in that their being bound together did not depend on their having already done something together, such as agree, plan, or promise. In spontaneous joint action, agents are bound together immediately.

Explaining how exactly this happens is not the goal of this chapter. My answer to that will come in Chapter 3. Instead, this chapter has two primary aims. The first is to show that spontaneous joint action is not a peripheral type of joint action and in fact occupies the core of the phenomenon. In virtue of this, there is a requirement on theories of joint action to explain *spontaneous* joint action. I call this the 'spontaneity requirement'. The second aim of the chapter is to argue that the spontaneity requirement seems to be in tension with the standard requirement on theories of joint action, that is, what I call the 'togetherness requirement'. To meet the togetherness requirement, a theory must explain how agents are bound together such that they can count as acting together intentionally. This is the traditional task upon which theories of joint agency have focused. The spontaneity requirement, on the other hand, has not been addressed, even though examples of spontaneous joint action occasionally appear.[14] The spontaneity requirement, as I characterise it, says that a theory must explain how agents can act together intentionally *immediately* or *without prior binding interaction*. I argue that these two requirements are in tension because many, if not all, ways in which the togetherness requirement may be satisfied rule out the satisfaction of the spontaneity requirement. If agents must, for instance, agree, promise, or express readiness to each other, then they will have engaged in some form of binding interaction. This is because it is the interaction of coming to agree, exchanging promises, etc., that binds them together. But this is

[14] (Searle 1990; Tuomela 2009, 299; Satne 2020, 6).

precisely what is ruled out by the spontaneity requirement. I do not mean to suggest that people cannot act together after they agree, etc., but by making binding interaction *necessary*, we seem to miss out on this core feature of joint action. On the other hand, without some form of binding interaction, we are left wondering how agents who are acting together spontaneously are *in fact* bound together such that they count as acting together at all. By all appearances, when we accommodate spontaneous joint action, it seems that we lose a grip on how agents count as acting together. And when we specify how agents are bound together, we seem to lose the ability to explain how spontaneous joint action is possible. The tension between these two requirements is no small matter because it poses a challenge to the very idea of joint agency.

The chapter is structured into six sections. In section one, I flesh out the togetherness requirement with the aim of clarifying the notion of *togetherness* at work in joint action. The idea of *being-partnered* is at the centre of this notion. In order to elucidate the ideas of *being-partnered* and *togetherness*, the second section develops, what I call, the co-agential perspective. To be partnered with someone requires that we occupy a perspective where we view each other as co-agents or, rather, those *with* whom we engage in agency. In this section, I distinguish the co-agential perspective from other perspectives from which we may view other people and from other seemingly similar concepts in the joint action literature. The third section presents an independent argument, 'the agency argument', in favour of including the co-agential perspective in our thinking about joint agency. The main thrust of the argument is that the co-agential perspective allows for a unified account of rational agency across singular and joint agency. The fourth section characterises the relevant sense of 'spontaneity' at issue in this dissertation. Next, in section five, I argue that so long as an explanation of how co-agents legitimately occupy the co-agential perspective and, so, can exercise agency together is available, then an explanation of spontaneous joint action is available. Moreover, because both the central notion of being-partnered and that of spontaneous joint action

are conceptually related via the co-agential perspective, we should view spontaneous joint action as central to the phenomenon of joint agency. This interrelated web of concepts establishes the spontaneity requirement on theories of joint agency. In the final section, I argue that there is a tension between the two requirements and offer a suggestion on how to resolve it. This tension poses a challenge because if a solution is unavailable, then we must revise at least one of the basic conceptions involved in the idea of joint agency: togetherness and being-partnered, spontaneity, or rational agency. The trouble with a revisionary, or even a sceptical or pluralist, response is that it requires us to change how we think of joint agency, and this change may clash with our experience in practice. Ultimately, a revisionary, sceptical, or pluralist response can be avoided so long as we alter how we think about joint agency.

1. The Togetherness Requirement

The primary requirement of a theory of joint action is to explain the joint sense of *together*, that is, how individuals are bound together such that they constitute a unit that acts together as opposed to individuals acting as a disparate aggregate.[15] This task has been notably difficult but at least one thing is clear: when individuals act together, they use the first-person plural pronoun to refer to themselves as the agent(s) of the joint action. So, a natural, initial account is to say that individuals are acting together when they refer to themselves with 'we'. But this is insufficient. The first-person plural pronoun is used for many purposes and often not in the sense that expresses the joint sense of *together*. For instance, a queen might use 'we' to refer only to herself, and she alone cannot act together. Similarly, Russell, Frege, and Wittgenstein might accept the claim "we are philosophers", which refers to all three of them. But the mere fact that a statement (truly) predicates the same thing of a set of individuals does not bind them together in the joint sense, even though it classifies them together. The problem here is that the pronoun 'we' is ambiguous. For instance, if Russell were to

[15] I do not intend any substantive ontological reading of the term 'unit'.

say "We are writing on names" while referring to all three of them, then this sentence has both a distributive and collective or communal reading.[16] The distributive reading is true in virtue of an aggregation of true sentences about some set of people. If Russell, Frege, and Wittgenstein were all writing on names independently, then Russell still spoke truly. But once again, the distributive 'we' does not track the joint sense of *together*. The distributive reading is true, even when they write nothing together jointly, insofar as they each write on names.[17] One of the guiding intuitions in joint action, as I understand it, is that to act together is to act *with* others intentionally in the *same* action.[18] In the Russell-distributive case, there would be three distinct papers instead of one. Seeing as the distributive reading of we-sentences sheds no light on how individuals are acting with each other in the same action, it does not get a grip on the joint sense of together and, hence, joint action.[19]

One must be careful to avoid thinking that just in case agents bring about the same event by the same type of action that they have acted together jointly. When agents perform the same *type* of action that produces one result, it does not mean that they are bound together in joint action. Consider two people walking side by side on the sidewalk heading to the subway station. They are walking at the same pace, neither pulls ahead nor falls behind. They are going to the same place and arrive at the same time. They even acknowledge each other and account for each other's movements, one adjusts her step in a way that offers space for the other because another person is coming towards her. Moreover, someone from a window above peering down at them might be inclined to think that they are going somewhere together. But they are not walking together. What

[16] Kirk Ludwig favours calling such a reading collective in opposition to the distributive (Ludwig 2014, 113), while Mathias Haase calls such readings communal (Haase 2012, 248-252).
[17] When I use 'together jointly' following an action verb, I use it to emphasise the joint sense of togetherness. 'Jointly' is redundant when following 'together', as I mean it here.
[18] What will count as being a part of the same action will depend on the description of that action: if we are making dinner together, we do not need to chop the same carrot at the same time.
[19] I note some caution here. There are prominent theories of joint action, for instance, (Bratman 2014), which are only committed to a distributive reading of we-sentences. Nevertheless, they typically have more to say about how the joint sense of together is captured while relying on the distributive 'we'.

the walkers are doing can be described as acting in parallel. In cases like this, agents act in the same ways, perhaps even cooperatively, resulting in a single event, but they do not act together. Here, the walkers arrived at the subway station, but, despite appearances, they did not do it jointly. Other examples of parallel action are cars stuck in a traffic jam on the highway or traders acting for their own ends on the stock market. In those cases, agents are not making money together or creating a traffic jam together in the joint sense. As individuals, we often contribute to a particular event intentionally and yet we do not act together intentionally. Joint action is more than parallel action and even cooperative parallel action.

Distinguishing joint action from parallel action gets us closer to the joint sense of *together* and coming back to Russell's statement will help us get there. His statement "We are writing on names" also has a collective reading where they jointly write *one* paper about names. Here, there is no aggregation of independently true individual statements which results in a collective reading. It would be false for Russell to say "we (collectively) are writing on names" if he were writing a paper on names independently of Frege and Wittgenstein's collective efforts. The point is that collective we-sentences refer to a unit of individuals who intentionally perform one action with each other.[20] For example, if "we are dancing the tango" is true, the two dancers must be dancing *with each other* in a way that counts as dancing the tango, since each of them could not dance the tango individually. When tango dancers dance, they dance the tango jointly. Notice how joint action that evokes the joint sense of *together* is distinct from parallel action, which permits distributive readings of 'we'. Though the two people heading to the subway, from the previous example, are acting in parallel, it is a different thing for them to be walking together to the subway. For instance, there might be an expectation to wait for one if the other must stop suddenly to tie her shoe. In parallel action, such expectations are absent. If one walker gets ahead of the other and happens to catch the earlier train,

[20] Kirk Ludwig makes this point well (Ludwig 2014, 114).

their action did not fail because they were never acting together.[21] In the joint example, the collective 'we' can be legitimately used while in the latter, parallel case, it cannot. For a theory of joint action to satisfy the togetherness requirement is for it to elucidate how two or more people can count as intentionally performing one action with each other—that is, can count as acting together jointly. This does not mean that they must be doing the same type of action, as parallel actors often do, but there must be a true intentional description under which what they do can be described as acting together. The collective sense of we-action sentences puts the joint sense of *together* front and centre. These sentences are true if and only if co-agents are acting together.

There is some debate about whether the collective 'we' should be used at all since it may have unattractive consequences, such as imposing a vicious circularity or entailing a dubious group mind.[22] Despite these worries, investigating the different senses of the first-person plural pronoun has been helpful since it allowed us to home in on the special social bond at issue in joint agency—the joint sense of *together*. This bond is characterised as thinking and acting with others in the same action. But 'with', just as 'we' and 'together', is ambiguous since it can be read in a distributive and collective way. For instance, it can be appropriate to say "I am riding the train with strangers" and not be involved in a joint action. Regardless of which reading of the first-person plural pronoun a theory adopts, we can identify the joint sense of *with* so long as the agents share a particular state of mind. Many philosophers appeal to similar ideas in order to elucidate the difference between joint action and parallel action: for example, being partnered, engaging in practical intersubjectivity, and treating another as an intentional co-participant.[23] These conceptions are ways in which agents who act together *view* and *treat* their fellow co-agents. For instance, an Algerian protestor does not think

[21] I do not want to say anything substantive about normativity and joint action. However, it is not prima facie unreasonable to think that, while doing something together, each co-agent has some minimal expectations with regard to the completion of their joint action.
[22] (Petersson 2007; Searle 1990).
[23] These ways of clarifying the idea are from (Gilbert 1989, 155–57; Kutz 2000b; Roth 2003; Bratman 2014, 48).

of her fellow protestor beside her as just some other person who is voicing similar concerns as she does. She does not view her fellow protestor primarily as another individual who stacks up her dissent on top of her own. Rather, she views and treats her fellow protestor as a partner who is demonstrating *with* her. Consequently, and generally, by viewing another agent as a partner, I think of us as bound together, as joined. What she does matters to me and what I do matters to her. My co-agent who I view as a partner is not merely another person whose actions affect me, but rather someone to whom I am bound. This is true even in competitive cases of joint agency.[24] Consider two boxers who, despite having the aim of knocking the other out, view each other as partners with roles to play in their joint activity. It is still possible to view a competitor as a partner so long as the joint activity is governed by an overarching aim, such as to fight. In competitive cases, there are still rules or constraints that agents joined together must recognise insofar as what they do counts as joint activity rather than parallel action.[25]

A theory of joint action must meet the togetherness requirement and to do this, it must show how agents are in fact partnered such that they can perform an action together. Of course, *conceiving* of another as a partner is insufficient to bind me to her—I am not her partner merely because I think that I am. A theory of joint action must explain how we may move from merely thinking that we are partners to the truth of those thoughts. However, what is not made clear in theories that appeal to these decisive conceptions is what it means to conceive of other agents as partners. Note that this is distinct from explaining what *in fact* makes us partners or what binds us together. The telling feature between joint and parallel action, that is, conceiving of each other as

[24] One difficulty in the joint agency literature is how to draw the line between competitive or antagonistic cases. At what point does antagonism rule out jointness? The perspective that I develop in the next section goes some way, I believe, toward drawing the line. If the activity in which both agents are engaged does not permit each to view the other as a partner, then they cannot be acting together.

[25] Looking ahead, spontaneous competitive joint activities will be rare, though not impossible. For instance, two children can spontaneously begin a game of tag and one starts chasing the other. The infrequency of this kind of joint action is simply a product of the absence of mutually acknowledgeable overarching aims that make sense of the competition. But we need a theory on board to see how this kind of joint action is to be explained.

partners, has received, if anything, minimal treatment. For instance, Margaret Gilbert explains our being partners in terms of our being jointly committed to a goal and Michael Bratman tells us that we have a specific set of attitudes whose contents interlock and mesh in appropriate ways that count as engaging in shared agency.[26] These descriptions of viewing others as partners seem to run two separate ideas together: that of *being* bound together and that of *thinking* of being bound together. Though it seems innocuous, *thinking* of others as partners requires attention. Take Gilbert's characterisation: if I am not in fact party to a joint commitment and thus have no obligations to others, what does it look like to think of others as partners when the commitment that would structure my thinking and action is absent? The danger of ignoring this special point of view on others, especially before developing a theory of joint agency, is that it may tell in favour of a way of thinking about joint agency. To glean this point of view from the status of being bound together may, depending on the theory, make us miss important features of the natural, intuitive way of thinking about the phenomenon of joint agency. Indeed, the idea of viewing others as partners may be another site for revision in the theory of joint agency. The mere fact that viewing others as partners tells the difference between joint and parallel action recommends that it needs a characterisation before the development of a theory.

2. The Co-Agential Perspective: Viewing Another as a Partner

The idea that, when we act together, we view and think of each other as partners implies that co-agents occupy a peculiar perspective from which they view each other.[27] What exactly is the perspective that is occupied when we view each other as partners? I call this the 'co-agential perspective'. There are three aspects of the co-agential perspective. The first is that, from this

[26] (Gilbert 1989; 1990; Bratman 1993; 2014)

[27] One might worry that in larger joint actions, as in my Algerian example, we cannot possibly think of each co-agent in particular. This problem of knowing our co-agents in this way is an issue for most theories of joint agency, so it is not my problem alone. However, I think it is suitable to think of your co-agents in these large actions generically. For instance, the Algerian protesters joined in action think of *Algerians* as their co-agents.

perspective, co-agents view each other as partners in action. To view another as a partner just is to occupy the co-agential perspective. To elucidate the perspective, though, it will help to distinguish it from other ways of viewing other people. One way that we might view people is from a moral perspective where we conceive of them as, for instance, being owed respect, having dignity, being responsible agents that one can trust, etc.. Viewing people from the moral perspective constrains how we think about and interact with them. For instance, from this perspective, I should acknowledge that I should not verbally abuse a stranger because she is owed respect. Likewise, from the moral perspective, I expect that my fellow Sunday afternoon strollers will not suddenly attack me because I trust that they are responsible, decent human beings. So long as I view others in this way, my patterns of thought and action with respect to them are morally constrained. However, even though the moral perspective forces me to conceive of people with a moral valence, it does nothing to guarantee that I think of them as partners in action.[28] Typically, partners in action will be treated with respect, they should not be coerced or deceived, but the moral perspective does not require that I think of other agents as my partners. The moral perspective is consistent with a world where there is only singular action, that is, actions undertaken by individuals independently. Each of us could have our own islands, be self-sufficient, and still view others as moral agents.

Another, rather different way to view others is to view them as barriers or conduits to your practical projects, and in general as observable evidence for your projects. I call this 'the exploitative perspective' because, from here, agents are viewed primarily as resources to be taken advantage of.[29] Imagine that you are late for a meeting and you need to get on *this* train but, as it is rush hour, there is a throng of people in your way. In this case, you view others as frustrations to your end, they are barriers to your doing what you intend to do. From this perspective, you might decide to push them

[28] Some readers might think that morality is a partnered endeavor and so the moral and co-agential perspective are not distinct. If this were true, and it would need quite an argument, it would not affect my overall argument.
[29] To exploit someone as a resource does not necessarily imply that she is treated immorally or disrespected.

aside in order to get on the train or you might come to terms with your failure and impending lateness. The evidence gleaned from the throng counts against the achievement of your end. Nevertheless, others need not be viewed as barriers from this perspective, for they may be viewed as conducive to your projects. For instance, a receptionist can tell you where the doctor's office has moved when you arrive for your appointment. In this way, the receptionist is a source of information that facilitates the achievement of your goal. The information that she offered is exploited by you for your projects. What distinguishes this perspective is that others are viewed not primarily as agents but as bits of the world that need to be negotiated or accounted for in one's own projects—they are viewed primarily as resources. Just as the road sign can inform your plans and projects, so can people. When viewing people strictly this way, I need not respect their intentions because, whether I do or not, I can use them as evidence for how best to achieve my ends. Because of this, this perspective is distinguished from the moral perspective. It is also distinguished from the co-agential perspective because viewing others in this way primarily excludes their agency from your projects. To view others in this way is the furthest thing from considering them as partners.

To view others from the co-agential perspective as distinguished from these two perspectives is to view them as part of a project that we collectively aim to complete.[30] From this vantage point, other agents and their actions and thoughts are not primarily resources or evidence

[30] At this point, one might think that the co-agential perspective is similar to Strawson's participant attitudes (Strawson [1962] 2008). Being opposed to objective attitudes, as I understand them, participant attitudes are attitudes that one takes towards another that treats them as mattering to or bearing on the attitude or its content. But these attitudes are consistent with the moral perspective described above and, so, have a far wider application than the co-agential perspective. Participant attitudes do not isolate the co-agential perspective. Moreover, I want to resist the thought that perspectives are reducible to attitudes. I think of perspectives along the lines of Nagel's 'points of view' or 'perspectives' where, depending on one's perspective, certain things *matter* more or less (Nagel 1986, Introduction). The rough idea is that the same attitude can be held from different perspectives and, depending on the perspective, certain inferences are warranted, and others are not. For instance, if you are tying your shoe beside me on the platform and I believe that you are doing so, what follows from this belief depends on the perspective I take towards you. From the moral perspective, I also believe that I shouldn't unnecessarily interfere with your tying. From the exploitative perspective, I might infer that I should not walk in your direction since you may hinder my activity. And from the co-agential perspective, I infer that I should wait for you until we are both ready to board the train. It doesn't follow from the other perspectives that I should wait for you to finish tying your shoe before I board the train. So, the very same attitude has different inferential relations depending on the perspective taken.

for my project, they have an integral agentive role in its success. For instance, the stranger who provides me with directions to the town centre is not viewed by me as a partner in the joint action of reaching the centre—she merely facilitates the achievement of my end. Likewise, I do not use my partner's attitudes as evidence for the judgment that I have a higher probability of bringing about some event, our joint action, over which I do not have full control. In both examples, we are not on equal footing as agents—you are not viewed as part of my project or, rather, *our* project. For us to be co-agents, you cannot simply provide me with evidence for my project—I must view you as an agent who is capable of and responsible for doing her part. You are not a barrier or a conduit to *my* project—we are agents together in *our* project. Indeed, the co-agential perspective forces me to view myself in a different manner, that is, as your partner in action. I am not merely acting to achieve my end. I am acting to achieve our end. Insofar as our end structures my practical activity, I am not alone in determining how our end is to be achieved. And because of this, I am, in a sense, beholden to you when occupying the co-agential perspective. When it comes to our end, you can make a claim on me, and I can make a claim on you.[31] I can expect things of you given that your activity is structured by our end. This expectation is not a prediction about your future activity; I expect you as an agent to be responsive to our end and, in virtue of that, to me. The co-agential perspective forces us to view each other and ourselves primarily as agents who are rationally constrained to some degree by our joint endeavor.

The second aspect of the co-agential perspective is that each co-agent views each other's *actions* as contributing to their joint effort. When we are in the process of acting together and, hence, occupying the co-agential perspective, what you do as a part of our joint action is not viewed as just someone else's doings. Instead, it is a contribution to our action, a part of the whole joint endeavor. Others' actions, even though they may facilitate our end, do not constitute a part of our joint action.

[31] This claim is neutral regarding normativity in joint action (Gilbert 1990; 2008; Bratman 1999c; Tuomela 2005).

For instance, we aim to give our friend a gift together, so we decide together what to get her, I go online and make the purchase while you prepare the card. Each of these actions partly constitute *our* joint action. But the delivery person who facilitated our gift giving by dropping off the gift at my house did not participate in our joint action. One way to think of the delivery is as an addendum or supplement to our joint activity. Despite our joint action depending on the delivery, it was not a *part* of our gift giving. The delivery person was not intentionally giving a gift with us. To return to the main idea, from the co-agential perspective, when you think of my action of ordering the gift, it is not an addendum to our gift giving but a proper part of it. When I occupy the co-agential perspective, I view your actions as constituting a part of what we are doing.[32] When we occupy the co-agential perspective together, we view each other's actions as parts of our single joint action. One way to put this point is that I see *our* agency in your actions, and vice versa.

Nevertheless, there is an interesting point to raise here given that actions are 'worldly' entities: one might think that the actions of co-agents provide evidence for their partners. And, in a way, they do: you might see that the gift has been delivered when you are over at my house. But that is not your primary orientation towards my actions. You do not need the arrival of the gift to serve as evidence to support your card buying, that is, to do your part. Instead, you expect that the order will be placed, and you can perform your part as you see fit. When occupying the co-agential perspective, we already think of each other's actions as practically related to and ordered towards our joint end. We may see each other acting and have evidence of our accomplishments, but the evidence is not necessary and, occasionally, would even frustrate the success of our endeavor.

The final aspect of the co-agential perspective is that it uses the collective sense of the first-person plural pronoun. Since I do not view my partners primarily as evidence and their actions are

[32] Naturally, I don't think of everything that you do as contributing to our action, but only the necessary parts. Sometimes joint actions take time and people do other things in between. Typically, I do not worry about your other doings unless they get in the way of the success of our action.

not viewed as addendums to my own, I do not view them as practically distinct from me, with respect to our joint action. When I refer to us by using 'we', I do not think of us as 'I' plus 'you' because the distributive 'we' would divide our actions. I would not think of your actions as practically bearing on my own, rather they would function as evidence, barriers, or conduits for me. For instance, when we are dancing the tango, your step is not something of which I, first, take account and, then, to which I react. Rather, I expect your step to go where it should, according to our end of dancing the tango. Of course, in joint action, each of us can fail to perform the expected part actions and we will need to adjust and to take it in stride in order to succeed. Further, I do not view you, my dancing partner, in the same way as the other tango dancers on the dance floor. With them, we may coordinate our movements so that we give each other adequate space and do not interfere with each other's activity. But, importantly, I do not respond to another dancer's misstep in the way I respond to my partner's misstep. I will not adjust my dancing because a dancer across the room continues to step on his partner's feet. Nevertheless, I would adjust my dancing if other dancers would crash into us, but this is because I am predicting that they will do so—I treat their behaviour as evidence. However, if you are stepping on my toes, I might suggest how to dance with me or adjust my steps to accommodate yours. And you should also recognise that, for us to dance the tango successfully, you should find a way so that these missteps do not occur. The general point is this: when we engage in joint activity, we (collectively) are doing something together and I cannot think of you as merely another person with whom I am coordinating. This would be to treat you primarily as offering evidence for my own behaviour. I am not merely making my actions and attitudes harmonious with yours, we are doing and thinking jointly.[33]

[33] Characterising the co-agential perspective as necessarily involving the collective 'we' seems to rule out many accounts of joint action or shared agency, namely, those that depend on the distributive 'we'. In the next chapter, I will show how prominent distributive accounts fail to adequately explain spontaneity. The basic point is that, even though we may coordinate our attitudes in a way that allegedly constitutes shared agency, I must be treating you primarily as an evidentiary source in order to coordinate my own attitudes. The appeal to treating another as a partner is not enough

These three aspects, viewing others primarily as agents, viewing their actions as part of a joint action to which I am in part contributing, and referring to my partners and I together with the collective 'we', constitute the co-agential perspective. This perspective helps us draw the line between joint and parallel action because when we are only acting in parallel, we do not occupy this perspective together. In parallel action, I may treat you as a moral agent, and also primarily as a bit of evidence. For example, when you and I are independently on our way to the train and happen to be side by side, we do not view each other as partners, though we may respect our respective goals that we predict the other has. When we act jointly, on the other hand, we view each other as partners in action from the co-agential perspective. I view you as a part of our agency, and I do not treat you primarily as evidence for my project—our agency is joined. There are two points in need of emphasis here. The first is that the co-agential perspective goes beyond mere respect for others, it goes beyond the moral perspective. It seems necessary that to be your partner, I must also respect you in the way that morality requires because, in a sense, I view you and your actions just as I view myself and my own. *We* are equals when it comes to the co-agential perspective. But insofar as the moral perspective is concerned, we are not on an equal *agentive* footing with respect to our actions. Respect, and the moral perspective, is often a prohibitive or exclusionary orientation towards others and their actions—'do not interfere' is the watchword. The co-agential perspective, on the other hand, is a participatory or inclusive orientation towards others and their actions. To leave your partner alone with her actions is, in some sense, a failure from the co-agential perspective. You have drawn an unwarranted inference. Consider the scenario where we agree to play chess, but I simply watch you set up all the pieces, respecting your activity. It would not be odd for you to think that I am exploiting you since we agreed to do this together. The necessary inclusiveness of the co-

because, as I have been arguing, to treat another as a partner is primarily to view their activity and thinking as non-evidentiary.

agential perspective is why it cannot be reduced to the moral perspective.[34] The second point follows from the inclusivity point and, I believe, aligns with a consideration forcefully raised by Annette Baier. She complains that theories of joint agency are committed, either explicitly or implicitly, to the primacy of the individual over the social.[35] A theory built around the co-agential perspective avoids this commitment because occupation of the co-agential perspective, which enjoins 'joint-thoughts', is the starting point for thinking about joint agency. When we occupy this perspective, we do not pit ourselves against each other; I do not exploit you or try to predict how your thoughts and actions will align with my own activity. We consider ourselves joined together. Indeed, we employ the collective 'we'. We do not add up each other's agency, arranging it in a particular way, and call it 'joint'. The co-agential perspective puts the social first because it requires us to think of ourselves as part of a collective. As a note of caution, these comments would not completely satisfy Baier since it is still an open question how we are joined together and that may involve an individualist approach. However, I think that the co-agential perspective starts us off on a 'social' footing that makes it possible to avoid Baier's complaint.[36]

Being merely a perspective, though, we cannot possibly be joined together by simply occupying the co-agential perspective. That is, we are not *in fact* partners just because we are thinking of each other as partners. This implies that there are two ways of occupying the co-agential perspective that do not amount to being partnered. First, the perspective can be occupied individually, alone. For instance, days or even weeks before the uprising, an Algerian can think of demonstrating with her compatriots, thinking of them as partners in action, while no one else is thinking about protesting. This kind of occupation is more than mere fantasy or hypothetical

[34] If there is more substantive normativity involved in joint agency than mere instrumental normativity, then there may be interesting ethical dynamics in joint agency to explore, and perhaps the co-agential perspective can be employed to think through certain ethical problems, one of which is the prisoner's dilemma.
[35] (Baier 1997). Sometimes this is called Methodological Individualism, but Baier believes even those who deny Methodological Individualism, such as Gilbert, give the individual primacy.
[36] The theory developed in Chapter 3 should satisfy Baier on this point.

calculation. There may be good reason for Algerians to occupy the perspective together, and it just so happens that she is the first to recognise this fact. In light of this, it involves at least some degree of commitment or readiness to act as partners. The occupation of the perspective does not require that those about whom you think also occupy the perspective with you. Second, two people may independently occupy the perspective and still not be partnered. Consider this example. Two high school students have a crush on each other, yet, importantly, both are ignorant of that fact. In class, they each think about going on a date with the other, about going to a movie together, one picking up the other, buying popcorn, and deciding where to sit. They are independently and merely coincidentally thinking about the other as a partner. In these moments, each does not view the other primarily as an evidentiary source and their actions would constitute a joint action if they were partners. In this example, the features of occupying the co-agential perspective are present, but clearly the two students are not joined together. Each has no reason whatsoever to think that the other is thinking the same thing as she is—they are not in a position to draw conclusions that are sufficient for joint action. Although they are not alone in occupying the co-agential perspective, as is the budding Algerian protestor, the students, as two distinct, unbound agents, form a set that is unified only by the fact that they happen to think the same thing. Hence, not only is thinking individually about another as joined and partnered insufficient to be in fact joined together, so is the fact that your potential partner is thinking the very same thing as you. Something more than being in the same state of mind is required for joint agency.

Though it illuminates how to draw the distinction between joint and parallel action, one conclusion to draw from the co-agential perspective is that it alone does not satisfy the togetherness requirement. The fact that we can occupy the co-agential perspective in these two ways without being joined implies that there is further a legitimate or warranted occupation of the co-agential perspective. The high school students are illegitimately occupying the co-agential perspective, as is

the Algerian protestor. This does not mean that they are doing anything wrong in thinking of the other from this perspective. But it does mean that they are not in fact joined together and, because of that, they would be wrong to draw conclusions concerning their joint action from the co-agential thoughts that they have. For instance, it is not warranted for one student to show up at the other's house with the expectation that they are going to the movies and it is not warranted to gripe about the other's unpreparedness.[37] These conclusions, or at least ones like them, are warranted if and only if agents are in fact joined together. Legitimate occupation is a third way of occupying the co-agential perspective, one in which we are in fact partnered.

Before moving on, it is helpful to illustrate how the co-agential perspective is distinct from other similar concepts in the joint action literature. I will focus on two of them, namely, the application of Anscombe's concept of practical knowledge to joint action and Tuomela's we-perspective.[38] Recently, there have been attempts to show how an Anscombean action theory applies to joint action.[39] Whether or not this extension of Anscombean ideas is successful, the fundamental concept of *practical knowledge* may appear to play a similar role as the co-agential perspective.[40] Because there is significant disagreement about how to extend the account, I will describe it only to the extent necessary. Anscombe's basic idea is that when an agent is acting intentionally, she has a special kind of self-knowledge that allows her to answer the relevant 'Why?' question.[41] In the individual case, an agent *knows* what she is doing while she is doing it, and this does not involve self-prediction. The agent acting intentionally does not need to view her arm moving towards the glass in order to know that she is reaching for the glass. In a sense, she knows

[37] In the case of dating, consent is needed to go on a date together. This renders spontaneous dating basically impossible.
[38] I will take up Tuomela's account in more detail in the next chapter, but here I want to focus on how the co-agential perspective differs from his main concept.
[39] (Seemann 2009; Laurence 2011; Satne 2020).
[40] For a dissenting view, see (Blomberg 2018).
[41] (Anscombe 2000).

it immediately rather than observationally. When it comes to intentional action, evidence does not help to determine whether an action is intentional. Further, if the agent is asked why she is moving her arm towards the glass, she can answer by placing the movement in an intelligible instrumental order—'in order to take a sip of water'. She may even go further and say, 'because hydration is good'.[42] What is important is that practical knowledge is non-evidential or non-observational. Extended to joint action, agents are acting jointly together intentionally if and only if they have practical knowledge of what they are doing. Somehow, each co-agent has immediate knowledge of the other's action without observation. This is the main hurdle that the extension of Anscombe's account to joint action faces: if we are making dinner together, how can I *know* without observation that you are chopping the onions and so that we are making dinner together? It is the focus on knowledge that distinguishes Anscombean practical knowledge of joint action from the co-agential perspective. The co-agential perspective does not necessarily involve knowledge (or warranted beliefs) about the other: I can think of being someone's partner in action without *actually* being her partner. The high school crush example shows just this. In order to occupy the co-agential perspective, it is not required that I know much about the person with whom I think of myself as partnered, let alone knowledge of her mind and actions. Again, there are legitimate and illegitimate occupations of the perspective and how that is understood remains to be explained. The distinguishing feature between the two concepts is that merely occupying the co-agential perspective does not involve immediate knowledge of another's intentions and actions.

Second, the co-agential perspective may appear to be similar to what Raimo Tuomela has called the 'we-mode', 'we-perspective', or 'we-attitudes'.[43] These are to be distinguished from the I-

[42] Anscombe seems to believe that the instrumental A-D order is intelligible only in light of some normative/evaluative property (2000, sec. 33). Recently, this requirement has been abandoned, see (Thompson 2008).
[43] (Tuomela and Miller 1988; Tuomela 2005; 2007). His is the most detailed version of what has been called a "mode account" (Schweikard and Schmid 2013). Tuomela develops what was nascent in work by Sellars and Searle (Sellars 1980; Searle 1990). Here, I use we-mode, we-perspective, and we-attitude interchangeably.

mode, I-perspective, and I-attitudes. Importantly, both I- and we-attitudes can have the exact same content, but the we-mode is supposed to capture a common sense understanding of being part of a group and the "social glue" that holds it together.[44] The I-mode does not have this social nuance, though it can have social content.[45] Using this paradigm to explain joint action, agents must have the relevant we-attitudes, namely, we-intentions to act together; in doing so, they occupy the we-perspective and think in the we-mode. Singular actions are rationalised by I-intentions, so, by extension, we-intentions are needed to rationalise joint action. The sense of 'we' involved here is the collective 'we', like in the co-agential perspective. But what is the difference between the co-agential perspective and the we-perspective? The difference between the two is that we-attitudes are attitudes that have peculiar social conditions for their possession whereas the mere occupation of the co-agential perspective does not. According to Tuomela, a person possesses a we-attitude if and only if three conditions are satisfied: (1) an agent has an attitude with content *p*; (2) she believes others in the group have the same type of attitude with content *p*; and (3) she believes that it is mutually believed that group members have the relevant contentful attitude.[46] It should be clear that the possession of we-attitudes are committed to a more robust sociality than is the occupation of the co-agential perspective. Indeed, to differentiate the possession of an I-attitude, which does not explain joint agency, from the possession of a we-attitude, Tuomela must include some further social beliefs since the content of the attitude alone will not explain the difference. To occupy the co-agential perspective, on the other hand, agents do not need to believe that another person has the same social thoughts to be in that state of mind. The budding Algerian protestor who is the only occupant of the perspective may be perfectly aware that no one else yet thinks of her as a partner in future protest. Because Tuomela is trying to capture the 'social glue' already involved in social

[44] (Tuomela 2002, 19).
[45] On Tuomela's account, the content of the attitudes is not what plays the social role; it is the mode of the attitude.
[46] (Tuomela 2002, 23)

behaviour, it is natural to think that agents share beliefs about each other. Instead, the co-agential perspective characterises a way in which agents can think of each other as a pre-condition for social behaviour. This is why it is better to think of the co-agential perspective as a form of thought rather than as a type of attitude.[47] Because of this, the co-agential perspective is less committal from a theoretical perspective than are we-attitudes. However, nothing has yet been said about what makes co-agential thoughts, inferences, and expectations legitimate. What the co-agential perspective does, however, is illuminate what it is to think of another as a partner, whether or not I believe that the other also has the same thoughts about me

3. The Agency Argument for the Co-Agential Perspective

Though the previous section proceeded largely by phenomenological description and conceptual analysis, it is the appropriate methodology when trying to make sense of an intuitive idea, viewing another as a partner. In this section, I want to provide an independent argument in favour of the co-agential perspective, and one that lends support to its essential aspects. It is not merely that the co-agential perspective helps us draw the line between joint and parallel action, though that is of utmost importance. It is that it preserves the unity and continuity of the idea of agency across singular and joint agency. This claim might sound similar to the 'continuity thesis' espoused by Michael Bratman.[48] His characterisation of that thesis states that joint or shared agency can be explained using only the resources available to singular agency. Bratman's continuity thesis is committed to methodological individualism. It is evident that the agency argument cannot aim to prove the same thesis since it supports the inclusion of the collective 'we' in the co-agential perspective. The collective 'we' is categorically not a resource of singular agency. But why should

[47] An agent can have an attitude whose content contains other people and, depending on the perspective she takes towards them, she can draw certain practical inferences. See footnote 20 for a description of why the co-agential perspective is more like a form of thought rather than an attitude.
[48] (2014, 3–9).

singular agency be given this preferential position when it comes to agency in general? The historical contingency that singular agency has been given more comprehensive treatment does not speak in favour of its preferential position.[49] Instead, one consequence of the agency argument is that, by adopting the co-agential perspective in thinking about joint agency, the general contours of agency are preserved in both singular and joint agency. There is no need for two senses of agency, one that applies to individual agency and another to joint agency, or a revision of our natural ideas.

The agency argument has seven premises. First, if rational agency in joint action cannot be expressed on the basis of binding interaction, then the co-agential perspective is required for joint action. Second, intentional action expresses rational agency.[50] 'Rational', here, does not mean all-things-considered or structurally rational. Instead, when I act intentionally, I act for a reason.[51] This means that there is a standard by which I successfully do what I intend to do. To act intentionally requires the possibility of failure, which in turn requires a standard by which the action is measured. Moreover, acting for a reason means to be guided towards or committed to the action. The standard does not merely adjudicate success from above or externally; it compels me to meet it. When engaging in rational agency, I can fail to do what I intend but, in virtue of having the intention, I am (in some degree) normatively pulled to succeed. If I fail to meet my standard of action, then I am in some sense answerable or responsible for failing. It is these features of compulsion and answerability that make agency *rational* rather than mechanical.

Third, rational agency does not depend on the world, in a qualified sense.[52] This premise does not deny the world a role in rational agency. Of course, I collect information and make

---

[49] Again, Annette Baier pushes this concern with some force (Baier 1997).
[50] Some representatives of this general claim are (Anscombe 2000; Davidson [1963] 2001; Bratman 1987).
[51] I leave aside the issue of whether agents can act intentionally for no reason, as Anscombe suggests (2000, sec. 17). I believe that what I say here is still agreeable to those who deny that reasons are required for intentional action.
[52] One might worry about this division between agency and the world. Aren't I and my rational powers a *part* of the world? That is certainly true. However, how to distinguish mind from world, internal from external, etc. is a fraught issue. What I mean by world, here, is the 'external' world or the world that is engaged predictively.

predictions about the world in order to decide what to do. It is prudent to attend to the world because it can facilitate or frustrate my actions. In general, the world provides me with the materials, including my historical and social background, that delimit the extent of my agency. Nevertheless, the world does not establish the standard for my action—my power for rational agency does.[53] In the first place, for example, the fact that there is no milk in the fridge might count in favour of going to the store but, equally, I can ignore it. The world does not determine or cause which intentions I form, though it certainly limits my options. In the second place, I may not get milk at the store because I get into an accident on the way. Here, the world stops me from satisfying my intention but it does not determine whether I maintain the standards that I set for myself. Instead, my rational agency depends on my powers *as an agent*. *I* decide whether to go buy milk, not the world.[54] I form my intentions no doubt on the basis of evidence collected from the world, but whether I set myself a standard for action depends on me, not the world. Note here that my intention's standard is distinct from normative reasons which may populate the world along with other facts. Successfully acting as I intend to do is different from acting as I ought to do. Forming my intentions on the basis of my normative reasons is up to me as an agent. It is another exercise of my power of rational agency: I choose, in some sense, which consideration becomes my operative reason. I determine my intentions and can even insist on acting contrary to my normative reasons. Rational agency, which is expressed in intentional action, is exhibited when I determine for myself what I intend to do and this power, in a crucial, qualified sense, does not depend on the world. Fourth, joint action is intentional.[55] The fact that a joint action is an action rather than merely an

[53] The non-evidential, non-predictive character of agency is highlighted by (Anscombe 2000; White 2018; Julius 2023).
[54] I have in mind here something like Velleman's discussion of settling (Velleman 1989; 1997). More detail on settling, deciding, etc. is unnecessary for my purposes here.
[55] This premise will be borne out by an intelligible account of joint agency, one that I offer in Chapter 3.

event depends on there being an intentional description of it.[56] Fifth, and consequently, joint action expresses rational agency.

At first glance, this consequence does not appear too radical. But given that it entails that, sixth, the rational agency in joint action does not depend on the world, in the qualified sense, it has some serious implications for joint agency. Consider two friends who bump into each other on the street and agree to go have coffee together. One asks the other, 'Shall we go have coffee?', and the other responds, 'Yes'. Let's assume that this exchange, which is a binding interaction, establishes their intention to act jointly. That they are bound together intentionally depends on the fact that their question and answer is used as evidence of each agent's willingness and for the formation of their intentions to have coffee together. Each does not form the intention to have coffee together without the binding interaction. And it does not seem possible to have relevantly bound intentions without the exchange. This point requires some attention. A person independently, and without binding interaction, may have an intention to have coffee with another, but the intention, on its own, does not join them together. Even if two people have intentions to have coffee together, without binding interaction, it does not add up to joint intention. There needs to be something that binds these intentions and agents together in order for the rational agency in their joint action to be expressed. Without a binding element, there is no joint agency and, indeed, no rational agency expressed in joint agency. But if the binding element is a kind of binding interaction, then joint agency, unlike singular agency, depends on the world, in the qualified sense. In order to act together intentionally, we must exploit each other as evidence rather than treat each other as agents—indeed as co-agents. Importantly, this requires us to view our partners in action from two incompatible perspectives, from the co-agential perspective and the exploitative perspective. Seventh, the result of this is that rational agency in joint action cannot be expressed on the basis of binding interaction.

[56] (Davidson [1963] 2001; Anscombe 2000).

We are not agents together in action if we treat each other primarily as evidence that authorises the relevant intentions, and binding interaction serves just that purpose. We must view each other as partners in order to exercise rational agency in joint action. Thus, the co-agential perspective is required for joint action.

One might think that this argument depends on a sleight of hand in the move from individual to joint expressions of rational agency. Why can each co-agent not establish their intention to act with another, hence, exercise her power of rational agency independently, and then they act together jointly? Indeed, they both share the relevant intentional descriptions, so why do their intentions not express rational joint agency? The problem is explaining how their intentions are joined together. Just because I independently act from an intention with joint content, such as 'I intend to have coffee with you', does not make my action part of a joint action or an expression of joint agency. On the basis of this intention, if I go to have coffee and you are absent (i.e., you never intended to be there), I have only acted as an individual, never as a partner in a failed joint action. Partners' intentions must be joined together in some way in order for their actions to count as part of a joint action and to express rational joint agency. Suppose that you and I each form these intentions independently and then go and ask each other to have coffee. Here, we would have exercised our individual rational powers and subsequently joined our intentions together. In this way, we have not exploited our partners in the exercise of our rational agency because, allegedly, we *already* exercised it before binding our intentions. This two-step model of joint agency, where first we independently exercise rational agency and then bind our intentions together, cannot accurately describe what goes on in joint agency. The question is whether we can distinguish the exercise of rational agency from the binding element in joint agency. So, what role does the subsequent binding interaction have? We have already established that without some binding element, two or more intentions are not joined together. The reason is that we are not in a position to believe that we are

partners without some warrant—this is the purpose of binding interaction. Granting that some binding interaction must be necessary, the intentions prior to binding must only qualify as wishes or hopes to have coffee together. And wishes and hopes are not the motivating stuff of joint action. The interaction warrants us to *intend* to have coffee together. On this model, binding interaction has a necessary rational role in how co-agents determine their intentions. It only makes sense for me to set myself the standard to act together with you insofar as I exploit you as evidence rather than treat you as a partner. Once again, if rational agency is only expressed individually and then constructed into joint agency, then it forces us to occupy two incompatible perspectives on our alleged co-agents. To avoid this conclusion and preserve the continuity of rational agency, we must be able to view each other from the co-agential perspective, and legitimately so, before engaging in joint agency. Without the co-agential perspective, from where we can exercise rational joint agency, I would always exploit my fellow agents as evidence, and they would never truly occupy an agential position with me. We would be agential solipsists, merely piling our individual actions on top of each other.

4. Characterising Spontaneity

Not only does the co-agential perspective make sense of the crucial notion of being partnered in joint action, but it also preserves a general account of rational agency, one that does not give the world an intention-determining role. To the main point of this chapter and in addition to helping us distinguish between joint and parallel action, the co-agential perspective helps to make sense of spontaneous joint action. In this section, I characterise the concept of spontaneity. There are a few uses of 'spontaneous' that are not applicable when talking about spontaneous joint action. The first usage relates to speed or a momentary impulse, such as when someone, suddenly, orders the more expensive wine. It is common to describe this behaviour as spontaneous. But this usage is mundane when it comes to intentional action generally because whether an agent decides quickly or

slowly does not change the character of their action or intention. For instance, two friends passing a bar may 'spontaneously' stop to have a drink together: "Wanna stop for a drink?", "I was just thinking the same thing". The second usage of 'spontaneous' that is inadequate is that which implies a natural tendency, as when a flower spontaneously blooms in spring. It does not make much sense to apply this usage to joint action since it implies a lack of intentionality. Unless there is a serious defect, the tendency will actualise of its own accord. Of course, some natural tendencies can be intentional, but they need not be, and that is why this characterisation is inadequate. This usage is closely related to a third connotation of 'spontaneous', where spontaneity is aligned with thoughtlessness. This is clearly inadequate for joint action since intentional action requires thought. These three connotations of 'spontaneous', impulsive, natural tendency, and thoughtless, are all related in some sense and yet pick out distinct features that make them inappropriate for intentional joint agency.

On the other hand, the sense of 'spontaneity' that accurately characterises some joint actions is that of *immediacy*. When the protesters in Algiers went out to the street together, rather than as disparate individuals, they did so because they were bound together immediately. 'Immediately' does not mean quickly; rather, it has logical register. Agents are bound together immediately so long as they do not depend on a mediating factor to be in fact bound together. For instance, the protesting Algerians did not need to confirm with each other to protest together that morning—they were bound in action immediately. There was no need for the confirmation to mediate between, for instance, their wanting to protest together and their protesting together. They simply went out into the street and protested together spontaneously. This sense of spontaneity gives a positive and negative characterisation of spontaneous joint actions. The positive characterisation is in terms of immediacy, namely, a spontaneous joint action is one in which co-agents are bound together immediately and then act. On the other hand, the negative characterisation may be more

informative, namely, a spontaneous joint action is one in which co-agents are bound together without binding interaction. Binding interaction is a form of interaction, such as promising, agreeing, expressing willingness, that plays the role of providing agents with the appropriate epistemic standing to act together.[57] By engaging in binding interaction, each co-agent knows that the other is acting with her. And part of its role is to warrant co-agents to make further judgements and to perform actions that serve their joint action. Despite this seemingly important role, the reason why binding interaction sullies spontaneity is because it is a mediating factor. If I require, for instance, your agreement to act together with you, then I cannot move from the intention to act together to acting together with you without your evidential mediation, without exploiting you. The move from intention to action is illegitimate. Binding interaction always forces agents to look outward for validation instead of acting immediately with another. Spontaneous joint action cannot depend on binding interaction.

5. Defending the Spontaneity Requirement

Everything is now in place to argue for the spontaneity requirement on theories of joint agency. The argument is simple: because they involve legitimate occupations of the co-agential perspective, spontaneous joint actions are paradigmatic instances of joint agency; therefore, theories of joint agency must explain spontaneous joint actions. For the purposes of establishing the spontaneity requirement, it needs to be assumed that spontaneous joint action is possible, that is, that the togetherness requirement can be met in a satisfactory manner. My aim here is to show that the requisite resources for the satisfaction of the togetherness requirement, the central demand on theories of joint agency, are the same resources needed to satisfy the spontaneity requirement. Before illustrating this connection, it is helpful to remember that these requirements can seemingly

[57] I include as a form of binding interaction what Tuomela calls the Bulletin Board View, that is, a publicly shared plan for joint action that is not communicated directly (Tuomela 2005, 336), because it is an act of informing.

come apart once a theory is provided. For instance, the togetherness requirement may be satisfied on the basis of binding interaction, but this would not satisfy the spontaneity requirement. Even though the basic resources for satisfying both requirements are the same, there appear to be ways of satisfying one without the other. But this apparent possibility does not diminish the centrality of spontaneity to joint action and, hence, why spontaneous joint actions are paradigmatic expressions of joint agency.

Recall how the idea of joint action is distinguished from merely parallel individual actions. It was argued that this is explained by the fact that agents who are acting jointly occupy the co-agential perspective legitimately and, in virtue of that, they view each other as partners in action. When two individuals are merely acting in parallel, they do not view each other as partners. The idea of joint action is isolated by the co-agential perspective that agents take towards each other. As I said before, agents can occupy this perspective legitimately or illegitimately. What is needed for agents to be bound together is for them to occupy the co-agential perspective legitimately and, then, they can act together intentionally. The crucial feature of the co-agential perspective is that co-agents do not view each other as evidence—they exercise rational agency jointly. This non-evidentiary relation between co-agents is what clears the way to spontaneous joint action, which is characterised by its immediate bond or its lack of binding interaction. Now, suppose two agents do not treat each other as evidence: they have not agreed, signalled their willingness, nor planned to act together. Instead, they legitimately occupy the co-agential perspective (assuming that this is possible) without binding interaction. From that perspective, they can have their joint action in mind, think of each as partners, and perform their part of the joint action all without having engaged in binding interaction. For example, the Algerians in one part of the Kasbah did not engage in binding interaction with those in the far end of the Kasbah and, yet, they went out into the street bound together, thinking of each other as partners in protest against French colonial occupation. The two sets of people did not

provide any evidence for each other that would establish the relevant bond and epistemic standing between them. Despite this, some were willing to run into French gunfire and sacrifice themselves for their collective freedom. What allowed them to act together in this spontaneous way was the fact that they could legitimately occupy the co-agential perspective. The connection between joint action and spontaneity is the special agential perspective that forces agents to view each other, not as evidence, but as agents who are their partners in action.[58] The very resource that distinguishes between joint and merely parallel action is the same resource that allows for the possibility of spontaneous joint action. These spontaneous cases are not outliers of joint agency, they occupy a central place in the phenomenon. It is required for a theory to explain them because that which makes joint action possible is what makes spontaneous binding possible. If an appeal is made to the role of viewing each other as partners to explain the phenomenon of joint agency, then spontaneity is an open possibility.

Were the possibility of spontaneous joint action to be denied, a revision to one of the fundamental concepts that I have been discussing would be required. Without putting forward a particular theory for how agents are bound together, spontaneity is a requirement for such a theory so long as it aims to utilise the natural, intuitive notions that I have been describing. Spontaneity is part of the web of concepts that is involved in our concept of joint agency. Without revising our natural notions of togetherness, viewing another as a partner, or rational agency, an explanation of spontaneity is required. Perhaps spontaneity is not as I have described, i.e., as immediacy or as the lack of binding interaction. In that case, it would not hold special interest to action theory and should not be singled out. But assume, for a moment, that spontaneity in some sense is possible *permitting* binding interaction. The problem here is that co-agents would never legitimately occupy

[58] Though by a distinct argument, one philosopher who recognises a connection between spontaneity and joint action is A.J. Julius (2023, chap. 10).

the co-agential perspective because they would exploit each other as evidence in order to bind together. And if they do not occupy that perspective, then the distinction between joint and parallel action cannot be made intelligible. By not being able to explain spontaneity as I have described it, the foundational notions of (joint) togetherness and viewing another as a partner would need to be revised or abandoned. There is a further damaging consequence to this revised conception of spontaneity. If agents do not truly occupy the co-agential perspective because they are bound through binding interaction, then our natural conception of rational agency cannot be applied to the domain of joint agency. The idea that the world does not *determine* an agent's exercise of their rational powers would not be applicable in the joint case because the would-be co-agents' exercise of joint agency, by exploiting each other as evidence, would be, in part, determined by the world. If this is right, then the natural conception of rational agency would not apply to joint agency and, consequently, a distinct conception of rational agency must be adopted in the domain of joint agency. By neglecting spontaneity as I have described it, our conception of rational agency would need to be revised or end up as a pluralist concept. I believe that this line of reasoning amounts to the fact that the natural concepts of togetherness, viewing another as a partner, rational agency, and spontaneity form a web of ideas that must be accepted together as a whole. To alter one of the concepts has a negative effect on the other concepts. After clarifying these basic notions that guide inquiry about joint agency, it is clear that an explanation of spontaneity is lacking. A theory of joint action must explain spontaneity in order to truly make good on these other basic notions. The explanation of spontaneous joint action is indeed a requirement.[59]

[59] There may be a more direct route to the spontaneity requirement. The power of rational agency is a spontaneous power. The ability to set my own standards of action is something that I do immediately; I am not determined by some outside force to do so. There is no mediating factor between my deliberation about reasons and my intending to act. Likewise, the ability for us to set our own standard without exploiting one another as evidence is something that we can do as long as we can legitimately occupy the co-agential perspective together. Thus, insofar as rational agency is exercised in joint action, there is a requirement to explain how spontaneous joint action is possible.

6. The Challenge of Spontaneity

Now that the two requirements on theories of joint agency have been established, I want to state explicitly the challenge that spontaneity poses. In order to satisfy the spontaneity requirement, the most common ways of meeting the togetherness requirement are ruled out because they involve binding interaction. Likewise, the common ways to satisfy the togetherness requirement seem to make spontaneous joint actions impossible. Let's take a few examples to see the conflict. First, Margaret Gilbert argues that we can act together by forming a plural subject.[60] A plural subject is a unified subject that is constituted by the pooled wills of its members and is not reducible to any constituent will/s. To form the plural subject, agents must express to one another their readiness to act together.[61] The natural way to understand this is that each co-agent must express their readiness and understand the other's readiness in order to form a plural subject. The plural subject seems to satisfy the togetherness requirement: each co-agent will legitimately view the other as a partner, once the plural subject is formed. But the necessity of expressing willingness to each other violates the spontaneity requirement: the other must be a source of evidence in order to establish the joint commitment at the heart of the plural subject. Mutually expressed willingness is a form of binding interaction. The trouble is that each co-agent must, at the same time, view each other as both partners and evidence. Gilbert's method of satisfying the togetherness requirement conflicts with the satisfaction of the spontaneity requirement. Second, Michael Bratman proposes a reductive account of shared agency. To act together, co-agents must have individual intentions with appropriate contents that mesh together while viewing each other as intentional co-participants.[62] The best way for co-agents' intentions to mesh together is to plan their action together. If we are

[60] (Gilbert 1989; 1990; 2009a; 2014b). Much more will be said about Gilbert's account in the next chapter.
[61] (Gilbert 2014b, 29). It is these individual expressions, as a starting point, that Baier finds problematic.
[62] (Bratman 1993; 2014). In the next chapter, I will go into more detail with Bratman's account. Bratman may not be a candidate for criticism here because he is not actually concerned with the moment of being bound together. Though, I am confident that his account does not have the resources to adequately explain spontaneity.

painting a house together, we might plan that I go buy the paint and you rent the ladders. Plans such as these mesh together because they are consistent and interdependent courses of action that will bring about our shared end. The same problem rears its head for Bratman though, as it did with Gilbert. By planning our action together, we allegedly meet the togetherness requirement. But if planning together is required for our attitudes to appropriately mesh, then what we each offer to the other by planning is evidence and epistemic assurance. We have engaged in binding interaction. Planning together seems to rule out the possibility of spontaneous joint action. These two examples support my claim that there is a conflict between the togetherness and the spontaneity requirements, so long as the typical means of meeting the togetherness requirement are taken. This purported conflict founds a challenge to the theory of joint agency: the challenge of spontaneous joint action.

We might be left wondering at this stage how exactly do agents occupy the co-agential perspective legitimately without engaging in binding interaction. That is, how can agents be in the relevant epistemic position with respect to their fellow co-agents without having engaged in binding interaction? If co-agents do not have the right epistemic standing, why should they believe that they are acting with another at all? How can they, by their own lights, think that they are acting jointly with others instead of acting merely in parallel with them? Answers to these questions and the challenge of spontaneity depend on a satisfactory account of joint agency that meets both the togetherness and the spontaneity requirements. Here, I will only signal the path that I think is most profitable: to respond to the challenge and provide an adequate account of joint action will require looking *beyond* the attitudes of the co-agents and their interactions alone.

One possible response to the challenge that aligns with my suggestion and that I wish to dismiss at the outset is that of social norms, which legitimate the collective occupation of the co-

agential perspective.[63] For instance, two people together and spontaneously rush over to help an old man to his feet who slipped on ice and cannot get up because they are guided by a social norm to help those who have fallen and cannot get up. The social norm seems to legitimate thinking of the other as a partner in helping up the old man for each co-agent. In response, I grant that social norms can explain some spontaneous joint actions, but why should the range of spontaneous joint actions be limited to social norms? It seems plausible that people can act together spontaneously where there is no social norm or where they act together against a social norm. For example, two hoodlums might spontaneously attack a person together who is walking alone at night. There is no social norm to attack individuals walking alone at night. For a social norm account of spontaneous joint action to work, it must explain why there is this limitation on spontaneous joint actions. It may turn out that no account can be given that adequately satisfies our two requirements without appealing to social norms. But this would be a concession and it should only be taken after a thorough investigation has exhausted all other possibilities. There is no good reason before this investigation has been completed to limit spontaneous joint actions to those governed by social norms. Moreover, it seems that a social norm account of spontaneity ultimately rests on binding interaction. The social norms that we have are the ones that we agree, in some sense, to live by. For instance, there is a social norm that when an election is called, the citizens will vote in it because people have agreed to live in a democratic society. Of course, every individual does not explicitly agree to the norms but, somewhere down the line, the peculiarities of social norms depend on some form of agreement. If this line of argument is right, then the co-agential perspective, that is, not viewing others as a source of evidence, is not fundamental to joint agency. And this, again, I believe would be a major concession. With this account out of the way, I can move forward in Chapter 2

[63] Here, one might think of Baier's Humean remarks about a customary 'we' undergirding all activity (Baier 1997). Social norms could be interpreted as our collective customs or habits.

with surveying the most plausible extant theories' ability to answer the challenge of spontaneity. My answer to the question of the legitimate occupation of the co-agential perspective along with my account of joint agency will be provided in Chapter 3, following the failed result of the survey.

CHAPTER 2

Attempting to Meet the Challenge

The previous chapter imposed a challenge on theories of joint agency: the challenge of spontaneity. The challenge holds that an adequate theory of joint agency must meet two requirements that seem to be in tension. First, the togetherness requirement contends that a theory must explain how agents are joined such that they can act together. Second, the spontaneity requirement contends that a theory must explain how agents can be joined immediately or without binding interaction. The second requirement has not been noticed in the literature. Most theories view spontaneous joint action, if at all, as just another example of joint action. There is nothing special about spontaneity that would make it a requirement. This omission is due, I think, to the fact that spontaneity in this context has not received much thought from philosophers. However, as the previous chapter argued, spontaneity is entailed by the co-agential point of view, which in turn helps carve out the distinct concept of joint togetherness. The conceptual relation between togetherness and spontaneity is precisely why the challenge of spontaneity is a serious matter. If both requirements cannot be met, then the concept of togetherness at issue must be revised. The concept of togetherness, as I presented it, leads to the interesting sense of spontaneity, so the concept of togetherness at issue here cannot be maintained while revising this conception of spontaneity. Moreover, by revising the concept of togetherness, the phenomenon of joint agency seems to vanish since the distinction between joint and parallel action cannot be maintained. Failing to meet the challenge would be a grave result for the cogency of the concept of joint agency.

In this chapter, I will survey the prominent accounts in the literature in order to see whether any can meet the challenge. The chapter is divided into four sections, following an organisational heuristic proposed by Schweikard and Schmid.[64] This heuristic carves up accounts of joint agency

[64] (Schweikard and Schmid 2013).

into three categories: mode, subject, and content. The purpose of these categories, derived from the three categories in the statement "we intend to phi", is to distinguish accounts based on where they locate collectivity.[65] Mode accounts place the collectivity in the attitude, namely, 'we-attitudes', in contrast to singular 'I-attitudes'. The main proponents here are John Searle and Raimo Tuomela. Subject accounts, such as Margaret Gilbert's, rely on the collective, irreducible 'we' rather than the distributive 'we'.[66] And content accounts insist that the action bears the collective burden. Content accounts are offered by the reductionists in the joint agency literature, such as Christopher Kutz and Michael Bratman. In sections 1-3, I will assess these theories. Ultimately, I conclude that these theories cannot meet the challenge of spontaneity. In light of this result, a different way of thinking about joint agency is required, but I save this for the following chapter.

1. Mode Accounts

I begin with mode accounts for two reasons. First, as a matter of historical contingency, they appeared before the others, less with the aim of accounting for joint agency and more to provide a foundation for theorising sociality in general.[67] Second, the reason why mode theories fail is instructive for theorising about joint agency in general. Searle develops his account to avoid two undesirable views, according to him.[68] The first is that collective intentionality can be reduced to the aggregate of individual intentionality.[69] Consider that a fire in the theatre may cause the audience members to rush towards the exits thinking only of their individual safety, but, in another case, they may also head to the exit with a plan to do so in an orderly, safe fashion. Even though the exact

[65] For a more detailed breakdown, see (Schweikard and Schmid 2013; Schmid 2018).
[66] The collective 'we' is not exclusively used by subject accounts, though mode accounts do not make it the focal point.
[67] Searle is rather explicit about this (1995, 26). Also, this is not to say that no one was thinking about sociality and joint action before this. The history of philosophy is populated by many philosophers who confer a foundational role to sociality, e.g., (Aristotle 1941; Hegel 1977; Marx 1994). Indeed, Searle and Tuomela develop Sellars' somewhat inchoate concept of 'we-intentions' (Sellars 1980) and Gilbert explicitly develops her theory based on work by Georg Simmel (Simmel 1971a; 1971b).
[68] (Searle 1990; 1995, 23–26).
[69] I have been using 'joint' for what Searle describes as 'collective', though there may be some differences due to theoretical details.

same bodily movements were performed in each case, there is a difference between the two: the latter involves collective intentionality. So, it cannot be the case that a set of individual actions or intentions equals a collective action or intention.[70] Collective intentionality is irreducible. This seems to imply the second undesirable view, namely that collective intentionality is located in a group agent that exists over and above the individuals performing the behaviour. The idea that aggregated individual intentions cannot explain collective action seems to require that another singular entity, a group agent, performs the action. Searle is convinced that singular *and* collective intentionality can only occur in the minds of individuals in part because intentional causation happens through individuals alone. His solution is to posit a primitive kind of attitude, we-attitudes. When an agent we-intends, she does not merely have an intention but an intention that necessarily already involves other people without considering the content of the intention. This is what makes Searle's account a mode account: the collectivity figures in the attitude and, in particular, an attitude in the collective mode rather than the singular mode. To avoid the consequence of an irreducible group mind, Searle contends that primitive we-attitudes are simply another kind of attitude that individuals possess. We-intentions explain how we act together because two individuals who possess matching we-intentions that refer to each other will act collectively when they act from their we-intentions. The we-intentions can be bolstered by mutual beliefs about the other's attitudes, but this is not necessary for Searle. Indeed, he believes that relying on mutual beliefs or common knowledge to achieve togetherness is a form of reduction to individual attitudes. The appropriately referring we-intentions do the trick and this is what allows his account to possibly explain spontaneous joint action.

Searle actually provides an example of spontaneous joint action:

[70] Searle also points out that the contents of individual intentions and collective intentions will differ. An individual cannot intend to perform a collective action since she cannot do it on her own (Searle 1990, 403).

> I see a man pushing a car in the street in an effort to get it started; and I simply start pushing with him. No words are exchanged and there is no convention according to which I push his car. But it is a case of collective behaviour. In such a case, *I* am pushing only as part of *our* pushing.[71]

Here, each individual has a we-intention that implies reference to the other person. The content of their matching we-intentions is to get the car started by pushing it (together). By instrumental reasoning, each can derive what she is to do individually, e.g., 'I intend to push from the left side'. Searle's mode account explains their collective action because their matching we-intentions which successfully refer to the other agent cause their individual actions that constitute part of their collective action. Togetherness is achieved by the successfully referring we-intentions and spontaneity is achieved because there was no need for binding interaction. The challenge of spontaneity appears to be resolved.

However, Searle's theory has quite a few difficulties.[72] But the question that I think is most relevant to spontaneity is whether togetherness is really achieved by the successful reference of agents' we-attitudes.[73] What makes we-intentions more effective than regular I-intentions that accommodate other's intentions? Why is this any different than parallel, strategic action? I do wonder why the initial pusher has a we-attitude with accurate reference without engaging with the other pusher; after all, the second arrives subsequent to his initial pushing and, so, his intention in

---

[71] (Searle 1990, 402). I think it is plausible to push back here on Searle's non-conventional claim. Though there is no convention to help *him*, there may very well be a convention to help those whose cars are stalled.

[72] First, Bratman complains that it violates parsimony because there are ways of explaining joint action without inflating our ontology by including we-intentions. Of course, Searle believes that we-intentions are the necessary result of avoiding the two undesirable views. But there is no independent argument for the existence of we-intentions aside from this avoidance. Second, the ability of we-intentions to avoid the undesirable views is questionable. Schmid has argued that mode accounts ultimately are versions of subject accounts (Schmid 2017). Third, Searle's account violates the own action condition on intention which says that an agent can only intend her own action and not (also) those of others (Baier 1970), though this objection has plausible responses (Bratman 1999b; Ferrero 2013). Fourth, his account is charged with being too subjectivist because the collectivity is all in the heads of individuals on their own—in a sense, we-attitudes stipulate collectivity instead of joining individuals together (Mathiesen 2002; Meijers 2003; Schmid 2003; Zaibert 2003). Finally, we-intentions are not primitive such that singular attitudes must be involved in them (Salice 2015)

[73] Note that, for Searle, reference to others occurs outside of the scope of the content of we-attitudes. The reference is built into the attitudes (Searle 1990, 408).

action. But let's grant the relevant we-attitude can be possessed spontaneously. I think this example only works, if it does, because the agents are in close contact and can see the other's activity. Without this, the intuitiveness of the explanation is lacking. Consider a case where we have mutually referring we-intentions to go to the museum. As a spontaneous case, we do not agree to do so. Instead, I covertly overhear your plan to go to the museum in the afternoon and you then spy my plan to go there from my diary. Neither of us are aware that each has the relevant we-intention. Afternoon arrives, and we both head to the museum separately. Have we gone to the museum *together* just because we each have mutually referring we-intentions and have gone to the museum? We may not even be in the same room or know where the other is, and yet our we-intentions are satisfied. I think that something more is required than the successful reference and causation of we-intentions. Spontaneity puts this problem in full view: co-agents may have interlocking attitudes, but if they have no reason to think that the other is acting with them, why credit them with acting together instead of in parallel? It seems clear that to act together each agent needs to be in an appropriate epistemic position with respect to the other. Many have found it necessary for a theory of joint agency to account for this epistemic constraint, and I follow their lead.[74] In spontaneous cases, it is very difficult to maintain that two people are acting together without each having the appropriate beliefs about the other. It seems necessary that the joint sense of togetherness involves having, in some sense, epistemic standing towards one's co-agents.[75]

Moreover, Searle indicates a further problem with we-intentions.[76] When I have a we-intention, I believe that you also have a we-intention because I have a we-intention for *our* action. And I might be wrong about your state of mind. This is not in itself problematic but, because my

[74] E.g., (Chant and Ernst 2008; Gilbert 1990; Bratman 1993; Tuomela and Miller 1988; Satne 2020; Schmid 2018). For a dissenting view, see (Blomberg 2016b).
[75] In Chapter 3, my account will need to explain how it meets this epistemic constraint.
[76] (Searle 1990, 408).

we-intention makes reference to you *irrespective* of the content of that attitude, I can be in error about my own state of mind and, thus, what I am doing.[77] This is a serious problem, and Searle seems to think that it is resolved by abandoning the Cartesian assumption that we can only be wrong with respect to the world. The problem however is that this attitudinal form of error conflicts with rational agency. I am the one who sets the standard for my intentional agency: I determine what I am doing. I determine whether my action is intentionally part of our action, which never comes to fruition, or whether I am acting independently. What I do intentionally depends on the attitude I set as my standard of action. If I can be in error about my own state of mind and what I am doing intentionally, I do not seem to have a grip on what I do intentionally. So, *I* cannot be credited with determining what I do intentionally. And this, at a minimum, makes it very difficult to be in a position to know if I am intentionally doing what I intend to do. Spontaneity exacerbates this problem because it rules out any form of engagement that would allow me to be certain that another person has the relevant we-attitude. In the end, these two problems with reference in relation to spontaneous joint action render Searle's mode account deficient.

But perhaps a mode account that does not appeal to primitive we-intentions or ignore mutual beliefs can succeed. Raimo Tuomela has developed such a view. Tuomela still holds on to we-intentions, or we-attitudes more broadly, though they are not primitive in Searle's sense.[78] We-intentions are one part of what he calls the We-mode, in opposition to the I-mode, and what characterises it is thinking or acting "as a group member".[79] So, when I have a we-intention, I believe that I am a member of a group with others who are also aiming to satisfy our intended aim.

---

[77] I am not exactly sure what it means for an attitude to be in error when the error is not a matter of content. On a metaphysical reading, the false reference might render the we-attitude non-existent—I just don't have the we-attitude when it fails to refer and that's why I do not know what I am doing. On a weaker epistemic reading, the error means that I simply do not know what I am doing but I am still in that state of mind.

[78] (Tuomela and Miller 1988; Tuomela 1991; 2002; 2005; 2007). Tuomela believes that the we-mode is irreducible to the I-mode, but the possession of we-intentions requires that an agent have beliefs about her partner's state of mind.

[79] (Tuomela 2003, 93).

In order to think of myself as part of a group, I must believe that there is a group of like-minded people. Here is Tuomela's (expanded) account:

> You and I *(jointly) intentionally performed X jointly* if and only if
> A. X is a collective action type, that is, an 'achievement-whole' divisible—either *ex ante actu* or *ex post actu*—into your and my parts;
> B. we jointly intended to perform X jointly;
> C. we performed X jointly in accordance with and partly because of our joint intention to perform X (or some 'closely related' action) jointly;
> D. you and I mutually believed—or at least shared the belief—that A, B, and C;
> E. B in part because of D.[80]

As you can see, there are many minute qualifications in this account, and I will for the most part pass over them since they are unimportant for my purposes. Let's see why this account does not suffer from the two problems associated with Searle's. First, condition D introduces a stronger bond between the co-agents than mere reference. You and I have beliefs about each other's beliefs about our joint action. Second, I cannot, under typical circumstances, be mistaken about my state of mind and, so, be confused about what I am doing intentionally. The reason for this is that the possession of a we-attitude on Tuomela's account does not depend on your partner having the same we-attitude. Instead, *believing* that you have the relevant attitude is enough to warrant my formation of a we-attitude. Tuomela's account seems to resolve the problems in Searle's account.

Now, one could ask whether condition D flouts the spontaneity requirement since our mutual possession of the relevant beliefs seems to depend on binding interaction. But the weaker qualification of merely sharing beliefs seems to avoid that problem. The shift to the weaker epistemic condition might make one wonder, as in the museum case above, whether this account meets the togetherness requirement. But let's grant that Tuomela's account meets the two requirements. It still needs to be an adequate account in its own right. The question that should immediately arise is whether this account is viciously circular. After all, the analysans makes use of

[80] It is expanded from the basic account in order make joint action rational (Tuomela 2007, 112–14).

concepts that are supposed to be explained or analysed: *joint intention* and *performing X jointly*. The blatant circularity of the account does not escape Tuomela's notice, and he provides three reasons that count against its perniciousness. I will address each in order.

First, there are evolutionary considerations that provide joint action an "objective basis".[81] Without getting into details, joint action appears to be an adapted disposition (preconceptual, in some higher species). I find it hard to see how this reason should make the account's circularity more palatable. The fact that a phenomenon exists does not mean that we can dispense with providing an explanation of the phenomenon, even if it depends on evolutionary preconditions. For instance, that we have (adapted) dispositions to perform arithmetical functions does not excuse the need for an explanation of how we in fact perform arithmetic. Moreover, even if some circularity is permissible, more argument is needed than pointing to evolution. Second, early-childhood psychology provides evidence that joint activity begins early in the life of children.[82] Once again, this empirical point does not satisfy the need for a non-viciously circular account of joint action. Third, "joint action is deeply ingrained in common sense".[83] Again, a phenomenon being common sense does not warrant a viciously circular 'explanation'. The fact that something is common sense often compels deeper philosophical refection because it is so ubiquitous and yet its explanation eludes us. With these three considerations, Tuomela does not so much make the circularity of the account more palatable but instead shows the difficulty of providing a satisfying explanation for a quotidian phenomenon.

The vicious circularity of Tuomela's we-mode account of joint action shows us that more than an elucidation or a rich description of the phenomenon is needed to understand it.[84] The fact

[81] (Tuomela 2007, 110). I assume that, by objective basis, Tuomela means that joint action is a real phenomenon.
[82] (Tuomela 2007, 110).
[83] (Tuomela 2007, 110).
[84] The idea that Tuomela is offering an elucidation rather than an explanation is discussed by (Schweikard and Schmid 2013).

that joint action is common sense does nothing to appease the sceptic who would deny joint phenomena. A viciously circular account would play right into her hands. To provide an *explanation* of joint action requires that we avoid vicious circularity.[85] This would, at least, give the sceptic something to think about. It has been helpful to survey mode accounts because they reveal two constraints on a theory of joint action: that there needs to be some sort of epistemic relation between co-agents in order for them to count them as acting together and that it cannot be viciously circular. The failure of mode accounts to offer a convincing explanation of joint action is not a terrible result because it has been suggested that mode accounts ultimately boil down to subject accounts.[86] Let's consider those accounts now.

2. Subject Accounts

Subject accounts differ from mode accounts in that the collectivity is found in the subject rather than the attitude. Subject accounts often appeal to a 'plural subject' or agent 'who acts as a body'.[87] This characterisation seems to raise questions about super-agents about which Searle was worried. However, there is no need to assume, some have argued, that a plural subject is mysterious or even implies an agent over and above the co-agents. For instance, David Velleman argues that agents can *literally* 'pool their wills' by sincerely uttering the appropriate statements that settle a practical matter and thus commit the agents to perform the intended action.[88] Meaningful utterances certainly are not mysterious occurrences. And Hans Bernhard Schmid pushes back on the implication of super-agents: "If you and I go for walk together, it's just the two of us out there, not three: it's not that there's a third subject ('the we') somehow coming along with us—in the exact same way, if I go for a

[85] For more on circularity in joint action, see (Petersson 2007). In Chapter 3, I will address the problem of circularity when my own account is presented.
[86] (Schmid 2017).
[87] (Gilbert 1989; Schmid 2009; 2014; 2016; 2018; Rödl 2014; Roelofs 2017).
[88] (Velleman 1997). Velleman borrows this phrase from Gilbert. I will not present more of Velleman's positive view since it clearly contravenes the spontaneity requirement: his account requires agents to exchange statements of their intention to act together.

walk alone, it's just me out there, all by myself, there's no such thing as 'the I' coming along with me".[89] These considerations should, at least, diminish the fear of super-agents in subject accounts.[90] In this section, I consider Margaret Gilbert's plural subject account of joint action to determine whether it can respond to the challenge of spontaneity. I consider only her view because it is the most influential and developed.[91]

The crux of Gilbert's account is her notion of a joint commitment. When we act together intentionally, we act on the basis of being jointly committed to act as a body.[92] We no longer act as individuals who have unilateral control over our actions—we are answerable to each other in virtue of being joined. As a result, we have obligations and entitlements with respect to each other. I can expect that you will perform your part in our action because we are jointly committed to it, and I am obligated to you to perform my part. Naturally, these are not moral obligations, but they are obligations insofar as commitment obligates.[93] Being joined together permits us to use the collective 'we' to refer to ourselves as a body, or a group. Importantly, the creation of a joint commitment provides us the epistemic warrant to believe that we are acting with another and that we are obligated and entitled in the appropriate ways. When we share a joint commitment, we are appropriately bound together. In order to judge whether Gilbert's account can meet the spontaneity requirement while meeting the togetherness requirement, more needs to be said about joint commitment and, in particular, how it is created.

---

[89] (Schmid 2018, 241).
[90] Some other objections to subject accounts are raised by, to isolate a few, (Bratman 1999c; Miller 2001; Baltzer 2002; Bittner 2002; Robbins 2002; Zahavi 2008; Kopec and Miller 2018).
[91] Some views that may fit best into the subject account category are Anscombian views (Seemann 2009; Laurence 2011; Satne 2020; Roessler 2020) and team reasoning views (Sugden 1993; Rovane 1998; Bacharach 1999; Sugden 2003; Gold and Sugden 2007; Pacherie 2011; 2013; Gold 2018). The main problem with Anscombian accounts is that there must be some account of plural practical knowledge. The closest to a successful account, in my opinion, is Glenda Satne, but her account relies on epistemological externalism to overcome the problem. Her account stands or falls with the plausibility of epistemological externalism. The main problem with team reasoning views is that they are circular: in order to act together, we must already be doing something together. Natalie Gold suggests that the reasoning is sub-personal (this theory has its roots in economics), but then we should wonder why it helps to explain intentional action.
[92] (Gilbert 1989, chap. 4; 1990; 2008; 2014b)
[93] (Gilbert 2008, 495–500).

To properly understand joint commitment, a distinction needs to be drawn between what Gilbert calls 'personal commitments' and 'individual commitments'.[94] A personal commitment is a commitment of a person that she can create and rescind unilaterally, and she is answerable to herself alone with respect to it. For instance, my commitment to submit my paper to journal *X* by the weekend is a goal that I set for myself and I am answerable to myself alone with respect to my pursuit of it. Perhaps at some point during the week, though, my basement floods and I rescind my commitment to submit my paper by the weekend. This recission has agential bearing only on myself because I have unilateral control over my personal commitments. On the other hand, individual commitments are the commitments of a person over which she does not have unilateral control or answerability because they are dependent on joint commitments. The content of individual commitments refers to the actions of the individual as opposed to the plural subject. For example, when we paint the house together, I can have the individual commitment to collect the paint from the garage while you can have a distinct individual commitment to set up the ladder, and both of these depend on our joint commitment to paint the house together. With this distinction in hand, joint commitment can be isolated because personal commitments are not involved in joint commitment and joint action at all.[95]

Joint commitments have six main features.[96] The first is 'holism', which says that "joint commitment is the commitment of two or more people".[97] On my own, I cannot have a joint commitment because, being isolated, I cannot be joined to another. This, for Gilbert, is the core feature from which the others follow. The second is 'answerability'.[98] Due to being jointly

---

[94] (Gilbert 2009a; 2014c).
[95] Gilbert has characterised this distinction elsewhere as the "disjunction criterion" (Gilbert 2008, 491–93). Of course, I may at the same time have personal commitments, but they are logically distinct from joint commitment.
[96] (Gilbert 2014c, 40–41).
[97] (Gilbert 2014c, 40).
[98] This feature is directly related to what Gilbert has elsewhere called the 'obligation criterion' (Gilbert [1997] 2000; 2008; 2009a). This normative element is criticised by (Bratman 1999c; Miller 2001, 88–89).

committed with others, each partner is answerable to the others for any violation of the commitment. Answerability is a feature of commitment in general and, since the commitment is not mine alone, I am answerable to my partners who are committed with me. Were I to fail to do my part, I would need to answer to my partners for my failure. The third feature is 'creation', such that each resulting partner must participate in the creation of the joint intention. I must take part in settling, to use Velleman's language, the joint commitment in order for me to be answerable to it. Gilbert notes that equal participation may not always occur given that there are situations where individuals have been authorised to make decisions for others, such as representative institutions. In such cases, though, the joint commitments are accepted as a result of granted authorisation. The fourth feature is 'recission', which states that joint commitments must be rescinded by all partners and cannot be rescinded by one or some partners unilaterally.[99] To rescind unilaterally would require that one partner has authority over the others. But no partner has authoritative standing over her partners by joint commitment. Notice that features two through four are the same features that personal commitments have but they differ because 'holism' strips away any unilateral agential control over joint commitment.

Feature five concerns the partners' subsequent commitments once a joint commitment is created. Each partner who is subject to a joint commitment has 'dependent individual commitments'. As discussed above, each partner is responsible for some part-action which will partly count towards satisfying the joint commitment. There must be a commissive tie between acting as a body and a partner's individual actions. This tie would be severed without dependent individual commitments since a partner could be jointly committed without being committed to do anything individually. Joint commitments would not be satisfied if they did not produce dependent

[99] This feature meets what Gilbert has called the "concurrence criterion" (Gilbert 2008, 494–95). Gilbert notes in two person cases that a violation of the joint commitment may serve as the basis for a seemingly unilateral recission. However, the violation can serve as an agent's conditional recission of the joint commitment.

individual commitments that move the members of the plural subject. Naturally, these dependent individual commitments are not personal commitments since each partner does not have unilateral control over them. Just because individual commitments are *for* partners individually does not make it the case that they have unilateral control over their individual commitments. The personal/individual commitment distinction is helpful for Gilbert because it allows her to avoid worries concerning conditional intentions.[100] In particular, if we are conditionally committed to act together such that our intentions display part of an exchange, e.g., 'I'll act with you if you act with me', they will never add up to a joint commitment.[101] The two intentions can only aggregate and, crucially, are still under the full control of each of us alone. In such cases, we only share a certain form of interlocking *personal* commitment which cannot constitute a plural subject. Moreover, dependent individual commitments are interdependent since they are produced by the same joint commitment—they "must arise and fall together".[102] They are simultaneously created with the joint commitment because they are entailed by it. Dependent individual commitments are ultimately the commitments that partners in particular take on when they create and are thus subject to a joint commitment. The final feature of joint commitment is its content, which is an action, broadly construed, performed as a body.

These logical features of joint commitment do not yet reveal when agents are in fact jointly committed and, so, can act together. For that, we need to know the conditions under which joint commitment is created. But I want to pause and focus on the dependent individual commitments and, in particular, their simultaneous creation. This feature of joint commitment seems to be a problem for spontaneous joint action. Consider Searle's case where I see you pushing your car and I join to help you and we spontaneously push your car together. As described, there has been no

[100] (Gilbert 2014c, 42–45).
[101] See (Robbins 2002). A similar argument, though not directed at Gilbert, is defended in (Julius 2013).
[102] (Gilbert 2014c, 41).

interaction between us at all, so it seems unlikely that *we* are jointly committed to pushing your car. Now, when I arrive and start pushing, you might exclaim, "Finally!" or "I've been waiting for someone to push with me". It seems, by your exclamation, that you are *already* committed to a joint action and, possibly, you were already committed before I even saw you pushing. And, in that case, I could not have had a dependent individual commitment to push the left side of the car, for instance. Where did the joint commitment come from, with its correlative dependent individual commitments, if we do not even know of the other's existence? It seems that Gilbert's account of joint commitment rules out this spontaneous joint action since it requires the dependent individual commitments to be possessed before a joint commitment is established. Perhaps, though, the joint commitment is created when you see me and exclaim, and I push and acknowledge your effort. Unfortunately, this is a form of binding interaction that rules out the relevant form of spontaneity. Spontaneously pushing the car together appears to be impossible from the perspective of joint commitments. Without any specifics with respect to creation, the logical features of joint commitment already seem to, at least, reduce the scope of legitimate spontaneous joint actions.

Now, let's see how a joint commitment is created, according to Gilbert, because perhaps there are some resources to help in this case. Gilbert is aware that explicit agreements already seem to involve the jointness that is to be explained: agreements appear to be joint decisions. For this reason, agents must express readiness to be jointly committed and act as a body.[103] When I ask you to dance, I express my readiness to dance together. And when you say 'Yes", you express your correlative readiness. The same is true if we exchange notes or speak through delegates. Our expression, when it comes to joint agency, is constrained by what we *intend* to express. Of course,

[103] (Gilbert 2014b, 29). Originally, Gilbert used 'willingness' interchangeably with 'readiness' (Gilbert 1989, 179; 1990, 7). It's worth noting that this creation condition already goes beyond merely correct reference, a problem with Searle's account. Expressions of readiness potentially exhibit something between us, or as Gilbert credits Charles Taylor, "entre nous" (Gilbert 2014d, 94; Taylor [1980] 1985, 263–69).

we can unintentionally express things, but unintentional expression will not convey readiness since the agent may not even be capable of doing what was unintentionally expressed. Expressions of readiness convey a state of the will, being ready to be jointly committed.[104] However, expressions of readiness are not sufficient to ground a joint commitment.[105] Imagine two individuals in separate households expressing in their mirrors that they are ready act with the other. Both have expressed readiness, but neither is aware of that fact, so it is unreasonable to think that they are now subject to a joint commitment and are thus answerable to each other. The one would certainly be surprised if the other starting making demands on her. What is needed beyond the mutual expression of readiness is common knowledge of each other's readiness and that they both know that the other knows that they are both ready. Common knowledge will render them jointly ready to act as a body. At this point, they will be jointly committed and can act together intentionally.

Gilbert has a distinct and technical account of common knowledge.[106] In simpler terms, Gilbert characterises common knowledge with respect to joint commitment as expressions of readiness being "out in the open" for the relevant individuals.[107] The result of common knowledge is that each partner knows that everyone has expressed readiness and each partner knows that each knows that everyone has expressed readiness and all of this knowledge is, in a sense, out in the open. In the paradigmatic case, partners are in a "special perceptual relation to each other and to a fact such as the fact that each has expressed quasi-readiness to share in some action".[108] For example, when I ask you to dance, you see and hear me and I see and hear your response; we have common knowledge insofar as our expressions are out in the open for each of us. One result of common

---

[104] (Gilbert 1989, 186).
[105] Gilbert sometimes refers to expressions of readiness as representing an agent's "quasi-readiness" (Gilbert 1989, 185–86). She occasionally uses "conditional commitment" to describe this state (Gilbert 1989, 198; 1990, 7).
[106] The account appeals to set theory in order to resolve the regress of *n*-order beliefs discussed by Lewis (Gilbert 1989, 186–97; Lewis 1969, 52–60).
[107] (Gilbert 2014b, 29). Openness, or 'openness*' is a technical term for Gilbert involving ideal agents who can know the relevant facts.
[108] (Gilbert 1989, 188).

knowledge in paradigmatic cases is that agents are not surprised when one of them acts or infers on the basis of that knowledge.[109] I am not surprised when you grab my hand and head towards the dance floor, I expect it. Likewise, after asking you to dance, you would be surprised if I did not go with you to dance—I have violated the joint commitment that we created. Of course, there will be non-paradigmatic cases where this special perceptual relation does not hold. Common knowledge may be achieved by other means, for example, a news cycle constantly discussing an upcoming election. To create a joint commitment, then, requires that agents express their readiness to act together and everyone involved has common knowledge that everyone is ready to act together.

Do these creation conditions help explain the spontaneous car pushing? Let's start with the common knowledge condition, since common knowledge of our expressions of readiness will warrant spontaneous inferences and actions. Suppose that you, the car pusher, notice my expression of readiness when I begin pushing the car, and now we are joined and have common knowledge of our readiness. This implies that the second before you noticed my intentions that I was not acting on our joint commitment because our readiness was not common knowledge. At the time of instituting common knowledge, my commitments changed. But it seems obvious that before and after your noticing, my intentions, and therefore, my commitments did not change: I was simply intending to push the car with you. Gilbert might say that I started with a personal commitment or maybe a conditional commitment that switched to a dependent individual commitment once common knowledge of our readiness is instituted and we only began acting jointly at that time. I accept that this is a possible response, but it does not seem to match the phenomenology of a very simple case. Do my commitments really change because I have been noticed? Am I really intentionally doing something different as a result of you noticing me? There is no question that we

[109] (Gilbert 1989, 194). Gilbert's example is that of a person removing a whistling kettle from the stove and nobody being surprised that she did so because it is common knowledge that the water is boiling.

are not performing an action together until we are both pushing the car, but being subject to a joint commitment with others does not require performing the action. The most intuitive reading of the car pushing case is that, before you are aware of me, I am *already* committed to acting with you. I would need to hear an excusing reason to think that I am not already joined to you, in some way.[110]

Supposing that this does resolve the simultaneity problem of the car pushing case, the explanation appears to vitiate the *spontaneity* of the joint action. The need for common knowledge, that our readiness is out in the open, seems to fail the spontaneity requirement. Even though the car pusher and I did not exchange words or explicitly agree, we still had to play an evidential role for the other to establish the common knowledge and, by extension, joint commitment. Your noticing my coming to push and my acknowledgment of your readiness to push together serve as the kind of evidence on which spontaneity cannot rely. Our common knowledge depended on this exchange of 'intentional' evidence, even though we did not explicitly agree to act together. Because of this, I cannot *immediately* act with you. I must glean evidence from you and you from me before we can legitimately think of each other as partners in action. In these kinds of close-quarter cases, common knowledge appears to be achieved by the non-spontaneous evidential exchange of expressions of readiness. The real problem for Gilbert's account appears to be that the mutual expression of readiness and its uptake is a form of binding interaction. I have no reason to think that we will act together unless I have evidence that you are ready to do so. Though perhaps, extentionally, these close-quarter cases don't count as spontaneous on Gilbert's account. Perhaps common knowledge can be achieved by other means and spontaneity preserved for some cases. However, in the paradigmatic cases of common knowledge, agents must perform the evidential role that is anathema to spontaneous joint agency.

[110] I think an interesting point is that the spontaneity at issue in joint action is not temporally construed, and Gilbert's account builds temporality into joint commitment as a matter of its logic.

Gilbert claims that her account can accommodate the creation of joint commitments by previous explicit agreement, prior implicit agreement, and those where neither of these are the case.[111] Previous explicit agreement is ruled out by spontaneity. Prior implicit agreement would be ruled out by spontaneity as well because there would still be some previously provided intentional evidence for the agreement. However, the cases where neither of these things occur seem like appropriate candidates for spontaneous joint actions and, indeed, Gilbert suggests that an 'agreement' is not necessary. How does common knowledge in such cases come about? Gilbert draws a distinction between two types of common knowledge. The first kind, which we just saw, is common knowledge amongst agents *X*…*Z*. She calls this "individual common knowledge" because it makes direct reference to particular individuals.[112] When we are jointly committed to dance together, the common knowledge we have directly refers to you and me. However, to accommodate large groups, like nations, Gilbert introduces 'population common knowledge': "someone may notice that some fact is open* *among the Xs*, where one is an X if one possesses some general feature such as the feature of *living on this island*".[113] Population common knowledge removes the requirement that individuals must know of each other in particular. For example, not all Ontarians know each other in particular but they can have (population) common knowledge that Ontarians will elect the next premier together. A joint action based in population common knowledge falls into the third category of joint commitment creation because members of a population do not explicitly or implicitly agree with each other. Indeed, many of them will be entirely unknown to each other. Turning back towards spontaneity, there does not appear to be binding interaction between agents because there is no explicit or implicit agreement between them. Ontarians will act spontaneously when they elect a new premier.

---

[111] (Gilbert 2014b, 33).
[112] (Gilbert 2014c, 51).
[113] (Gilbert 2014c, 51)

Population common knowledge does expand the size of a group who can act together jointly.[114] It is likely impossible to think of 15 million people in particular. And it seems that members of these populations do not engage in binding interaction, if they have never even met. Nevertheless, the object of common knowledge here remains the same: intentionally expressed readiness (of members of a population). Gilbert suggests that it is not necessary to meet someone or be in their vicinity to know that she has expressed readiness.[115] She provides an example of testimony to demonstrate. I can know the intentions of a person that I have never met as long as a different person tells me her true intentions. True testimony is good grounds for knowing the intentions of another. So long as there is some testimonial system in place, members of a population who have never agreed, neither explicitly nor implicitly, to act together can have common knowledge of their readiness.[116] On this basis, individuals seem to be able to act together spontaneously. Admittedly, the extension of this account will be rather small and leave out many paradigmatic spontaneous joint actions. But it seems to be able to explain the togetherness and spontaneity of some joint actions.

Confidence in this explanation should be rather weak, and not only because of its gerrymandered extension. The question is whether relayed expressions of readiness count as binding interaction. Simply because my intentions are expressed via a conduit, a friend's testimony, a text message, etc., and not face to face does not resolve the problem of spontaneity. The fact that there are layers of mediation between individuals who form a joint commitment does not change the situation. What was problematic about binding interaction was that agency is treated primarily as evidence. Whether the interaction is face to face, through governmental delegates, or playing Bach

[114] The number of members is not crucial to this form of common knowledge. A population could have only two members. However, it is useful to Gilbert since she applies her theory of joint commitment to large scale groups (Gilbert 2006; 2009b; 2014a).
[115] (Gilbert 2014c, 52).
[116] In addition to testimony, I think that artefacts can be the basis of common knowledge for Gilbert. A letter, a recording, a constitution all can express readiness to act together.

to the far reaches of the universe and waiting for a response, binding interaction still occurs. Ultimately, the reliance of Gilbert's account on expressions of readiness to ground joint commitment is what precludes its ability to meet the spontaneity requirement.

Having taken Gilbert's plural subject account to represent subject accounts of joint action, its failure to resolve the challenge of spontaneity speaks against the ability of subject accounts in general to meet the challenge. Now, I will move on to content accounts to evaluate their ability to meet the challenge. These accounts are already a step behind because they tend to rely on the distributive 'we' and, hence, explaining the togetherness requirement is a more difficult matter.

3. Content Accounts

The goal of content accounts is to show that the action is where collectivity resides and what is needed is to show how the collective action is caused in the right way without assuming other collective notions. Content accounts are reductive because the collectivity only appears in what needs to be explained, the collective action. Minimally described, a collective action is caused by the aggregated agency of individuals. Somehow, content accounts need to pull togetherness out from the individualist hat. Some questions for content accounts are: why is the collective action not merely a set of individual actions? Why consider it a unit that a group of individuals have produced jointly or at least on the basis of a shared intention? In this section, I focus on two accounts, namely, Christopher Kutz' minimalist account and Michael Bratman's robust yet still reductive theory of shared agency.

Kutz' aim is to provide an account of joint agency that can explain the most typical two person examples, like dancing the tango, all the way up to large scale joint actions in hierarchical institutions, such as passing a bill. Moreover, he explicitly notes that spontaneous joint actions should be explained and joint agency generally involves seeing "the other as an intentional

participant in a collective action".[117] In order to meet this extensional demand, Kutz develops a minimal theory of joint agency that has two basic conditions. At first glance, Kutz' account of joint action seems to address the concerns raised in Chapter 1, in particular, emphasising the role of viewing your partners in action in a particular way. However, this other-regarding feature is quickly eschewed. As I will argue, losing sight of how we see our partners in joint action leads Kutz to fail to meet the togetherness requirement.

The first condition of Kutz' minimal account is that agents who act jointly act from 'participatory intentions'.[118] These intentions consist of two elements: individual role and collective end. The individual role accounts for the individual's action that is *part of* or *fosters* the collective end. For instance, my role in our car pushing is to push this side of the car. There must be some action that an individual performs if she is to count as acting jointly with others, that is, to participate in bringing about a collective end. We can think of the individual role as similar to the means in means-end relations, the difference being that the individual action is not sufficient to produce the collective end. On the other hand, the collective end is a state of affairs, satisfying an action description, constituted or caused by distinct agents' actions.[119] For instance, parents who jointly give their child a birthday gift will have produced the collective end of the child receiving the gift. Depending on the type of action, each agent's contribution to the collective end will either constitute (e.g., tango dancing) or cause (e.g., elect an official) the collective end. Moreover, collective ends do not need to be intentionally produced—one agent may aim at a collective end and another contributes to it unawares.[120] Indeed, no one may aim at a particular collective end and two or more people may produce it. What is important is that at least two agents' behaviour produces

[117] (Kutz 2000b, 4, 9). Kutz discusses his account and its relation to accountability in large scale collective actions in (Kutz 2000a).
[118] (Kutz 2000b, 10–13)
[119] This is a very thin conception of collectivity and it allows Kutz to avoid problems of circularity.
[120] (Blomberg 2016a, 357) emphasises this point.

the state of affairs and there is an applicable action description of it. Unlike individual roles, which depend on being a part of participatory intentions, collective ends can exist distinct from participatory intentions because they can be the causal result of two or more individuals' behaviour. The collective end in a participatory intention has two important functions. First, it rationalises the action that fills the individual role. It makes sense to describe my action as 'doing my part' only insofar as it is intended to contribute to the collective end. Second, the idea of the collective end includes other people. My action alone cannot produce the collective end because a single person cannot produce it, others must contribute. Because of the collective content of a participatory intention, agents' individual actions can be joined together and, so, they can act jointly intentionally.

The second condition in this minimalist account is that agents' participatory intentions must overlap to count as intentionally acting together.[121] The most obvious way for participatory intentions to overlap is to have the same collective end, that is, we both aim to contribute to the same state of affairs conceived in the same way. For instance, I intend and you intend to push the car. The fact that the agents have the same conception of their matching collective end is important because there is an ambiguity in action descriptions that matters when it comes to joint action. Two people may employ the same concepts in their intentions, that is, describe their actions in the same way, *pushing the car*. But when it comes to intentional action, an agent also has a conception of the action, e.g., as getting the car started or as proving strength. The conception of the action matters because it figures into the satisfaction conditions of the intention. While it may be true that we pushed the car and thereby proved our strength, we still did not get the car started. Recall Davidson's discussion of getting on a plane to London.[122] Even though I intentionally got on a plane to London, this plane is going to London, England rather than London, Ontario. My

[121] (Kutz 2000b, 17).
[122] (Davidson 2001c).

intention has not been satisfied since my action, according to my conception, was never performed. This is an important point because agents may have participatory intentions that employ the same concepts and yet we would not count them as acting jointly intentionally unless their conceptions of their collective end, at least, overlap. However, for intentions to overlap does not mean that agents acting together must share the very same conception of their collective end. That is, they need not aim at the same end, but their ends must at *some* level of description match. Kutz provides the helpful example that we may drive to a friend's house for dinner together (and do that jointly), but your end is to have dinner there and my end is to go to your surprise party.[123] Both of our ends are collective ends that refer to the same state of affairs and, at some level of description, our ends intersect. It is true that we are driving together intentionally because our individual actions constitute a part of driving together, but we are not both intentionally going to your surprise party together. Nevertheless, it is still true that we jointly intentionally went to a friend's house for dinner.[124] For participatory intentions to overlap, it only requires that, at some level of description, our collective ends are shared. When we act from the appropriate intentions, we perform the action jointly intentionally insofar as we share collective ends. Joint action, on this minimal proposal, has no need of common knowledge amongst co-agents—each must only intentionally participate in the production of a collective end, described at some level in the same way.

Kutz' account seems capable of explaining spontaneous joint action. The fact that we only need to be in the appropriate states of mind and to be involved in the collective production of the same state of affairs, conceived at some level in the same way, does not involve binding interaction. I can head towards the car with the participatory intention to push it and the other guy can already be doing his part of the same collective end. What is important is that our actions are rationalised

[123] (Kutz 2000b, 21).
[124] What I do intentionally has a broader scope than what I intend to do.

by our collective end that, at some level of description, intersect. The question is whether the account meets the togetherness requirement. Recall that Searle's account couldn't meet the requirement because successful reference alone does not make the action jointly performed, so it is worth noting why Kutz' account does not fail in the same way. In Kutz' account, it is not necessary that participatory intentions refer to people who will perform the action with you. We do not need to share attitudes that mutually refer. Rather, it only requires that we each intentionally participate in an event brought about by us and we conceptualise the event as, at some level description, the same collective end. The collectivity is in the state of affairs produced in the right way. But this is where things start to go wrong, I believe. Recall that Kutz draws our attention to the importance of viewing the agents with whom you act jointly as intentional participants in the action. Where does this fit in the minimal account? It appears to be an unnecessary component of joint action, if the minimal account is right.

One consequence of not sharing the same conception of the collective action is that we individually do not need to care about the satisfaction of others' collective ends. After all, I do not necessarily aim at your end. It must only be the case that at some level of description we similarly conceive of the state of affairs that satisfies our collective ends. Consider the example from Chapter 1 where the delivery person dropped off a package that two parents jointly gave to their child. Now, assume that the delivery person, because the package is wrapped, has a participatory intention to give a gift. She is doing her part in the (collective) gift giving by dropping off the package and the parents will complete the act by giving it to their child. Her individual role is causally related to the state of affairs that satisfies her collective end—the gift giving. The parents have a different collective end—to give *their child* a gift. Both collective ends are realised by the same state of affairs that was collectively brought about. All three people had participatory intentions, each performing an action that led to their respective rationalising collective ends. And there is a description of their

collective end that they all share—giving a gift. On Kutz' account, the parents and the delivery person jointly intentionally gave the gift to the kid. They each intentionally contributed to the same collective end, so each can say that they intentionally gave a gift. And because the state affairs was produced by the three of them, they did it jointly. In this case, the parents are not conceiving of the delivery person as an intentional participant, in part because they are ignorant of her participatory intention. But this appears to be irrelevant to attributing to the three of them the performance of an intentional joint action. In typical cases of joint action, we view those acting with us as acting in the same action as our partners in that action.[125] But Kutz' account gives up on this other-regarding feature because the parents can view the delivery person as merely a facilitator to their joint action rather than including her as a partner. Even if it is true that under some description the delivery person intentionally gave a gift, and this involved other people, it looks as if her action is merely parallel to the parents' joint action. Recall that meeting the togetherness requirement means to parse parallel action from joint action. This case shows that Kutz' minimal account fails to do this.[126]

Though this minimal content account fails to explain togetherness, and so does not resolve the challenge of spontaneity, there are more robust content accounts.[127] I will now consider Michael Bratman's account of shared agency since it is the most comprehensive account available. Originally, he provided an account with necessary and sufficient conditions for shared intention and action.[128] However, his most recent and developed account is less ambitious since it provides only

---

[125] Blomberg notes in a similar vein: "In intentional joint action, each participant cares about the realisation of what the other intends to contribute to because this is what she herself intends to contribute to. They do not treat each other's agency merely as a resource to exploit for their own separate ends" (Blomberg 2016a, 359). Blomberg goes even further by showing how the most charitable reading of Kutz' account is neither necessary nor sufficient for joint action.
[126] Kutz does discuss how collective content helps parse the distinction (Kutz 2000b, 15–16). But, there, the agents share a conception of their collective end—they all intend to bring about the very same thing under the same description. To include this amendment would radically alter the account, namely, it would give up on reduction. And even still, there is the worry about how you can act intentionally jointly with another when you are entirely unaware of their intentions. The means of becoming aware would likely violate the spontaneity requirement.
[127] In between Kutz and Bratman on the content spectrum, there are a variety of accounts that I will not consider. For instance, (Miller 2001; Roth 2004; Tollefsen 2005; Pettit and Schweikard 2006; Alonso 2009; Ludwig 2016).
[128] (Bratman 1993).

sufficient conditions for what he calls "modest sociality".[129] Modest sociality consists in small scale cases of intentional joint action where there are symmetric authority relations between the same members over time and where there is an absence of legal institutional complexities. My focus will be on Bratman's most recent account with reference to previous work where it fits with the current position. It should be noted that Bratman's goal is to vindicate the continuity thesis according to which individual agency contains all that is necessary to explain group agency.[130] This is why he aims to provide only sufficient conditions: even if his account turns out to be partially wrong, a robust enough account of individual agency can sustain modifications to the conditions for shared agency in the account.[131] Further, Bratman's account aims to satisfy three roles that he believes shared intention must play: (1) the interpersonal coordination of action, (2) the planning in pursuit of a group action, and (3) the structuring of bargaining and shared deliberation with regard to the action.[132] Bratman's strategy is to provide a structure of interrelated individual attitudes that can play these special shared intention roles. Despite only appealing to individual capacities and attitudes, their structural relations permit the attribution of shared intention.

Bratman's account has five jointly sufficient conditions:

A. *Intention condition:* We each have intentions that we J; and we each intend that we J by way of each of our intentions that we J (so there is interlocking and reflexivity) and by way of the relevant mutual responsiveness in sub-plan and action, and so by way of sub-plans that mesh.
B. *Belief condition*: We each believe that if the intentions of each in favor of our J-ing persist, we will J by way of those intentions and relevant mutual responsiveness in sub-plan and action; and we each believe that there is interdependence in persistence of those intentions of each in favor of our J-ing.
C. *Interdependence condition*: There is interdependence in persistence of the intentions of each in favor of our J-ing.
D. *Common knowledge condition*: It is common knowledge that A-D.
E. *Mutual responsiveness condition*: Our shared intention to J leads to our J-ing by way of public mutual responsiveness in sub-intention and action that tracks the end

[129] (Bratman 2014, 7–8).
[130] This thesis was introduced in Chapter 1.
[131] (Bratman 2014, 36).
[132] (Bratman 2014, 27).

> intended by each of the joint activity by way of the intentions of each in favor of that joint activity.[133]

The intention condition contains two ways in which the co-agent's attitudes are bound together that result in 'meshed' attitudes. First, you and I each have an intention with the content 'that we J'. I intend that we J and you intend that we J. These intentions can be added together to say 'we intend that we J'. The first-person plural pronoun 'we' in the content of the intentions is read distributively, that is, it merely refers to the conjuction of my J-ing and your J-ing.[134] The distributive reading of the content 'we' is crucial because it eliminates any threat of circularity seeing as it does not imply the concept of shared intentionality. Because of the distributive reading, each of our intentions has a content that refers to the other's intention. But having such intentions is not sufficient for shared intention. Bratman provides 'the Mafia case' to show this: two mafiosos can each have intentions that the two of them go to NYC, but each intends to stuff the other in the trunk. Note that to satisfy these intentions each must act from them and produce their token J-ing. No matter who ends up in the trunk, their intentions will be satisfied by arriving in NYC. For this reason, when we share an intention, we need to view each other as intentional co-participants. By doing this, our intentions are semantically interlocked.[135] This requires us to intend to carry out our intentions "by way of" the other's.[136] I need to have knowledge of or accurately predict your intention and act in a way conducive to *you* satisfying it. This means that I need to treat your intentions with respect insofar as we are intentional co-participants. It is worth emphasising that the

---

[133] (Bratman 2014, 103). This is the compressed version of the account. For the longer version, see (2014, 87).
[134] (Bratman 2014, 41). He notes that we can still appeal to groups in a very loose sense, such as, we in the supermarket are paying. But on its own, this does nothing to get beyond the distributive reading of 'we' because it is an accidental aggregate of people.
[135] Bratman notes two assumptions here: 'we' picks out particular individuals rather than a description and both individuals have a common conception of J (2014, 41–42). Bratman views the common conception constraint very loosely—he mentions that I can intend that we go to the city of the Mets and you the Yankees. What happens if, due to ignorance, co-agents cannot substitute those terms? It doesn't seem enough to merely refer to New York City. Bratman does accept that gross divergence in conception is incompatible with shared intention, but he offers no principle to settle the matter. This will matter when it comes to spontaneous joint action.
[136] (Bratman 2014, 49).

distributive 'we', by avoiding circularity, forces me to view your intentions as a resource in carrying out my intention that coincides with yours. Treating you as an intentional co-participant, negatively construed, means that I do not bypass your intention to satisfy our end.

The second constraint on meshing intentions is for our sub-plans to be responsive to the other's. For Bratman, intentions are plans: that is, to have an intention is to have a plan to do something.[137] Once an agent has an end, she plans to realise it through instrumental means. Some ends require additional instrumental means, or sub-plans, in order to realise them. For instance, if I plan to catch the bus at 7:00AM for campus, then I should plan to pack my bag the night before and to set my alarm to a suitable hour. The difference between sub-plans and intentions is that sub-plans are not settled in the way that intentions are since they will change or vary with the circumstances insofar as they are relevant to an end. Intentions can have greater complexity depending on the number of sub-plans required to realise the end. Now, when each co-agent has an intention that we J, their goal of J-ing will often require sub-plans. If we intend on dancing the tango tonight, one will need to bring the music and the other will need to learn the steps. These actions that fill in the sub-plans contribute to the success of realising the shared goal and must be consistent with one another in order to mesh. If one of our sub-plans changes, in order to achieve our end, the other's sub-plans will need to be responsive to the change. These two forms of meshing intentions partly reveal how two people share an intention. Bratman notes that because the content of each person's intention refers both to the other agent's intention and her own, the intentions involved in shared intention are reflexive.[138] Two important features of shared intention are revealed by meshing intentions. Because each of us intends that we J by way of the other's intention that we J, each of us can expect the other to do her part in bringing about the end seeing

[137] (Bratman 1987)
[138] (Bratman 2014, 52).

as one expects agents to do what their ends require. This expectation, though, is based on prediction—there is no normative constraint on either of us except for the fact that we share an end.[139] In addition, people who act together are often disposed to help each other realise their ends and this is explained by the fact that each agent's intention can only be fulfilled if the other does her part. So, if one agent is having difficulty, the other will typically help because the fulfillment of her intention depends on it. Each person's behaviour with respect to the other is governed by their intentions that we J.

For meshed intentions to have their proper effect, that is, to produce a joint action, we should have two corresponding beliefs (the belief condition). First, each of us must believe that if each continues to hold her intentions, then she will do what she intends to do.[140] This belief is constitutive of having the intentions with the relevant content in a shared intention; that is, it ensures that our intentions will cohere and mesh in the proper ways. If I did not believe that you intend to perform your part, there is no reason for me to be responsive in the relevant way to your intentions. Moreover, it seems difficult to have an intention to do something with someone without believing that the other person intends the same thing. The second belief regarding the relation between the two intentions is that they are "persistence interdependent".[141] It would be irrational for me to continue intending that we J when I know that you have no corresponding intention. I could never realise my intention under such conditions, so each intention's persistence depends on the other to ground a shared intention.[142] Bratman notes that this condition must be read counterfactually such that I can rationally maintain my intention when I hope that you will form a corresponding one or will intend once again after rescinding your original intention. The

[139] Bratman does believes that, in many cases, there is normativity involved in joint action, but it is not necessary (Bratman 1999c).
[140] (Bratman 2014, 76).
[141] (Bratman 2014, 77).
[142] Bratman notes that this dependence condition is not about the onset of intentions, which is a "puzzling idea" (2014, 68), but instead is about the continued persistence of the intentions constituting a shared intention.

counterfactual reading of persistence interdependence makes it, at least, not irrational for me to hold an intention when you currently lack the corresponding intention. Finally, a shared intention exists when we have the appropriate intentions and beliefs, our attitudes are appropriately meshed, they exhibit persistence interdependence, and we know all of this on the basis of it being publicly available. For the shared intention to lead to joint action, each of us must be mutually responsive to the intentions of each and, hence, the sub-plans of the others all while the action is taking place. Insofar as the appropriate attitudes are in place, each of us knows it, and each is responsive to the other in carrying out her part of the joint action, then a joint action has occurred.

Bratman's account does not necessarily fail to explain spontaneity because we do not need to form our intentions via binding interaction, we can simply have the relevant intentions.[143] Indeed, Bratman completely avoids discussing the *onset* of intentions in shared agency and, instead, focuses on what must be in place to credit agents with shared agency. Given that agents can have the relevant intentions without binding interaction, they can also mesh because each of us is or would be responsive to the other's intentions and sub-plans, and they are persistence interdependent. Each of us is disposed to be ready to adapt our sub-plans if something changes. And all of this is 'out in the open'. These adaptations do not contravene spontaneity because we *already* share an intention. Bratman gives a helpful example of being at the symphony and when the orchestra concludes, each audience member applauds together with the others.[144] They applaud together because the wonderful concert was manifestly deserving of applause, so each person forms the intention to applaud with all the others.[145] But the connection between applause and the concert is established by background common knowledge, namely, the common knowledge that those who like this sort

[143] Given that Bratman construes intentions as states that settle deliberative questions, intentions do not typically pop into our heads without reason. This suggests that the agent should have some 'intentional' evidence for forming the intention.
[144] (Bratman 2014, 68).
[145] Bratman makes use of Stalnacker's concept of *manifest event* to establish the 'out in the openness' of the intentions (Stalnaker 2002). The event warrants expectations of the right sorts of intentions.

of music will recognise that this performance was deserving of applause. In this spontaneous case, the audience members have the appropriate intentions that mesh and are persistence interdependent and, therefore, they applaud jointly. It is important to note that the concert only establishes the expectation of the right intentions in others by creating a common ground between audience members based on a background of relevant audience norms. This positions Bratman's account of spontaneous joint actions within the social norm account I discussed at the end of Chapter 1. The problem is that a spontaneous joint action is not possible without an established background of norms. This severely limits the extensional scope of the theory. Without a background of social norms to provide common ground, it is mysterious why we would form the relevant intentions to act together spontaneously or why I would think another is acting with me as opposed to doing something else intentionally. Moreover, in complex situations, there are likely competing norms where deciphering the relevant common ground without binding interaction is extremely difficult, if not impossible. If the relevant common ground cannot be disambiguated from others, then our conceptions of our action may radically differ and, thus, rule out shared agency.[146] It would be hard to count the intentions of others in complex situations as out in the open without engaging in binding interaction. Though Bratman's account does not seem to rely on binding interaction, it appears to be limited in its applicability. This limitation warrants investigating a different approach to spontaneous joint action.

Nevertheless, it is worth delving a bit deeper into some implications of the account. The two implications that I consider here show that Bratman's account requires the revision of some of the common-sense concepts that, as I argued in Chapter 1, should be preserved if possible. I think that these two implications warrant looking beyond content accounts to a different way of meeting the challenge of spontaneity. The first implication is that the imposition of the distributive 'we'

[146] Once again, the issue of common conception that is left unresolved by Bratman plays a significant role here.

forces us to always view our co-agents as bits of the world, bits that facilitate our (collective) ends. I treat your intentions as evidence that you will act as you intend, that is, I predict your future behaviour on the basis of your current intention, and I can *expect* you to act as I predict. However, the kind of expectation I have of you from this predictive point of view is not the kind of expectation that is had of agents. The expectations that we have of agents are grounded in their reasons for action. Indeed, 'expectation' seems to be the wrong word. For instance, when I tell you that I expect to have a coffee in the morning (as result of my making it), my agency is presented as alienated from what I do. Instead, I should say that I *will* have a coffee in the morning, because I will make it so. As a rule, I do not predict my behaviour. Moreover, my commitment to having a coffee is not blind or unguided. I have reason to have a coffee, perhaps because it will wake me up. In a very weak sense, I should have a coffee. From my perspective, when I respond to reasons, I do not expect that I will act so, rather I am rationally *compelled* to do so. This means that when I fail to have a coffee, it is not the case that I fail to do what I expect to do but that I have failed as an agent. From an external perspective, when I expect another to have a coffee *because* she has reason to have a coffee, I do not view her success as the fulfilment of the laws of nature and I do not chalk up her failure as the result of a conspiracy of causal forces. Typically, I employ normative standards to evaluate her actions because it is the mark of agency.[147] This matters for Bratman's account because reasons for action do not enter into his explanation of shared agency at all. And without them, my only means of evaluation of my co-agents is in terms of whether they carry out their intentions by way of mine—whether they act as I expect them to in light of their intentions.

The question is whether we should allow this predictive mode into an account of shared or joint agency. One reason to allow it is based on the distributive 'we'. If we do not allow this

[147] Of course, she may act on stronger reasons of which I am ignorant or she may actually be carried away by causal forces.

predictive mode, then a reductive account of joint agency cannot be given. This may be persuasive, that is, *if* we aim to give a reductive account. The benefit of the distributive 'we' is that we do not need to bloat our ontology by including any collective entities. Its cost, however, is that we have revised our conception of agency. In individual cases, prediction is anathema to agency. But in joint or collective cases, prediction is a necessary aspect of shared agency. Though reductive accounts preserve an austere ontology, they seem to require a revisionary and pluralist conception of agency. Recall in Chapter 1 that the hope was to avoid revision of any of our fundamental concepts with agency taking center stage. Singular and joint agency should express the same agential capacities and features. Bratman's reductive account renounces a unified concept of agency. This may be a necessary cost, but it is surely too high if there is a viable non-revisionary option available.

The second implication of the account is that viewing another as an intentional co-participant is consistent with viewing her as a resource to exploit. This, I suggest, amounts to a revision of 'togetherness'. Recall that Bratman fleshed out the idea of treating another as an intentional co-participant by appealing to interlocking intentions and the 'by way of' relation. Despite viewing each other predictively, we treat each other as intentional co-participants so long as our intentions and actions take account of each other's intentions. Intentions, though, can satisfy the 'by way of' condition without amounting to a shared intention. On the busy subway platform, we each intend to take the train. When the doors open, I could throw you aside to make sure that I get on *this* train. I would satisfy my intention, but I would not respect your intentions by doing so. Instead, I can satisfy my intention to get on the train by way of your intention to get on the train insofar as I respect your intentions. But I still use you as a resource for satisfying my intention because I predict where you will be, what you will do, etc.. Thus, the 'by way of' relation does not actually get us to viewing the other as more than a bit of the world to be exploited. Simply taking a

respectful stance to another's intentions does not make us partners in action.[148] But Bratman will rightly insist that the 'by way of' relation must be coupled with intentions with distributive 'we' contents. The fact that we aim at the same collective end *and* think and act by way of each other's intentions is the difference maker. I treat you as an intentional co-participant in *our* action.

But does this really change my perspective on you? Are you no longer a resource for me to exploit? I do not think so. If I have an intention that you and I J, that is, for you and I both to intentionally perform J, then, from my perspective, I should not interfere with your intentions since it runs the risk of rendering your action unintentional, e.g., interference by coercion or deception.[149] The trouble is that this is compatible with an extreme or sociopathic self-interestedness. Imagine that I am some sort of value solipsist: I am the only agent who is valuable and whose actions have any value. I still believe that others are agents and that they have intentions. But I have no reason whatsoever, unless they affect me, to care about them or their actions since I believe that they have no value. What is of value is whatever is *mine*. Now suppose that I intend to dance with you, and what I want is for you to intend to dance with me. Because of this, I will not force you to dance with me, etc.. I want us to intentionally dance together. If you do not intend the same as me, then I will not have satisfied my intention. In this case, I treat you as an intentional co-participant since I 'respect' your intention and, so, intend by way of your intention and the contents of our intentions interlock. Despite the fact that I treat you as an intentional co-participant in dancing and I am disposed to help you satisfy our shared ends, I still only view you as a resource to satisfy my intentions due to my idiosyncratic worldview. I only care about our dancing because it is *my* end, and my ends are important whereas yours are not. The value of your intention that we dance is

[148] I can take a respectful attitude towards objects as well. Just because a book is not an agent does not mean that I would burn it as a means like any other to achieve my end.
[149] Bratman allows that some degree of coercion or deception is permissible in shared intention, but there is a limit. Perhaps the limit is based on the conceptions of J. But Bratman does not tell us how to parse legitimate divergent conceptions or this limit.

derivative of mine. Even though I view you as an intentional co-participant in satisfying my end, there is still something off when it comes to our being partners in action. The problem is that I do not treat you as an agential equal because I view your intentions as serving my end. Note that this inequality is not about authority, the kind of inequality that Bratman ruled out. It is an inequality of perspective: I only view you as a resource to exploit for my ends. Your intentions merely need to match up with mine in such cases. Given that Bratman's conception of treating another as an intentional co-participant is compatible with this perspective of gross exploitation, this is another revision of the web of concepts in joint agency. It is a revision of the concept of togetherness because the value solipsist, though satisfying Bratman's conditions, never really views the joint action as *our* joint action. Because Bratman's account includes too many revisions of the concepts that we are trying to address, we should look to another account.

4. Conclusion

In this chapter, I have argued that none of the prominent accounts of joint action has been able to meet the challenge of spontaneity. The accounts that appear to be able to explain spontaneity fail to meet the togetherness requirement. Typically, this was because they left agents with insufficient epistemic resources or they are viciously circular. And those accounts that meet the togetherness requirement fail to explain spontaneity, in large part due to relying on a form of binding interaction for yielding agents with an adequate epistemic standing. What could the solution to the challenge be? Has the bar for joint action been set too high? Though it is difficult to meet the challenge, I do not think that it is impossible. Instead, we need to think about action and what is involved in its rationalisation in a different way. Invariably, the accounts surveyed aimed to identify the right kind of mental states that cause joint action in the right way.[150] For some of them, in order to cause in the right way required a degree of coordination, consistency, or interdependence amongst the

[150] Of course, the Anscombean accounts do not present a causal theory.

attitudes. This approach to action explanation displays a certain prejudice: that normative reasons have nothing to do with it. This prejudice is rooted in the distinction between motivating and normative reasons.[151] No doubt this is a fruitful distinction. There is a difference between what happens (intentionally) and what ought to happen. In singular action, it is possible to explain an action with reference to an agent's attitudes alone, e.g., a belief-desire pair or an intention. Indeed, some even believe that action can be explained by evil or the absence of normativity.[152] However, most philosophers, save for Hume perhaps, believe that our motivations can be guided by our normative reasons.[153] The prejudice against normativity's role in action theory is not warranted solely on the basis of the motivating-normative reasons distinction. It's warrant, when it comes to singular action, derives from the fact that normativity does not appear to be necessary, or rather that motivating reasons are sufficient. But the challenge of spontaneity, in addition to challenging the very idea of joint agency, counts against the sufficiency of this merely 'psychological' approach when it comes to joint action. The surveyed accounts that adopt this approach have failed to meet the challenge. My proposal is that the explanation of spontaneous joint action must make an appeal to normative reasons of a particular kind. If the right kind of normative reason guides partners in action, then they can be bound together and act spontaneously.

[151] Naturally, the role of this distinction is not universally accepted, e.g., (Dancy 2002; Alvarez 2013; 2018), though it has wide acceptance.
[152] (Velleman 1992).
[153] Of course, the normativity at issue here is more robust that the normativity produced by instrumental rationality. Moreover, it is a noteworthy challenge to show how this is even possible (Smith 1994).

CHAPTER 3

Acting Together Spontaneously

An adequate theory of joint agency must meet the challenge of spontaneity because, by doing so, the very idea of joint agency along with our natural conceptions of rational agency, togetherness, being partnered, and spontaneity are preserved. We do not need to make any conceptual sacrifices in our thinking about joint agency if the challenge can be met. Recall that the challenge is composed of two requirements on a theory of joint agency: the togetherness requirement and the spontaneity requirement. The togetherness requirement says that an adequate theory of joint agency must explain how two or more people can count as intentionally performing one action with each other—that is, can count as acting together jointly rather that merely acting in parellel. And the spontaneity requirement says that a theory must explain how agents can be joined immediately or without binding interaction. The reason why these requirements pose a challenge is because they do not seem to be able to be met at the same time. The means of meeting the togetherness requirement, such as mutual promises or expressions of willingness, seem to rule out the possibility of spontaneity. Likewise, it is mysterious how the togetherness requirement can be met without some form of binding interaction on the part of the agents because they need to possess some epistemic standing with respect to their joint action. In Chapter 1, I argued that the key to both requirements is the *co-agential perspective*, which agents can occupy independently or together. The co-agential perspective shines a light through the difficulty of the challenge because agents who occupy it together, that is, they legitimately occupy the perspective on each other and their joint action, are bound together in a way that can produce spontaneous joint action. A theory that can explain how agents occupy the co-agential perspective together will meet the togetherness requirement. Moreover, if a theory can explain how agents can legitimately occupy the co-agential perspective without binding interaction, then the spontaneity requirement can be met. The question is: how can

agents *legitimately* occupy the co-agential perspective? The task is to elucidate the requisite resources and to offer a theory that meets the challenge of spontaneity.

After the failure of extant accounts to meet the challenge of spontaneity, it is worth pursuing my suggestion from the end of Chapter 1, namely, that we must look beyond only the attitudes of the agents involved in the joint action in order to meet both the togetherness requirement and the spontaneity requirement. Recall how I suggested that social norms would not do the trick because, first, they would unduly limit the scope of spontaneous joint actions and, second, that social norms themselves seem to depend on joint action, thus, incurring a vicious circularity. Instead, I proposed that we look beyond the attitudes which constitute motivating reasons towards normative reasons for action. Normative reasons are facts about what an agent ought to do, and their truth does not depend on whether an agent actually acts on them. Nevertheless, an agent responding appropriately to her normative reasons for action is put into the motivated state of mind for which the reason calls. Unfortunately, the term 'normative reason for action' is ambiguous with respect to the agent(s) to which it applies. Normative reasons for an *individual* to act will not help us meet the challenge of spontaneity. For instance, if I have reason to go for a walk and respond appropriately, then I will not engage in a joint action because the action that I am rationally compelled to perform is not a joint action but a singular action. Indeed, individual reasons typically make no mention of other people and they are not typically under the authority of those reasons. What is needed, rather, are normative reasons for *joint* action or, as I will refer to them for short, *group reasons*.[154]

Normative reasons for joint action call for more than one agent to perform an action (jointly) together. For example, the clear, sunny sky is (group) reason for *us* to go for a walk *together*. Note how normative reasons for joint action directly apply to more than one person, in that they are

[154] When I say "group reason", I mean a normative reason for joint action. These are equivalent terms for me. When I use "group reason", I am *not* referring to the reasons of groups over and above the individuals who perform group-singular actions, such as supreme courts, governments, clubs, and, in general, institutions and corporations.

under its authority, and the recommended action is a joint action as opposed to an individual action. One might be concerned, in the context of an explanatory project, about an appeal to normative reasons, whose primary job is typically to justify. I respond to this concern in section 2, where I provide the *reasons account of joint action*. In short, the model of agency employed here does not sever normative reasons from motivating reasons and, moreover, the appeal to normative reasons fleshes out the content of our motivating reasons.[155] Indeed, we want a model of agency that has a robust connection between what we ought to do and what we do or are motivated to do.[156] What would the point of ethics be if its requirements cannot be put into action by agents? Normative reasons do not merely have a justificatory role, but they should *guide* agents by their normative authority. As I will argue, group reasons bind agents together by directly applying to more than one agent and guide them towards their joint action. Group reasons provide the normative structure to understand how agents view each other as partners and are moved to act together jointly.

This chapter has two primary aims. First, in section 1, it explains how a theory that incorporates normative reasons for joint action can meet the challenge of spontaneity, thereby preserving our intuitive conceptions of togetherness, being partnered, rational agency, and spontaneity. Second, in section 2, I develop a theory of joint action on the basis of group reasons, the reasons account of joint action, that explains spontaneous joint action. And, in section 3, I discuss how the reasons account of joint action is not susceptible to some objections typically levelled against theories of joint agency. In particular, I discuss concerns about circularity and parsimony. The section concludes by raising a concern for the reasons account.

[155] This model is based on Donald Davidson's later philosophy and developed by Robert Myers (Myers and Verheggen 2016). This model is not the typical Davidson, because the way Davidson is typically understood is not how Davidson in fact developed his practical thought. Davidson is an extremely systematic thinker and his concern with the holism of the mental carries over to pro-attitudes. More will be said about Davidson later, especially, in Chapter 5.
[156] This is one part of Michael Smith's 'moral problem' (Smith 1994).

1. The Challenge of Spontaneity, the Co-Agential Perspective, and Group Reasons

In order to meet the challenge of spontaneity, a theory of joint agency needs to show how co-agents can legitimately occupy the co-agential perspective together. By doing this, a theory satisfies the togetherness requirement because legitimate occupation of the perspective requires that co-agents are bound together, and it satisfies the spontaneity requirement because legitimate occupation means that co-agents did not exploit each other as resources to bind together and, so, can act together immediately. What warrants or legitimates their occupation of the co-agential perspective is a group reason that applies to them. Three questions must be answered in order to show how group reasons play this important role. First, how do group reasons bind agents together? Second, how do they help to do that spontaneously? And, finally, how do they warrant or legitimate co-agents' occupation of the co-agential perspective together?

Group reasons have two primary functions. The first is to recommend or guide agents in the performance of a joint action.[157] And the second is to bind agents together by having direct normative force over more than one agent.[158] The simple fact that group reasons directly apply to more than one agent binds those agents together as the agents of a potential joint action. To see this, it will help to define a group reason:

> A reason for action is a group reason iff (1) it is a consideration that counts in favour of φ-ing; (2) it is a consideration for at least more than one person to φ; and (3) φ is a joint action.[159]

The first condition is fairly straightforward as normative reasons generally are defined as considerations that count in favour of acting in a particular way.[160] The second condition states that

---

[157] The strength of the normative reason will depend on the particular reason, agents, circumstances, etc. The point is that that group reasons exert some degree of normative force on the agents.

[158] I use 'direct' here to indicate the agents of the action. As we will see in Chapter 5, a reason may exert normative force indirectly to other agents who are not the agents of the recommended action.

[159] The account offered in section 2 might seem circular based on this definition: joint actions are explained in terms of group reasons, while group reasons are defined in terms of joint actions. This objection is discussed in the final section of the chapter.

[160] See (Scanlon 1998, chap. 1). Of course, there are normative reasons for belief as well, but I bracket those off for the time being.

the agents for which the consideration is a reason is a set or group of individuals rather than one agent. There is nothing in principle in favour of the thought that reasons only apply to single individuals as opposed to a plurality of individuals. In Scanlon's reason relation, the agent variable can be filled by any thing that is rationally responsive, that is, any entity or entities that can count as an agent or agents.[161] To be clear, this does not mean that the plurality of individuals is a distinct ontological entity, a so-called group agent. Instead, the group reason picks out individuals who are in circumstances that provide reason to act together with others. For instance, the group reason to demonstrate against French occupation picks out the native Algerians and tells them to protest jointly. The final condition on group reasons is that they recommend a joint action as opposed to an individual action. It would be absurd for a group reason to pick out a group of individuals to perform an individual action. Seeing as only one person can perform it, what would the others be required to do?[162] Ultimately, a group reason is a consideration that counts in favour of the performance of a joint action by a plurality of individuals. Group reasons help to explain how agents are bound together because they *already* bind them together in the relevant sense. The group reason already picks out the individuals as the co-agents to act together. Insofar as each co-agent were to act on the group reason, they would be bound together in performing the same action because they are acting under the same intentional description and for the same reason. On the basis of the group reason, each co-agent can assent to the fact that they are acting together intentionally, and each co-agent understands each of their part-actions *as* contributions to the joint action. The normative group reason provides the description under which each co-agent knows her

[161] The reason relation states that a reason is a relation between a fact, a set of circumstances, an action to be performed, and an agent who performs the action. See (Scanlon 2014, 30–33).

[162] It is possible that one individual or a subset of the group could perform all the part actions required for a joint action, e.g., go buy the paint, rent the ladder, and paint the house all on her own. This possibility, though, does not indicate that a joint action is not recommended—it just means that the others are failing to do what they ought to do and one person is picking up the slack. Further, this possibility only works for a kind of joint action that can be broken up into different part-actions. It is impossible for the other kind of joint action where the individuals perform the same part-actions, e.g., walking together—an individual cannot walk together by walking alone.

action as part of a joint action and she is warranted in thinking that others are similarly acting on it because others are part of the 'we' to whom the group reason applies. That is, responding to a group reason that you share with another warrants your occupation of the co-agential perspective—you can legitimately think of the relevant others as your partners.

However, agents may fail to respond to their group reasons. Why should we think that group reasons *already* bind individuals together without their acceptance of the reasons? Consider a case where Shaq who is too heavy for any one individual to lift falls on the sidewalk and needs help to get up. You and another person are in the vicinity of the accident. Imagine further that there is no one else around and the other walks by Shaq while recognizing that he needs help. Intuitively, Shaq's falling and inability to get up provides reason for us to help him and because neither of us could help him up alone, it seems that we share a reason to help him up together. When the other walks past him, you do not typically think that she only failed to help him up but that she also failed to help *you* do what you ought to do. How could she leave you in a position to not be able to help Shaq to his feet? The fact that she has failed *you*, in addition to the fallen person, is an indication that the two of you were already bound together by the group reason to help him to his feet despite the fact that the other never accepts or recognises the group reason. Otherwise, all you would be permitted to think is that this rude person failed to help Shaq. But the other is not just failing in what she ought to do but she is failing *you* in what you together ought to do. The normative group reason grounds demands between co-agents such as these whether they are in the process of acting together or not.[163] The fact that group reasons help to explain these kinds of normative demands and interpersonal failures indicates that group reasons bind agents together irrespective of how agents respond to their reasons.

[163] Of course, there are overriding reasons that each individual might have at any given moment that can 'silence' a group reason. But that depends on settling what each agent's all things considered reason is as opposed to what sufficient reasons they have.

Second, group reasons bind us together spontaneously and allow us to act together spontaneously because we do not need to interact with each other to respond to them.[164] So long as agents have access to their group reasons without binding interaction, they can respond to their group reasons spontaneously. Consider the Algerian case: the circumstances of their occupation make the group reason to resist together patently clear. They did not need to interact with each other to recognise their group reason to protest—their reason was clear as day. By recognising their group reason in this way, they acknowledged that they were bound together. And from it, so long as others were responding to it, they could act together spontaneously. There is no need for co-agents in such a case to exploit their fellow co-agents as evidence and, paradoxically, to view them from two incompatible perspectives at the same time. Instead, they can view them from the co-agential perspective upon recognising their group reason—that is, they can be viewed purely as partners with whom they are engaged in joint action. The recognition of the group reason can move agents to action without the mediation of binding interaction.

Finally, group reasons help to explain the requisite warrant in occupying the co-agential perspective and acting from it because it is part of their function *as reasons* to warrant inferences in joint agency. Recall that occupying the co-agential perspective legitimately means that the agent is warranted to draw inferences about acting with a partner and in fact acting. And one of the primary functions of group reasons is to guide co-agents to perform a joint action and to infer the relevant thoughts about their joint action. By responding to a group reason, an agent is warranted in thinking about those who are also picked out by the group reason *as* partners. She may also infer, on the basis of the group reason, the relevant agential expectations that occur in joint action. For

[164] There are some special cases that complicate this point, such as joint actions that require consent. However, consent being given can create the relevant group reason to act together to which co-agents respond. The binding is done by the reason and not the interaction. These cases should not be viewed as counter-examples to my overall account because responding to the group reason is not treating your co-agent as evidence. Without the appeal to group reasons, we would have to treat our co-agent as evidence. On the reasons account, the prior interaction does not perform the binding role, so it does not count as binding interaction.

instance, I expect the other person on the street to help me raise Shaq to his feet because the group reason dictates that we help Shaq together. And, most importantly, she is warranted to perform her part of the joint action.[165] Group reasons warrant co-agents to occupy the co-agential perspective together because they guide agents to think of each other as partners and to act together.

Group reasons bind agents together. They also bind them together spontaneously. By responding to a group reason, co-agents can legitimately occupy the co-agential perspective and be warranted in inferring joint-action-relative thoughts and undertaking their portion of the joint action. Indeed, when two agents respond to the same group reason and occupy the co-agential perspective together, they are in fact partners. Moreover, by occupying the co-agential perspective legitimately, co-agents can be warranted to have joint-action-relative expectations because they are governed by the same group reason. Group reasons, then, help an account of joint agency to meet the challenge of spontaneity because they can bind agents together spontaneously. A joint action that is guided by a group reason can be a spontaneous joint action. Group reasons make spontaneous joint action possible and, thus, make good on our natural conceptions of togetherness, being partnered, spontaneity, and rational agency. By incorporating group reasons into our thinking about joint agency, we can resist any need for revision, and it certainly saves us from the threat of pluralism about rational agency.

2. The Reasons Account of Acting Together, Spontaneously

Seeing that group reasons can help to meet the challenge of spontaneity, it's now time to see how they can be incorporated into an account of joint agency. What's needed now is to understand their role in the explanation of a particular spontaneous joint action. I offer the *reasons account* of joint action. When co-agents act together spontaneously four conditions are satisfied:

[165] I will discuss a complication with this point in section 2.

A. Each co-agent recognises that we have normative reason to φ together. (Binding Condition)
B. Each co-agent believes that for us to φ together, individual actions of the type A/B…Z must be performed. (Joint Instrumental Condition)[166]
C. Each co-agent believes that for us to φ together, I ought to do my part, that is, individual action A/B…Z, of φ. (Individual Instrumental Condition)
D. Each co-agent desires to do her part, i.e., A/B…Z. (Motivational Condition)[167]

The basic idea is that in virtue of each co-agent recognizing a group reason, one that they share and that binds them together, each can derive an instrumental belief to perform an individual action, their part, and each desires to perform her part on the basis of their group reason. Their desires and beliefs would explain why they act as they do, and the instrumental relation between their individual instrumental belief and their joint instrumental belief explains why those individual actions are united and constitute a joint action. Instead of relying on some psychological feature to bind the agents together, the group reason does it—it picks out the agents who should act together.[168] When each of them forms the relevant beliefs and desires, on the basis of their group reason, and acts, they are bound together in the appropriate way such that their actions can be described as intentionally acting together. Even though they are immediately acting from a desire for an individual action, there is a rational, normative structure governed by the group reason in which the desire is constituted. When asked what she is doing, she can respond by describing her part-action as "We are φ-ing" just as when making a cake, you can describe measuring the flour as making a cake. Further, there is no barrier to understanding how co-agents could act together spontaneously on this account: so long as each responds to their group reason, there is no need for them to interact to

[166] I use '*A/B…Z*' to describe the two possible types of joint action. First, 'A' stands for one type where the co-agents are doing qualitatively the same thing, for instance, walking when walking together. Second, 'B…Z' stands for the other type where co-agents perform different actions which constitute the joint action, such as, the variety of actions that constitute protesting together. In such cases, one co-agent does, for instance, actions A and B while the other does C and, taken together, these actions can be described as φ-ing together.
[167] I note that there is no explicit condition for obligations between co-agents. However, group reasons would establish the kind of obligations that are incurred by legitimately occupying the co-agential perspective. See Chapter 1.2.
[168] The prominent accounts of joint action try to explain what binds agents together internally, or by psychological features that need to be related in relevant ways. See (Gilbert 1989; Bratman 2014; Tuomela and Miller 1988).

settle their intentions. The group reason provides them each with the rational direction that they need and, by acting on that group reason and not some other reason, they all count as acting together spontaneously.[169] Because the group reason contains the relevant information, i.e., who should act and what should be done, co-agents do not need to interact before acting together—they can already view their co-agents *as* co-agents.

To get a clearer sense of how the theory works, let's apply it to the Algerian spontaneous protest case. On the fateful morning, each Algerian recognises the group reason that they share with other Algerians. They have (group) reason to protest and demand their independence together (Binding Condition). As a result, they occupy the co-agential perspective together and view each other as partners in their fight for freedom. In order to satisfy their group reason, they must protest and demand their independence. As a result of recognising their group reason, each Algerian believes that each one of them, as a matter of their own actions, must go into the streets and demand their independence (Joint Instrumental Condition). From the normative group reason and the joint instrumental belief, each Algerian forms the belief that she ought to protest and demand independence (Individual Instrumental Condition). Because each views the performance of her part of the joint protest as necessary to the performance of the joint action, it is instrumentally supported and made intelligible by the group reason that picks her out as one of the many who ought to protest together. On the basis of this normative judgment that she ought to do her part in the protest, she forms the desire to protest and demand independence (Motivational Condition). The group reason along with this set of attitudes explains why the Algerians are moved to the street to protest and demand independence together. And they are not moved merely as individuals, they are

[169] One might worry here that by just responding to the group reason, the co-agents may never perform their joint action in some cases because there will be confusion over who needs to do which part. For instance, two co-agents may opt to do the same part-action while leaving a crucial part-action unperformed, thus rendering the joint action unsuccessful. But that is all that such confusion amounts to—an unsuccessful joint action. It does not mean that they were not intending to act together while performing their individual actions.

moved together. The account preserves the spontaneity of this joint action because the group reason binds the Algerians together and by recognising it, they can act jointly immediately, without binding interaction.

It is worth pausing to see why the Algerian protest is not an aggregation of parallel actions when their action is guided by their group reason. It is possible that each Algerian heads into the street in order to protest for her freedom independently. The fact that each of them independently are doing the same *type* of action at the same time makes it look as if they are acting together. Their critical mass obscures the ontology of their actions. However, when each of them independently waves their flags and marches on the guards, they are merely acting as individuals whose actions are in parallel. There is no joint action. Why does the group reason and their response to it change the situation? When the Algerians respond to their group reason, they are no longer merely performing the same type of action, they are acting in one action together jointly. The desire that compels them to each wave her flag and march is rationally constituted by her individual instrumental belief to do her part. And her individual instrumental belief derives from her joint instrumental belief about which actions must occur in order for them to protest jointly. Moreover, these beliefs all derive from the fact that each Algerian who is protesting together has recognised that they share a group reason to protest together. They together view each other as partners in protest from the co-agential perspective. The group reason provides the normative structure that unifies their individual attitudes and actions into a joint action. Though their behaviour may look the same from a distance, when they respond to a group reason, we can describe each Algerian's action as intentionally contributing to their joint action. In the parallel case, the Algerians do not have any special view on the other protestors. In the joint case, they view each other as partners and, partly in virtue of this, we can describe what they do together as an intentional joint action.

Before arguing for the reasons account, I want to discuss two concerns that typically arise in theories of joint action and show why they do not apply here. First, I discuss what is sometimes called the 'own action' condition on intention. This condition states that an agent can only intend her own action. Though this is not a universally accepted condition on intentions, if true, it causes trouble for some theories of joint action because it seems that co-agents must intend that 'we act' in order to act together.[170] Now, the reasons account does not appeal to intentions, but judgement-sensitive desires do perform many of the same roles as intentions. Nevertheless, the reasons account can avoid the 'own action' concern because, in this case, judgment-sensitive desires do not have plural content. What each co-agent desires to do is her part of the joint action. The plural content in the reasons account occurs in the accompanying beliefs, and beliefs can easily accommodate plural content. So, the reasons account of joint action is not threatened by the 'own action' condition.

A second concern is with the so-called 'settle condition' on intentions.[171] The settle condition causes trouble for theories of joint agency that appeal to shared intention. If joint action is caused by *an* intention that is shared by multiple agents, how is it in an agent's power to settle a shared intention? The problem arises for these views because for agents to be bound together, that is, to meet the togetherness requirement, requires that they all settle their shared intention and it may not be in their individual powers to do so. However, the reasons account does not appeal to shared intentions. What binds agents together is a normative group reason. The co-agents do not need to *do* anything together, such as settle an intention together, in order to bind themselves together. Nevertheless, in Chapter 1.3, I did make an appeal to the rational power of settling our intentions.[172] Indeed, it played a crucial role in the agency argument for the co-agential perspective. There, I said

[170] (Bratman 1999b; Ferrero 2013)
[171] See (Velleman 1997; Bratman 2014) for discussion of settling and joint action.
[172] I use 'intentions' loosely to cover judgment-sensitive attitudes.

that rational agency is in part characterised by an agent's power to settle which reason becomes her operative reason, instead of the world determining it for her. The reasons account does not threaten this picture of rational agency because the agent's rational power of settling 'intentions' is not tied up with binding agents together. Instead, agents are bound together by the group reasons that they share. That is a fact of having group reasons. It is still up to each agent to determine which of her reasons she acts on, that is, each still has the rational power of determining her own 'intentions'. She has the power to determine for herself whether she acts on her group reason. With these concerns out of the way, let's delve into each condition of the reasons account in more detail.

A. Each co-agent recognises that they have normative reason to φ together. (Binding Condition)

In order for us to act together spontaneously, we must each recognise the group reason that we share for three reasons. First, even though the group reason binds us together *already,* it would not lead to joint action unless it is recognised by the relevant agents. The group reason must have psychological uptake in order to play an explanatory role in the joint action. Second, recognition of the group reason causes co-agents to occupy the co-agential perspective—it makes us think of each other as partners. By recognising the reason, I recognise that I am bound to others, my partners, and so do they. And, from there, we are rationally compelled to act on its basis. The fact that co-agents recognise their group reason allows them to distinguish what they do as jointly done rather than merely parallelly done. It is the recognition of the group reason that allows identification of their action as properly joint. Third, I have emphasised the role of viewing each other as partners, but recognition of the group reason also supplies the appropriate description under which co-agents count as acting together intentionally. This description must be available to them as the co-agents of the joint action in order for them to be warranted in instrumentally relating their part-actions to the joint action. Imagine one co-agent is struggling to do their part. Because they are acting together intentionally, the other co-agent is warranted to help the other out within the context of the joint

action such that their joint action can succeed. The reason for helping the other is based on the fact that they share a reason to act together. Even if there are other reasons to help her co-agent, the helping action counts as part of the joint action, as opposed to some non-joint description, because she is doing so intentionally on the basis of the group reason. Thus, they must judge that they have reason to act together in order to have the rationally supported description available to them that warrants the connection between their part-actions and the joint action. For these reasons, in order to explain a joint action, the co-agents must recognise that they ought to φ together, that is, co-agents are related to their group reasons via recognition.

But if an attitudinal uptake of a group reason is necessary, then it is possible for co-agents to misapprehend their group reasons. They may act on no reason at all. This would appear to sever the attitudes from the normative support of the group reason. If this were true, then this condition rules out the possibility of acting together (spontaneously) when the co-agents misjudge their group reason.[173] For example, two muggers judge that they ought to mug this single mother together, but it is false that they ought to mug her together. The objection presses the reasons account to explain how it handles misjudged group reasons.

The first thing to consider is the deliberative stance of the co-agents. Agents have many reasons to act in a given situation, even though an agent may only consider a subset of them. Clearly, many of the reasons that we act on are not *conclusive* reasons and conclusive reasons, facts of the matter, may not be arrived at via actual deliberation. Nevertheless, the reasons with which an agent deliberates are pro tanto reasons, reasons which can be defeated by other reasons and yet are still normative. When co-agents judge that they ought to act together, sometimes they get it right and act on their conclusive reason but, at other times, they get it wrong and misjudge a pro tanto reason

[173] This objection might come from those who defend factive rationalisation: they insist that only facts can be reasons for action, e.g., (Alvarez 2018). After all, what does a falsehood explain?

for a conclusive one. The point here is that incorrect judgments about reasons do not always eliminate the normative force of the group reason. Pro tanto group reasons still have the power to bind agents together due to their normativity. Most cases of joint action will be based on this kind of normative reason. When two people ought not, all things considered, mug a single mother, this does not preclude the possibility of their being a pro tanto reason, however weak, to mug her. Pro tanto reasons appear to agents so long as there is *something* to say in favour of the action. That consideration may be so strong that it is imperative to respond or it may be so weak that you barely consider it in deliberation. But the (severe) weakness of a pro tanto does not eliminate its normative status. Hence, co-agents who respond to a group reason that is clearly very weak can still be bound together by it because there is still a normative fact of the matter that binds them together.

But what if there is not even a pro tanto group reason, as some might say in the mugging example?[174] There is *no* reason to do *this* together, and yet co-agents judge that they have group reason to do it together. In this case, the content of the attitude alone should not be able to bind agents together in the required way. To provide the warrant required by the co-agential perspective, the judgments require the support of a normative group reason. This is a tough question and I may need to bite the bullet. But I believe that the bullet is miniscule given the discussion in the previous paragraph. The number of joint actions that will not have a pro tanto group reason in favour of it will be negligible. However, I think that there is a satisfactory response to this question, though I cannot explain the argument in great detail here.[175] All that I can do is provide some broad strokes to, hopefully, allay the worry for the time being. The idea is that to have judgments about reasons at all depends on one agent interacting with another while focusing on some part of the world.[176] This

[174] This question recalls Anscombe's saucer of mud example (Anscombe 2000, sec. 37).
[175] The triangulation argument is laid out in Chapter 5.
[176] The response depends on Davidson's triangulation argument applied to reasons. For this argument, see (Davidson 2004c; 2004a; 2004b; Myers and Verheggen 2016).

triangular situation fixes the contents of the agents' judgments because each is caused by the world to think what they do and is in a continual process with the other to correctly represent the world. In this process, agents can have many incorrect judgments and move towards correct ones. With regard to normative reasons, the interacting triangulators are caused by a normative property in the world to have a normative thought. What they *think* they have reason to do may be incorrect, but they are still tracking normativity because they constitutively aim to get normative matters right. Despite their error in content, the background of normativity can still explain their joint endeavour because it will still, to their minds, warrant rational instrumental moves to perform their part of the joint action. The suggestion is that agents can be tracking an appropriate group reason and be compelled by its normative force while misapprehending its precise content. In the end, I think that this objection can be assuaged so long as the triangulation argument succeeds.

It is worth pausing to discuss the epistemic constraint on joint agency that was introduced in Chapter 2. I criticised Searle's we-attitude account of joint action because it lacks an epistemic relation between co-agents—all that was required to be joined together was to possess we-attitudes with the right content and reference. I argued that the successful reference of these attitudes was insufficient to explain the joint sense of together because the agents did not have any epistemic standing with respect to the other's mental states. When they act together, co-agents must be warranted in thinking that they are acting with others. How does the reasons account accommodate this epistemic constraint? In virtue of responding to a group reason that we share, I am warranted in thinking that my co-agent who shares this reason with me is acting on the group reason. I know that my co-agent should be acting with me because of our group reason. The group reason provides me with sufficient epistemic standing with respect to my fellow co-agent. Moreover, it is too high of a standard to require *certainty* that another is performing her part of a joint action. Of course, we may fail to succeed in our action for a variety of reasons. But my co-agent's mind is not transparent

to me, and I cannot expect for it to be so in order to be warranted in thinking and acting jointly. Instead, responding to a group reason should satisfy the epistemic constraint on joint agency because it warrants occupation of the co-agential perspective and thinking of another as a partner.

> B. *Each believes that for us to φ together, individual actions of the type A/B…Z must be performed.* (Joint Instrumental Condition)

The joint instrumental condition says that co-agents must have a belief about the actions that would collectively lead to their successful joint action. They must have a conception of how the joint action will proceed. This belief is a rationalising bridge between the joint action recommended in the group reason and the action that a co-agent performs as her part of the joint action. For instance, the Algerians believe that carrying flags, filling the streets, and chanting jointly constitute the joint action. Or, when we go for a walk together, we each have a belief about what each of us must do, that is, walk, so that we can walk together.[177] In both cases, this collective instrumental belief provides an understanding of the necessary part-actions, irrespective of what a co-agent personally does, in order to understand the success conditions of the joint action. In such cases where the co-agents perform the same type of action, e.g., walking (in order to walk together), there is little room for misunderstanding which type of part-action needs to be performed for the joint action to occur. If I did not believe that both of us had to walk (in the appropriate way) in order to walk together, it is hard to understand my conception of *walking together*. However, cases where different action-types are performed as part of a joint action are a bit more complicated because at least one or more co-agents are performing different actions than another co-agent. For instance, if we ought to paint the house together, of course there is an action-type, 'paint the house', that we could share in performing; but other part-actions might also be necessary, such as buying paint and

[177] The first case involves the performance of different action-types that constitute the joint action. The second case involves co-agents who perform the same action-type.

climbing a ladder. In these cases, each co-agent must believe that there is a set of part-actions necessary for the completion of the joint action.

This condition is necessary for three reasons. First, an action theory requires that an agent has an instrumental belief that can explain why what she is doing now counts as (part of) another action, her intended action. There must be an instrumental relation between what the agent does and what she intends to do so that what she does is rendered *intelligible* as part of or as identical to the intended action. For instance, the fact that I raise my hand to the light switch tells us nothing about whether I am intentionally turning the light on. The instrumental belief that connects my raising my hand to turning on the light in part explains why I move my hand towards the switch and why my moving my hand *counts* as turning on the light or, at least, attempting to do so. The same rationale is employed in the context of joint action. For an agent's action to count as part of another intentional action, the joint action, she must believe that her action is instrumentally connected to the joint action. However, in the joint case, each co-agent must be aware of at least some of the other requisite part-actions involved in the joint action, not only her own. The reason for this is that she cannot perform a joint action on her own and, so, she must have some conception of what the other co-agents must do for their joint action to succeed. It would be irrational for her to believe that her action alone could bring about the joint action. Indeed, the intelligibility of her instrumental action is suspect. Without a conception of other part-actions that serve the joint action, her participation in a joint action would be irrational. In non-spontaneous joint actions, the necessary part-actions might be up for discussion, such as when we plan on where to eat dinner together. But co-agents cannot proceed in the joint action until they have a conception of what their joint action looks like.

The second reason for the joint instrumental condition is that it is distinct from the individual instrumental condition. Whatever I believe is *my* part of our joint action is a distinct belief

from what I believe that *we* must do in order to succeed in our joint action. Co-agents each have an instrumental belief, that is, the individual instrumental condition, about what she must do as a *co-*agent, as an individual who is a group member. And this is a different belief from what *we* as a plurality must do in order to bring about the joint action. For instance, for us to have a coffee together, we might know that we must leave our house, head to the café, order coffee, and drink the coffee (together). This is a distinct instrumental belief from the individual instrumental belief that *I* must leave our house, etc., in order for us to have coffee together.

The third reason for the joint instrumental condition is that co-agents must have an understanding of the part actions instrumental to their joint action as a whole in order for group reasons to play their interpersonal, normative role, such as justifying scrutiny, praise, etc. of other co-agents. Suppose that this condition was denied and each co-agent had to possess only her own individual instrumental belief. How could one of them demand of the other that she do her part when she fails to do so? It seems entirely unwarranted without each co-agent having an understanding of the necessary part-actions of the joint action. And this is crucial to the rationality of joint action: if one co-agent is not doing her part, then it might not be rational for another to perform hers.[178] So, in some cases, for a co-agent to have an appreciation of the rationality of her part-action depends on her knowing what else must happen for the joint action to be successful.

But, to what degree must co-agents believe the same necessary part-actions to act together? At what point do their conceptions of the joint action diverge too much? Imagine a case where we are painting a house together but differ in our joint instrumental beliefs in terms of specificity. We each believe that paint needs to be bought, a ladder rented, and the house painted, but I

[178] Cases like these typically involve the moral caveat that if only one person does her part, then it is in fact worse than no one acting or both acting. Classic cases are the 'coordinated attack problem' and the 'electronic mail problem', see (Halpern 1986; Rubenstein 1989; Binmore and Samuelson 2001; Chant and Ernst 2008). I discuss this complication more in section 4.

idiosyncratically believe that the paint must be bought from a particular store and one side of the house needs painting first. It seems as if the co-agents have different joint actions in mind because they do not share a conception of the joint action. How could they act together? Accounts of joint action will have to allow for some variation in the thoughts of the co-agents, otherwise joint action would be too intellectually demanding in most cases. If co-agents had to share identical joint instrumental beliefs, then joint action would fail the moment that the world no longer cooperates. For instance, when one person gets a flat on the way to buying paint, does the other co-agent really need to know what she will do in order to get paint? What is relevant is that she buys paint, not necessarily where she buys it, what route she takes to the store, etc.. All things being equal, when there are a variety of ways to perform a part-action, it is irrelevant that each co-agent knows which specific course of action is taken. Some specific courses of action, though, will be relevant to know, such as when one course of action rather than another is detrimental to the joint action. But here all things are not equal and there is reason against taking that course of action on the basis of the group reason. So, I believe that co-agents do not need to share identical joint instrumental beliefs—they only must share those which are relevant to the joint action.

But what if co-agents' beliefs do not differ in specificity but rather in which part-actions are required? That is, do co-agents share a conception of the joint action if each co-agent believes that different part-actions are required? For example, I believe that we need paint and a ladder in order to paint the house, whereas you believe that we need paint and brushes but not a ladder. Now, here it really does seem as if the co-agents have different joint actions in mind, that is, different conceptions of how the joint action will proceed. However, successfully acting together is not the same as engaging in joint agency, spontaneously or not. To successfully perform our joint action, it certainly helps that co-agents share identical joint instrumental beliefs—things will go much more smoothly if there are no incongruencies. But the question is whether co-agents are sufficiently

bound together when they differ in joint instrumental belief. And they are, by the group reason that they share. Logically, the group reason, and their judgment about it, is basic when engaging in joint agency. Insofar as each aims to perform her part of the joint action on the basis of the group reason, they are sufficiently bound together. The fact that there is disagreement, explicit or not, around what must happen does not eliminate the fact that they are engaged together in an activity. When co-agents have different joint instrumental beliefs and yet are still acting on the same group reason, they might act together badly and, very likely, will not succeed in their joint effort. This is most clear in non-spontaneous cases: just because we do not have identical joint instrumental beliefs does not mean we cannot come to a later agreement about the best means to our end. There are many reasons co-agents might differ in joint instrumental belief, especially in spontaneous cases: they might see the lay of the land differently; one might have more knowledge of the situation that favours one course of action over another; one might have more know-how in a relevant area; etc.. None of this rules out the possibility of joint agency, even if it makes it very difficult to succeed at it without coming to agree in joint instrumental beliefs.[179]

> C. *Each co-agent believes that for us to φ, she ought to do her part, that is, individual action A/B…Z, of φ.* (Individual Instrumental Condition)

The purpose of this condition is to connect a constituent part-action from the collective instrumental belief to each one of the co-agents such that she judges that she ought to perform *this* individual action as her part of the joint action. For instance, on the basis of my judgment of the group reason to walk together with you, I further judge that I should walk at the relevant time and place in order to walk with you, and you also judge that you should walk at the same time and place in order to walk with me. How do co-agents derive these individual reasons from the group reason?

[179] Nonetheless, the degree of permissible divergence between different joint instrumental beliefs will be constrained by the group reason. If the conceptions differed in a way that made the success of the joint action impossible or incompatible with the relevant concept of the joint action, then the co-agents should not be considered to be acting together.

To answer this, more needs to be said about instrumental reason. A reason for action provides an agent with further instrumental reasons to perform actions required by the basic reason on the basis of instrumental rationality.[180] Consider a case of Nagel's: if I am going to Italy in six weeks, then I have to speak Italian in six weeks. If I have reason to speak Italian in six weeks, then I have reason *now* to learn it because I do not already know it. The thought here is that in order to do what an agent has reason to do, the basic reason provides her with further reasons to perform actions that would ultimately satisfy the basic reason. The derivative instrumental reasons are the ones that guide the agent along the rational path to satisfying what she has reason to do. Applied to joint agency, the group reason is the basic reason that then provides further instrumental individual actions to the co-agents. For instance, because the Algerians share a group reason to protest together, they are provided with further individual reasons to fill the streets and wave flags of independence, which in turn instrumentally make it the case that they perform the joint action. The rationality of the instrumental individual actions depends on the group reason.

A question arises here regarding how each co-agent receives her *particular* derivative individual instrumental reason. That is, how do individual instrumental reasons get distributed amongst co-agents? And why doesn't each derivative reason apply to each co-agent *equally*? If this were the case, it is possible that each co-agent selects the same derivative reason and performs the same part-action, resulting in the non-occurrence of the joint action. But it is a failure of instrumental rationality for each co-agent to perform the same part-action such that the joint action fails to occur. The derivative reasons are provided in order for the co-agents to successfully act together; that is, the reasons must be distributed amongst the co-agents such that, in good cases, they successfully perform the joint action. So, how are derivative reasons distributed? First, only

[180] This is a theory developed by A.J. Julius, which he dubs 'reason giving reasons', see (Julius 2023). The theory has its roots in the Nagel's thoughts about instrumental rationality in *The Possibility of Altruism* (1970).

the agents picked out by the group reason will be provided with a derivative reason. Second, which derivative reasons get applied to whom will largely depend on each co-agent's circumstances. In the case where an elderly person falls in front of two people, which arm each of them should grab will depend on where they are standing relative to the fallen person: if one is on the right hand-side, she grabs the right hand, and vice versa. Another example: if paint needs to be bought from a store across town in order that two people to paint the house together and only one of them owns a car, then she is likely to be the one who ought to go buy the paint. The circumstances in which each co-agent finds herself distribute the derivative reasons amongst the co-agents just as circumstances in general distribute reasons amongst agents. This consideration, I believe, accounts for the majority of joint actions.

However, what if, all things being equal, the circumstances do not differ from co-agent to co-agent? How, then, do the derivative reasons get distributed? First, in some cases, it can be warranted that *any* co-agent as long as she acts before the others can act on any of the derivative reasons and other co-agents must choose from the leftover options. The cases to which I think that this solution applies are joint actions that take time. For instance, if we need paint in order to paint the house next week and one person goes and buys it before the others, then that is a perfectly rational thing to do. For the best results, that co-agent should also tell the others that the paint has been purchased so that they do not uselessly double up on paint. However, when the joint action needs to happen immediately, and each co-agent must act at roughly the same time, this solution seems to offer no help because many of them might choose the same part-action when there is no time to be informed of the best courses of action. It seems that each co-agent can act on the group reason and yet they never perform the joint action. In this second type of case, I think that the distinction between acting together and *successfully* acting together plays a role as well as the fact that the reasons involved are pro tanto reasons. The fact that derivative reasons are pro tanto reasons

helps to explain away the problem here. Pro tanto reasons are sufficient reasons that are enough to act upon without considering other reasons.[181] Nevertheless, that fact does not mean that there is a rationally better option available taking all of an agent's reasons together. Suppose each co-agent judges that they ought to act together, that they must all act now, and that two different part-actions must happen to successfully act together. If each derivative reason applies to the two co-agents, then they each have sufficient reason to act on each derivative reason. When they act on different reasons, they will act together successfully, as long as the world cooperates. If they act on the same derivative reason, this just means that they have failed to perform the joint action but not that they were not acting together. They were still trying to act together. There is no barrier to them acting together on the group reason just because it is indeterminate which part-action each one ought to perform. The trouble here is that it appears that co-agents can act perfectly rationally and yet not succeed in their joint action. This might suggest that group reasons do not sufficiently lead to joint action, that there is a gap between what is rationally required and the performance of the joint action. What good are group reasons if they cannot compel us to act together successfully? Consider an example. You are making burgers with a friend for dinner. While preparing the toppings, you both reach for tomatoes to slice them. Is this really an objection to them acting together? The tomatoes had to be sliced by any one of you. The fact that you both acted on the same derivative reason does not invalidate your joint agency. In fact, you have reason to decide who slices the tomatoes in order to complete your joint action. It is because you are acting on a group reason that you ought to try again in order to perform your joint action successfully. The appearance of a possible conflict in the rational course of the joint action does not eliminate the possibility of acting together successfully, but it does lower the likelihood of spontaneous success.

[181] I take McDowell to say a similar thing with regard to akrasia and the reasons an agent acts upon: an agent taking a reason to be enough to act on and doing so is enough to explain their action even if there is a form of irrationality in the action. See (McDowell 2010, 426–27)

There is one final concern with the rationality of doing one's part. I claimed in section 1 that group reasons warrant a co-agent to do her part without engaging with another. This claim seems true when co-agents are responding *appropriately* to their group reasons. However, in a two person case, where one agent will not perform her part of the action, it does not seem rational for the other to perform her part because they will not succeed in their joint action. Such failures may have terrible consequences, so it is important to figure out the rationality of the individual instrumental belief in bad cases. It seems as if the fact that a co-agent will not act on the group reason bears on whether I should perform my part, whether my individual instrumental belief has normative support. This problem is not an easy problem, and I do not have a special solution to it. Instead, I will follow A.J. Julius' solution since it is a cogent solution.[182] That solution states that there is an asymmetry of rational dependence in the good and bad cases. In the good case, where co-agents are appropriately responsive to their reasons, responding to the group reason suffices to explain why we act together, even spontaneously so. In the bad case, though, where one (or more) agent(s) simply will not do her part, then it is irrational for me to do my part of a joint action since I cannot satisfy the joint reason alone.[183] This does not mean that the group reason no longer governs us, because the failing agent must still count as failing and that is only intelligible if the group reason retains authority over us. In the bad case, what my co-agent will *not* do determines what is rational to do. For my part, if the group reason is sufficiently strong, it may still issue individual instrumental reasons to me even though my co-agent will not do her part. Instrumental reasons change with the circumstances. The failure of my co-agent is now part of the circumstances and, so, I might now be required to, for instance, persuade her to act on the group reason that we share. In the bad cases, I can still participate in joint agency though what I am instrumentally required to do may change.

---

[182] (Julius 2023, chap. 10).

[183] In some cases where an agent *can* do all the part actions, then she can perform all those actions as an individual rather than as a partner. For instance, she can paint the house alone if the other will not act.

Accepting this asymmetry seems to have an unacceptable consequence for my view. Don't the bad cases involve treating my co-agent as evidence? Don't the bad cases sully the possibility of spontaneous joint action? I do not think so. Recall that the challenge of spontaneity is to account for how agents are bound together without binding interaction. But on the reasons account, group reasons already bind agents together. And so, recognising the failure of a co-agent and treating their failure as evidence is downstream from occupying the co-agential perspective legitimately. The point is that the reasons account already meets the challenge of spontaneity before this problem arises. Now, the bad cases do seem to sully the success of spontaneous joint actions, but this is different from being spontaneously bound together. We can be spontaneously bound together and then engage in interaction, planning, discussion, argument, etc., that is, the kinds of things that are ruled out *as* binding interaction. But on the reasons account, they no longer count as binding interaction. All this means is that the 'interactive' joint action is not spontaneous all the way down to the action. And this should be a pleasing result since the extension of *joint action* should include spontaneous and non-spontaneous cases.

D. Each co-agent desires to A/B…Z. (Motivational Condition)

Finally, each co-agent desires to perform her part-action, under the guise of the group reason. This condition adds the motivational element to explain how each co-agent is moved to perform her part of the joint action. As a whole, each co-agent has a desire to do her part on the basis of a derivative individual instrumental reason, supported by the group reason, that binds each of them together in their joint action. This condition is necessary seeing as judgment alone is insufficient to explain action: judgments help to specify what it is an agent is doing, but they do not explain *that* she is doing it intentionally.[184] But why does a co-agent only desire to do their part rather than to perform the joint action? For instance, I only desire to buy the paint and paint part of the house, not the

[184] For some dissent to this claim, see (McDowell 1979; Dancy 1993).

entire joint action. Why is this? There is a well-accepted constraint on desire that says that it must be possible for an agent to bring about her end for that motivational state to count as a desire, as opposed to a wish or some other conative attitude.[185] For instance, it makes no sense to say that I desire a drink in my hand when I already have one in my hand. This constraint helps clarify the motivational attitude involved in joint action because it is not possible for one co-agent to perform a joint action on her own, so she cannot strictly speaking *desire* the joint action. Imagine that I desire for us to go for a walk together. No matter what I do on my own, I cannot perform this joint action alone—the joint action requires more than one person to do it. Even if I walk alone on the basis of a group reason, it is not true that I am walking with you because no one else is acting on the group reason. It is better to characterise the attitude as a wish for a joint action but wishes do not help rationalise actions. Even in cases where it is possible that one person, with adequate time and effort, can perform all of the part-actions of a joint action on her own, she cannot satisfy a desire for joint action. It is possible that one person can paint her house, buy the paint, and rent the ladder without the help of another, but this does not make it true that she painted a house together with another. For a joint action to happen at least two people must act together and, for this reason, no individual can have a desire for a joint action. And, thus, the desires co-agents require are desires for their part-actions, which are instrumentally supported by their group reason.[186]

But what do desires have to do with our reasons, or our judgments about reasons? Is it not possible, perhaps even likely, that our desires are not constrained by our reasons? After all, aren't we just slaves of the passions and our beliefs about reasons are motivationally inert? This is the toy Humean picture of desires and their relation with beliefs. However, the theory of desire that I think

[185] See (Velleman 1992). This is analogous to the own action condition on intention discussed in section 1.
[186] This seems to reverse the normal causal story of action where an agent has a desire for some end and, in virtue of an instrumental belief, she desires and performs the means to that end. The reason for this reversal is because our ends are given by reasons to which our desires are sensitive.

is most justifiable is one where desires are *sensitive* to our judgments about reasons. I cannot get into all the details of why I think that this theory is best, but the basic idea is that in order to make sense of the nuances of desire, a theory of desire must incorporate a connection with an agent's judgments about her reasons, that is, her normative beliefs.[187] On the version which I support, normative judgments condition desires. For example, when an agent has a desire to win at chess that is because she judges that there is reason to win at chess. This agent will satisfy her desire and do what she believes that she has reason to do just in case she wins, that is, wins in any circumstances, e.g., against poor opponents, children, by cheating. However, if the agent also thinks that there is reason to win fairly and play suitable opponents, then the ways in which she can satisfy her desire to win at chess become more limited.[188] If she wins by cheating, then she has acted against her reasons and it appears to withhold the value of winning. And if the value disappears, it seems as if the desire was never satisfied. Thus, to account for all the nuances that each desire contains, the contents of desire must be conditioned by the agent's judgments about her reasons.

With this theory of desire in view, the reasons a co-agent has to perform her part-action condition her desire and explain why she acts as she does. Each co-agent, insofar as she acts on her reasons, has a desire conditioned by her individual instrumental reason, provided by the group reason, to do her part. Because she judges that she has both reasons, it would be irrational for her to attempt to satisfy her desire for the part-action in way that would be detrimental to the joint action, under suitable conditions of knowledge. This is the way that the group reason still constrains what co-agents do in service of the joint action. Even though they do not act *directly* on the group reason, it still has a rational bearing on what they do. The constraint provided by the group reason in part explains why co-agents who derivatively desire to do their own part are indeed acting

[187] For these arguments, see (Scanlon 1998, chap. 1; Myers 2012; 2017; Myers and Verheggen 2016, chap. 6)
[188] This is a consequence of holism of the mental, see (Myers 2012).

together: they are all intentionally acting to bring about the joint action. The group reason that they share normatively structures the entire enterprise by binding them together, providing the conditions of their collective agential success, and providing reasons that help rationalise each co-agent's behaviour.

3. Two Objections and a Concern

This section discusses how the reasons account of joint action avoids two leftover problems from which other accounts of joint action suffer along with a peculiar concern for the reasons account. I have already spoken to the requirement of an epistemic constraint, the own-action condition, and the settle condition in section 2. Now, I explain, in particular, how the reasons account avoids the circularity objection and the problem of parsimony. The first concerns whether the reasons account of joint action is viciously circular and, hence, sheds no light on the phenomenon of joint agency. The second concerns whether it is objectionable to bloat our ontology with group reasons (assuming that these are ontological entities). And the peculiar concern has to do with the nature of long-term group reasons and spontaneity.

First, circularity is a problem that plagues the joint action literature. The objection goes like this: if the goal is to explain how co-agents are bound together when they act together, then joint concepts should not appear in the explanatory resources of the account. For instance, suppose that co-agents intended to walk *together*. The jointness is already assumed in these intentions rather than explaining how co-agents constitute a group when acting on the right kind of intentions. The real question is how assuming joint content can help shed light on the nature of joint action when the task of a theory of joint action is to explain the *joint* phenomenon.[189] Is the reasons account of joint action susceptible to the circularity objection?

[189] For more on this objection, see (Petersson 2007; Bratman 1999a, chaps. 5 and 8; Kutz 2000a, 86–88; Schweikard and Schmid 2013).

Though the reasons account depends on group reasons, which is a joint resource, I do not think that the account is viciously circular. On the reasons account, joint action is not explained merely in virtue of the relevant attitudes with joint content. The explanation depends on the group reason to which those attitudes are responding. Co-agents are bound together by the group reason and they act together when they respond to it appropriately. Although their attitudes contain joint content, they are endowed with joint content in virtue of the group reason. Their joint activity is not merely explained by attitudes with joint content. It is explained by their attitudes which stand in the appropriate relation to and structured by a group reason. I consider this an externalist account of joint action because it begins from outside the co-agents' minds. In Chapter 1, I suggested that we look beyond the co-agents' attitudes to explain spontaneous joint action. The group reason is that which is beyond their attitudes, and it does the binding that brings co-agents together. It is the basis for their capacity to act together spontaneously. The fact that group reasons are distinct from co-agents' attitudes is important. Indeed, normative reasons in general are independent from how agents respond to them. So, when it comes to circularity, group reasons alone do not explain joint action; the other attitudinal conditions are required. What this means is that the *joint* element in the explanation of joint action is not sufficient to explain a joint action. The fact that there are other necessary conditions in the explanation of joint action indicates that the reasons account is not viciously circular. Given that the challenge of spontaneity demands that we provide a non-revisionary answer, it cannot be helped that there is some irreducibly joint element in the account. The benefit of the reasons account of joint action is that the joint element is insufficient on its own to explain joint action. Because of this, circularity should not worry the reasons account.

Second, the parsimony objection arises when we introduce some joint concept into theorising about joint action. It might be thought that the introduction of group reasons

unnecessarily bloats our ontology.[190] It is worth pausing to consider whether group reasons are ontologically problematic. I do not think that they are. I will consider two different parsimony concerns. First, perhaps the methodological individualist denies the existence of group reasons because she denies the existence of any joint phenomenon. However, the methodological individualist has been shown, at least according to extant accounts, to fail to meet the challenge of spontaneity.[191] When a methodological choice forces us to do injustice to the phenomena that we aim to explain, then the methodological choice is the problem. Even though the methodological individualist will find it objectionable, group reasons are necessary to meet the challenge of spontaneity, so it is not an unnecessary inflation of our ontology. Second, the introduction of specifically *normative* group reasons might be problematic. However, I do not see a strong reason why we should deny the existence of normative group reasons since normative individual reasons are unobjectionable.[192] Once methodological individualism is off the table, why is it so problematic for there to be normative reasons for joint action? Group reasons do not in fact introduce a new kind of thing into the world, such as normativity in a non-normative world. Rather, they expand a category that is already generally accepted. It would be a very odd situation if we admit individual normative reasons into our ontology, while they play a special role in our thinking about individual action, and, at the same time, hold that there are joint actions but there are no normative reasons for joint action. Ultimately, parsimony is not a telling objection to the reasons account because we require group reasons to explain spontaneous joint action.

It is worth thinking on a final concern with the reasons account and spontaneity. Thinking about the Algerian example, why do the Algerians act together spontaneously *when* they do? What

[190] The reasons account is also committed to the collective 'we'. But this is a by-product of the co-agential perspective, which was justified in Chapter 1.
[191] See Chapter 2, section 3.
[192] Of course, some will deny the *existence* of normative reasons, such as expressivists.

makes it rational for them to perform their joint action at this moment when presumably that reason existed during the whole period of French occupation? I do not have an answer to this question, but I think it is a hard question to answer. If it weren't for spontaneity, the same worry wouldn't arise. If agreement between agents or the like were permitted, then the Algerians would act at the agreed time. What makes the problem difficult is that it seems rational for the Algerians to protest *at any time* during the occupation. So why now? It might not seem appropriate to ask this question about an action theory, which the reasons account is, because it is tasked simply with explaining an action rather than explaining why that action is better than or more justified than others. Why the Algerians are acting on *this* reason rather than another at this moment seems to be an unfair question. But because the reasons account takes normativity to be an essential aspect of the account, it is burdened with answering why it would be rational to act on this reason rather than another because that is how we engage with reasons. Again, I want to emphasise how difficult the question is. Why, without interaction, would the Algerians pour into the streets and protest on that fateful morning? They could have done it the day before or even the day immediately after occupation. Why do they all respond to the group reason *now*? One stab at an answer is that when a reason lingers for so long and it is not acted upon, it can become oppressive. It can fill the air and one cannot see or think of anything else. I take it that the group reason in the Algerian example is like this. It did not exactly silence other reasons, but it took up everyone's attention because it could not be ignored any longer. Of course, this is no answer to the problem because an explanation of how some reasons become oppressive is needed.

4. Conclusion

The task of explaining spontaneous joint actions required of an account that could meet the challenge of spontaneity. The reasons account of joint action resolves the challenge and, importantly, without revising any of our natural conceptions. First, it explains how co-agents are

acting together as opposed to acting in parallel because the co-agents are responding to a group reason that they share. The group reason binds them together as the agents who ought to perform the joint action and, when they respond to the reason accordingly, they are in fact bound together in action. Second, it is in virtue of group reasons that the account meets the spontaneity requirement as well. When each co-agent recognises her group reason, there is no need for the co-agents to engage in binding interaction in order to act together. As long as they each think the joint action on the basis of the group reason, thereby occupying the co-agential perspective, and each derive her part-action from it, then they can do their parts intentionally under the description of the joint action. The question that is left for the next chapter is: what is the nature of group reasons such that they can help to explain spontaneous joint actions? This chapter assumed that group reasons were immediately available for co-agents, but for that to be true group reasons must have a peculiar nature. They must be real reasons that are mind-independent and normatively public. The following chapter makes this case.

CHAPTER 4

Real Group Reasons

The account of joint action defended in the previous chapter depended on normative reasons for joint action, or as I call them for simplicity's sake, group reasons. By appealing to *normative* reasons, as opposed to *motivating* ones, the reasons account of joint agency is able to respond to the challenge of spontaneity. Agents could legitimately occupy the co-agential perspective together by responding to their group reasons. A group reason gives each co-agent the warrant to think that the other is acting with her. Moreover, the togetherness requirement is satisfied because the group reason unifies their agency so long as they both act from it. But more needs to be said about the nature of group reasons in order to show that the spontaneity requirement is met. The previous chapter left the nature of normative group reasons unexamined; instead, it merely defined them. And there are many theories of normative reasons, some of which may violate the spontaneity requirement. For instance, take a very crude internal reasons position: if a theory said that normative reasons are based, in some way, on an *individual's* motivations, then the scope of the theory is too narrow to account for group reasons that co-agents share. In this case, normative reasons do not extend beyond the individual. Rather, two individuals may have overlapping singular reasons, but those do not explain joint action in the appropriate way. In order for the reasons account of joint agency to meet the challenge of spontaneity, the group reasons co-agents act on must not, in some way, violate the spontaneity requirement. The purpose of this chapter is to determine which theory of normative reasons meets the spontaneity requirement. I argue that realism about reasons provides the best account of group reasons, under these constraints, because group reasons would be mind-independent and, in some sense, public. These two features are essential to providing an adequate account of spontaneous joint action because agents would not need to engage in binding interaction so long as group reasons possess them.

This chapter will canvas some prominent views of normative reasons, making adjustments as I proceed to render them amenable to group reasons. I will discuss both relativistic and non-relativistic internal reasons, represented by Bernard Williams and Michael Smith respectively, and realism about reasons. I argue that realism about reasons is the only theory considered that can do justice to group reasons. I begin by showing that motivating reasons are insufficient for a theory of group reasons; this will largely collect points already made previously throughout the dissertation. Then, I argue that internalism fails to provide an adequate account of group reasons because it fails to explain how agents could be in an epistemic position to act on their group reasons—a fundamental commitment of internalism—without binding interaction. Next, I explain in more detail what normative group reasons are with an emphasis on rejecting the claim that normative reasons apply to *an* agent alone as opposed to a plurality of co-agents. Following this, I show how realism about reasons can yield group reasons that meet the spontaneity requirement. Realism about reasons provides an adequate account of group reasons.

1. Group Reasons are not Motivating Reasons

It is a natural thought that the reasons involved in joint action are motivating reasons. After all, motivating reasons are typically the kind of reasons appealed to in explanations of individual action. Motivating reasons explain and rationalise an agent's intentional action, often by appeal to certain attitudes of the agent. For instance, Jane is driving to 150 Bloor St. because she desires to go to the restaurant and believes that driving will get her there.[193] Motivating reasons are distinct from normative reasons in two ways: first, the content of the rationalising attitudes may be false whereas normative reasons are always true and, second, normative reasons do not explain what an agent is doing but rather what an agent *ought* to do. Moreover, motivating reasons can be distinguished from

[193] Without digressing into substantial argument, there are different accounts of what counts as a motivating reason. For examples see (Anscombe 2000; Bratman 1987; Smith 1987; Thompson 2008; Alvarez 2013; 2018).

explanatory reasons because explanatory reasons may have nothing to do with the *agent*. Occasionally, this distinction is overlooked and motivating reasons are referred to as explanatory reasons because the job of motivating reasons is to explain actions. It is prudent, though, to keep the distinction in mind because a person might do something because of an explanatory reason which does not render her action intentional. For example, a person might swing her arm in the air, flip the light switch, and thus turn on the light because she was startled by a shadow. Another way to express these differences is to say that motivating reasons provide a rationalisation of action rather than a justification (normative reasons) or a mere explanation (explanatory reasons). Motivating reasons do not justify the action undertaken by the agent, but they tell us what the agent is doing intentionally and why she does it.[194]

Since a theory of joint action should explain what co-agents do together intentionally and why they do it, can motivating reasons alone do the job? The answer is 'no'. Motivating reasons accounts will fail because either they cannot meet the togetherness requirement or, by meeting it, they fail to meet the spontaneity requirement; that is, they fail to live up to the challenge of spontaneity. Imagine that two people intend that they see a painting at the gallery and they each go to see the painting at the same time.[195] One person formed her intention by spying the other's calendar and learning that she was going to the gallery. Whereas the other person, having the intention to see the painting with the spy, follows her from her house to the gallery. Both of their intentions seem to be satisfied since they are together in the room viewing the painting. And their intentions explain, as motivating reasons, what each of them is doing and why she is doing it. The question is whether they count, on the basis of these intentions, as viewing the painting together in

[194] One might think that motivating reasons can justify in an internal sense because they are representative of the agent's best judgments. This would require denying the possibility of akrasia, something many action theories take as an explanatory goal.

[195] This example assumes, following Bratman that an agent can intend for others to act. See (Bratman 1999a, chap. 8).

the joint sense. The intuitive answer is no because the fact that their intentions matched and they both viewed the painting at the same time was an accident, that is, their intentions are not properly bound together. Neither agent has any reason to think that *the other* is thinking that they are going to view the painting together—they do not legitimately occupy the co-agential perspective. Even though the spy knows that the other plans to go to the gallery, she has no reason to think that the other will be there to view the painting *with* her. Likewise, even though the follower is following the spy to the museum, she also has no reason to think that the spy is there to view the painting *with* her. Because of they each lack relevant information, the agents are not acting together. Having attitudes with 'collective' contents, which constitute motivating reasons, is insufficient to meet the togetherness requirement. Both agents are still merely performing individual actions.[196] Furthermore, it might be wondered why, in a given situation, I cannot simply trust that another has the right motivation that would lead to our joint action. On the basis of what, though, can I place my trust in the other? Either she has provided me with evidence, which would violate the spontaneity requirement, or my trust is blind and I am in the same position as the spy at the museum. Merely trusting that others have the right motivations will either fail to meet the togetherness or spontaneity requirements.

But maybe further attitudes can be added to secure the right kind of connection between the two people's intentions. The problem in the above case can be solved by adding attitudes that constitute common knowledge. Above, both agents individually knew that they were going to view the painting 'together' but they did not mutually know or have common knowledge that they were doing so together. Many prominent theories adopt a common knowledge condition(s) to satisfy the

[196] Bratman presents a similar argument against having intentions with the same contents. There he argues that each agent has to view the other as contributing to *their* joint action (Bratman 2014, 43–44).

togetherness requirement.[197] According to these theories, a joint action occurs, very roughly, when agents have matching intentions and they somehow have common knowledge of those intentions and act on that basis. Let's grant that this meets the togetherness requirement: co-agents now view each other from the co-agential perspective. But this addition of common knowledge does not help because it violates the spontaneity requirement. How can co-agents possess common knowledge of each other's attitudes, such that they can meet the togetherness requirement, without interacting with each other? Immediately, there appears to be an impasse. For instance, I cannot *know* that my co-agent knows that I intend to view the painting with her and vice versa unless we have sufficient evidence for drawing that conclusion. This evidence will typically come by way of assent, agreement, or perhaps merely an expression of commitment. Each of these means of knowledge requires that co-agents interact with each other in some way, they serve as evidence for each other. Perhaps, in the above example, the spy knows that the other intends to view the painting with her because it is written in the calendar; and the other knows that the spy intends to view the painting with her because she found the spy's diary. But even if each of them knows that they intend to view the painting together, they do not each know that the other knows about her intention. Common knowledge goes beyond merely knowing another's intention because it requires second-order attitudes on the part of the co-agents about each knowing that the other knows about her own intentions. So, even if each co-agent has evidence of the other's intentions, more is required to have common knowledge of each other's intentions. This additional requirement will involve some form of interaction between the co-agents, and even then the right form of interaction is difficult to achieve.[198] The issue here is that any interaction that reveals co-agents intentions to act together

[197] (Bratman 2014; Gilbert 1989; Tuomela and Miller 1988; Chant and Ernst 2008). Common knowledge has been attacked recently (Blomberg 2016b; Schönherr 2019), so it is not universally recognised as a solution. Those who attack it, though, often have a different conception of what is required for meeting the togetherness requirement.
[198] The coordinated attack problem (Chant and Ernst 2008, 552–56) and the email game (Binmore and Samuelson 2001) make it clear that the bar for common knowledge is quite high, if not occasionally unachievable.

violates the spontaneity requirement. Indeed, even the appeal to diaries and externalised expressions of intention still counts as binding interaction.[199] So, in an attempt to satisfy the togetherness requirement by adding a common knowledge condition on joint action, such theories fail to meet the spontaneity requirement because common knowledge can only be secured by some form of binding interaction.

If the reasons for which we act together are motivating reasons, that is, merely a structured set of attitudes that are relevantly connected to the action, then rationalisation of joint action that meets the challenge of spontaneity is impossible. For this reason, we need to look outside of the attitudes of the co-agents alone to rationalise joint action. We must look towards normative reasons.

2. Internalist Versions of Group Reasons

Internalism about normative reasons was initially developed by Bernard Williams, who offered a relativistic version of the view.[200] According to his view, normative reasons must be appropriately related to an agent's motivations, what he calls her "subjective motivational set".[201] All of an agent's normative reasons are relative to the items contained in her subjective motivational set—e.g., desires, personal loyalties, or, more generally, "commitments of the agent"—because they are conclusions to which she would be lead if she deliberated soundly from those items.[202] So, an agent cannot have a normative reason if she would not retain or develop, under rational conditions, the motivation to perform the action. This is a relativistic position because people's normative reasons would not converge if they had sufficiently different subjective motivational sets.[203] The purpose of

[199] I argued this when I raised the issue of Tuomela's Bulletin Board View (Chapter 1, footnote 43).
[200] I have chosen Williams because his position is canonical. Sharon Street has a very similar position to him, except for the fact that only the attitude of valuings matter to the construction of normative reasons (Street 2012) Street is also more of a proceduralist about deliberation.
[201] (Williams 1981, 102).
[202] (Williams 1981, 105).
[203] This claim may not accurately represent Williams' real intentions given his argument against external reasons (Williams 1981, 108–9). There, he argues that any external reason must be psychologically connected to the agent in order for it to play an explanatory role, so internal reasons get a foothold once again by the psychological necessity.

relativising reasons to an agent's motivations is to secure an explanatory link between her reasons and her actions—when an agent acts on her reasons, they should explain why she acted as she did. This 'practicality' constraint on normative reasons is important, but it is not meant to make a normative reason of any motivation. Taking explanation to be paramount to an account of normative reasons misinterprets their proper role, that is, to account for the rationality of an agent. Normative reasons must be able to *justify* actions from the first-personal and third-personal perspectives as opposed to merely explaining why an agent did what she did.[204] They must have rational authority over other considerations for the agent and anyone can see that they do if they had access to all relevant information. In light of these two perspectives, an agent can be wrong about her normative reasons. For instance, an agent has a desire to rob a house because she believes it to contain valuable jewellery and, as a result, she believes that she has normative reason to do so. However, the house contains no jewellery and the desire is based on a falsehood. Third-personally, she does not have reason to rob the house. She also does not have a reason from her perspective even if she thinks that she does so. Specifically, a motivation based on a falsehood or an agent's incorrect belief about how to satisfy a desire cannot yield a normative reason. Were the agent to know all the facts, she would not be justified in pursuing an action grounded in falsity.

But rationality involves more than mere truth and falsity because there may be conflicts within the subjective motivational set. Internal reasons are those that can be arrived at via a "sound deliberative route".[205] This does not mean that normative reasons depend upon *actual* deliberation because an agent might deliberate badly or not at all and, all the while, still have normative reasons.

Nevertheless, his discussion of Owen Wingrave and the abusive husband suggest a strong relativism. I think that there is a fundamental tension in Williams between the explanatory and justificatory roles of normative reasons.

[204] The third-personal perspective here has to be considered from within the internal reasons position. Williams might say that an abusive husband has no reason to be kind to his wife, third-personally, whereas an externalist would say that he does have reason. The third-personal ascription of an internal reason depends on the facts of an agent's subjective motivational set; an externalist would not limit reasons to an agent's motivations.

[205] (Williams 1995, 35).

Normative reasons, rather, are facts about an agent's subjective motivational set with respect to the constraints of hypothetical sound deliberation. Beginning from an agent's motivations, were she to deliberate well from her motivations, she would come to conclusions about what she has reason to do. For example, if I only desire to be caffeinated and there is coffee in front of me, then, I not only have reason to be caffeinated but, sound deliberation would determine that I have reason to drink the coffee. Importantly, an agent can have many reasons at one time because sound deliberation produces many consistent or equally strong reasons. Depending on my motivational set and the conclusion of sound deliberation, I might have reason to drink coffee and, at the same time, have reason to abstain from coffee and drink water—deliberation may not help me to settle on which reason to act. Nevertheless, deliberation is quite an expansive capacity for Williams, ranging from time-ordering of ends to imagination. The role of sound deliberation, were it to occur, is to subtract from and add motivations to the subjective motivational set. For instance, Jane has motivations A, B, C, and D, where A-C are based in truth and D based on a falsehood. In this case, sound deliberation would bring the agent to realise the falsehood and to see that she has no reason to act on motivation D. It is constitutive of *sound* deliberation that it eliminates motivations based on falsehoods and depends on truth for its progress. If Jane actually deliberates in this way, then "the deliberative process can…subtract elements from S".[206] Further, motivations can be added to an agent's subjective motivational set by a variety of deliberative processes. As above, the realisation that I have reason to drink coffee can add the desire to drink coffee to my motivations.[207] But also the exercise of imagination provides quite a lot of leeway in which desires get added and removed. For instance, if I cannot imagine how bad being a serious heroin addict is, then my desire to try

[206] (Williams 1981, 104).
[207] Williams is rather murky here, but he seems to think that our motivations are causally related to deliberation. Although he provides no argument for how reasoning and motivations interact, he might think that deliberation is already infused with motivations because we deliberate from desires and our reasons have an essential connection to them.

heroin may survive deliberation or if I can imagine how beneficial a just and equal socialist society is, then deliberation may extinguish any desires for a libertarian society that I have. Imagination does not only concern what a state of affairs is like, it can actually make deliberative leaps. For instance, if I desire to avoid traffic, I might imaginatively leap to the belief that I have reason to find the nearest helicopter and form that desire even though this may not be my most efficient option. Even with this leeway, imagination must still be hemmed in by truth in order to constitute a part of sound deliberation. In the end, Williams considers deliberation to be a rather open-ended capacity in virtue of imagination's role as opposed to the more proceduralist versions of internalism.[208] Moreover, this process of adding and subtracting motivations just goes to show that an agent's subjective motivational set develops as she deliberates soundly. Though, if sound deliberation is only hypothetical, the development of her subjective motivational set ought to occur.

Now, what would group reasons look like on this theory? Each co-agent picked out by a group reason has a subjective motivational set, but an 'internal' group reason cannot be derived from one individual's motivations and still apply to the others. In order to play an explanatory role, there must be a connection between each co-agents' motivations and their group reason. If one co-agent lacks the grounding, or a suitably related, motivation, then that co-agent cannot be said to have a reason to act together. This is the core of Williams' relativist position. To be clear, this means that two agents who arrive at overlapping reasons, which may refer indexically refer to each other, e.g., reason to act with another, do not arrive at proper internal group reasons because each is not normatively constrained by the *same* group reason. Recall that the primary function of group reasons is to bind agents together and reasons that merely overlap in reference but not in normative force do not bind agent together.[209] So, group reasons must be relative to co-agents' *collective*

[208] For instance, (Smith 1994; 1995; Street 2010; 2012).
[209] Overlapping reasons would certainly justify cooperative parallel action. They may bring about an event while making use of each other. But this is not joint action.

subjective motivational set; that is, internal group reasons are relative to a single expanded motivational set comprised of the motivational sets of the co-agents.[210] This does not mean that co-agents must share the same starting motivations; instead, somewhere, their motivations must converge. For instance, one individual has a desire to dine with another and the other has a desire to dine and a separate desire to spend time with another. For the latter person, there is a clear deliberative route to the desire to dine with another. In such a case, their subjective motivational sets could collectively provide a group reason to dine together. Interestingly, it appears that if two people's motivations were never to converge, they would never have a group reason to act together. This limits the power of the theory to plausibly explain the amount of group reasons that agents presumably share, but this appears to be part and parcel of the relativistic nature of the theory. Next, all motivations rooted in falsehoods are eliminated as sources of group reasons and the co-agents could be wrong about their reasons. The most important feature of Williams' account though would be that there is a sound deliberative route from the collective motivational set to group reasons. Assuming co-agents are fully rational and aware of each other's motivations, they could deliberate towards group reasons to act together. Now, it is not necessary that they in fact do deliberate on their collective motivations; there just is a fact that, from their motivations, there is a route to a group reason to act together and this would be discovered via sound collective deliberation. One might think from this perspective that another's matching motivation simply cannot matter when deliberating about what to do because it has no connection to the agent's own subjective motivational set. That is, why would matching motivations produce a reason for *us* rather than two individual reasons? It is hard to know what Williams would say to this but were another's

[210] I want to note a methodological concern here regarding the pooling of subjective motivational sets and whether relativistic internalism permits the inclusion of another's motivations in the deliberative process. Given that an agent's reasons are relative to her subjective motivational set, it may be illegitimate to include another's set in service of deliberation towards group reasons. After all, the argument that I am making is that internal group reasons cannot get a grip on agents unless they interact in ways that flout the spontaneity requirement. This would be problematic because the whole motivation for internalism is to show how reasons have a grip on agents.

motivations to be excluded, then relativistic internalism could not yield group reasons. Nevertheless, in order to be charitable to the theory, I will assume the pooling of motivations is permissible. Excusing this concern, there does not seem to be anything troubling about deriving a group reason from two or more individuals' collective motivational set.

The question now is whether relativistic internal group reasons are consistent with the requirements on a theory of joint action. My claim is that relativistic internal group reasons conflict with the spontaneity requirement because for agents to be in an epistemic position to act together spontaneously on their group reasons requires that they would interact with their co-agents beforehand. The spontaneity requirement, again, says that co-agents need not engage with each other in order to have a co-agential perspective on each other and their joint action from which they act together. How will co-agents be warranted in adopting this perspective, on Williams' account? If co-agents do not interact, then in fact they only have their own motivations from which to deliberate. And I do not see any way for there to be an actual deliberative route from one's own motivations to a reason for one and another to perform a *joint* action. Even if your motivations include intentions that another φs with you, the reason would have no basis in the other's motivations and would still not be a reason for her to act together with you.[211] Such other-regarding intentions would only ground an individual reason to cause another do something. The only way for you to arrive at a group reason via sound deliberation would be to have access to your co-agent's motivations and this cannot be done without binding interaction.[212] Even though there might be a fact of the matter via hypothetical sound deliberation that you and another have an internal group

[211] These other-regarding intentions are defended by Bratman. A useful example is that a mother can intend that her child clean his room. Even if the mother has reason to make her child clean her room, it does not mean that the reason extends or transmits to the child for internalism.

[212] Imagine a co-agent who covertly reads another's MRI report and knows what the agent is motivated to do. Those motivations, though, on Williams' account are inadmissible because not even the agent can deliberate from those hidden or unconscious motivations. Moreover, this counts as binding interaction because this evidence and deliberation is part of the binding process.

reason to act together, for deliberation to do its necessary work the co-agents must pool their motivations together such that they are available to each other and deliberate from them. The problem is not whether Williams' theory can provide an account of group reasons, rather it is whether it can provide a suitable account that can accommodate spontaneity. Remember that it is necessary that reasons can explain action for Williams, and if the recognition of a group reason necessarily depends on binding interaction, then it cannot explain spontaneous joint action.[213]

One might object that an agent can assume truly that another has a particular motivation or relevantly related motivations and, coupled with her own, deliberate from there. Why can't co-agents arrive at group reasons on the basis of true assumptions without interacting? Clearly, deliberation often involves making substantive assumptions. Suppose two people assumed truly that they each desired to live decent lives. Why can't they deliberate on their own to the conclusion that they have group reason to structure their society fairly? I worry that assumed motivations cannot *soundly* lead to conclusions about group reasons. When an end or desire of another is assumed, all that can result is a further assumption containing a reasons statement. This is not the kind of conclusion that represents normative reasons—it is not actually a reasons statement. An assumed conclusion about reasons does not justify acting on the reason contained in it.[214] Recall that normative reasons primarily have a justificatory role. A further logical step is required for an agent to determine that they *in fact* have a reason, that is, the assumption must be discharged by verifying that her co-agent has the relevant end or motivations which would lead to the relevant end. In the individual case, an agent merely needs to acknowledge that she has a desire in order to discharge her assumption. The sound deliberation of an individual would not conclude in an assumed conclusion about reasons since the resources are internally available to the individual to make the further logical

[213] I think Williams himself would have this concern if he accepted my requirements on joint action.
[214] By 'assumed conclusion', I mean the conclusion of a sub-argument beginning with an assumption.

step. But, in the group case, it requires binding interaction because an individual cannot settle on her own whether her co-agent has the discharging motivation. An agent who truly assumes her co-agent's motivation is only warranted to derive an assumed conclusion about their group reasons. The (simplified) reasoning looks like this: (1) I desire joint action X, (2) if B and I each desire to X, then we have group reason to X, (3) I assume that B desires X, (4) so, I assume that we have group reason to X. This conclusion about a group reason does not *warrant* adopting the co-agential perspective because an assumed conclusion about a reason cannot warrant acting on that reason until the assumption is discharged. The conclusion has no justificatory force unless the assumption is discharged, that is, proven true. So, a co-agent would have to confirm the other's motivations in order to inferentially move beyond the assumed conclusion to a proper normative reasons statement. In the group case, each co-agent does not have access to all the relevant motivations needed to discharge the assumption on her own, as one does in the individual case. To do this requires binding interaction and the flouting of the spontaneity requirement.[215]

One might still object, though, that this argument does not succeed because it seems to depend on actual reasoning in the group case instead of hypothetical reasoning. It is a fact that co-agents begin with some motivations, and Williams' theory is only committed to there being a sound deliberative route from their starting motivations to group reasons if they were to soundly deliberate. Now, the question is whether hypothetical sound deliberation must respect the limitations incurred by the spontaneity requirement. Why can't hypothetical sound deliberation have access to all the relevant motivations? Let's assume that it does. Clearly, there would be truths about what the co-agents would have reason to do together. And if there are truths about what co-agents have reason to do together and if co-agents act on them, then Williams' internal reasons account succeeds at

[215] Why can't Williams rely on trust? The trouble is that there is no *reason* to trust the co-agent yet, that reason is the result of deliberation.

meeting the spontaneity requirement, even if it only explains a small number of cases. Due to the argument in the preceding paragraphs, co-agents cannot *actually* deliberate soundly on their own to their group reasons without flouting the spontaneity requirement. But it is still possible that they choose their group reasons 'out of the air', so to say, without *actual* deliberation and perform joint actions spontaneously. There is still a sound deliberative route to their reasons from their collective subjective motivational set and they have not interacted with each other. It appears, then, that they can act together spontaneously in such cases. Despite their luck at choosing the same group reason, there is still an impediment to this explanation based on the constraints of relativistic internalism. The impediment is that, in the actual case, each co-agent, *on her own*, does not have a sound deliberative route to the group reason. Considering the motivations that each agent has available to her, she will not arrive at a group reason, at most she can arrive at an individual reason that matches the other's reason. They must interact in order to see that their collective subjective motivational set leads to a group reason. The question here is: what is the point of the appeal to deliberation if it is not something that the co-agents can rely on to reveal their reasons? Pumping motivations through the normative standards which constitute hypothetical sound deliberation is what transforms motivations into normative reasons. But if co-agents can never deliberative soundly to their group reasons, why would they on their own recognise the *normativity* of the group reason when it comes time to act? That is, why would they see their group reasons as justifying that course of action? The group reason, by their own lights, would appear *ad hoc* and they would never be able to justify to themselves why they are rationally compelled to act on it and also view the other as a co-agent. Even if the group reason has the motivational connection, Williams needs to explain why, to the agent, it plays the crucial justificatory role of normative reasons. Indeed, group reasons would appear to the agent in these circumstances as, by Williams' lights, those problematic external reasons. Thus, for relativistic internal group reasons to get a grip on the agent and for her to be

warranted in adopting a co-agential perspective requires engaging in binding interaction. Relativistic internal group reasons cannot do the job that they are intended to perform, that is, explaining spontaneous joint actions.

Things are even worse for Williams' account if the role of imagination is more pronounced, as he tends to do later on.[216] Because imagination is constitutive of deliberation, Williams says that it "is a consequence of the account that the question whether a given person has a reason to act in a certain way may have no entirely determinate answer" and "there is an essential indeterminacy" when it comes to an agent's reasons.[217] Each agent will have different imaginative capacities, meaning that what her reasons are in any given situation depends on her ability, while respecting the other conditions of sound deliberation, to imaginatively connect her current desires to further desires. In the group context, even if rational co-agents knew everything about each other's starting motivations, they would need to deliberate in the same way to arrive at the same group reason. But it seems that imagination gets in the way of accurately predicting a co-agent's deliberative moves. For instance, one co-agent who has minimal imagination sticks to deductive logic and the other who is wildly creative prefers 'inference by imagination', where she ends up it is hard to know in advance. That these two people arrive together at a conclusion about their group reasons seems highly unlikely. Were they to converge at some point, that alone does not imply that they have reason on Williams' account. Deliberation does not stop because an agent likes where she has arrived or because she converges with others. Deliberation stops at its rational and imaginative end, and if there are further deliberative steps to take, then they ought to be taken. If one person can make further steps than the other, then they will have different reasons in any given circumstances. It is possible that the more imaginative person could help the other see the conclusion at which she

[216] (Williams 2001).
[217] (Williams 2001, 92).

arrives and then they could converge on their group reasons, but this flatly contravenes the requirement of spontaneity. If a sound deliberative route contains imaginative leaps on the part of the co-agents, it is unclear how any one of them could be in a position to accurately foresee the group reasons that they have without interacting. Williams' relativism extends beyond agents having different motivations, even their deliberative capacities are different. This is a major hurdle to overcome if relativistic internal group reasons are to explain spontaneous joint actions. Taking imagination seriously for Williams' account makes it difficult to see how people would have group reasons, except in rare circumstances, but even if the less radical version is the official view, then it still has to overcome the epistemic gap between agents in order to get sound deliberation going.

The problem with Williams' account appears to be its commitment to relativism. After all, none of us know at a glance what others are thinking or how their imagination works. How can we know the other's motivations (and deliberative capacities) without interacting with them? But what if deliberation had the power to make us all converge on the same desires and, hence, reasons? This is what Michael Smith has argued for, namely, that internalism about reasons is best understood as a non-relativistic theory.[218] In this case, even though deliberation begins with agents' desiderative profiles or subjective motivational sets, like Williams, we can be confident in the power of reason to arrive at the same conclusions about reasons for all of us. In fact, according to Smith, giving an agent's desiderative profile significant weight in the determination of reasons is a mistake because our reasons would then appear arbitrary and lack normative significance.[219] It seems, then, that a non-relativistic version of internalism might supply the goods because whatever are our starting motivations, reason determines what we ought to do and what we ought to do together, if we ought to do anything together. Thus, the non-relativistic version of internalism promises to preserve the

[218] (Smith 1994; 1995).
[219] (Smith 1995, 124).

motivational, i.e., explanatory, capacity of normative reasons while preserving their normative significance.

Internalism, generally speaking, is a position that contains a requirement on normative reasons that says that "the desirability of an agent's φ-ing in certain circumstances C depends on whether she would desire that she φs in C if she were fully rational".[220] When we say that an agent has a reason to φ, we mean that her fully rational self would desire that she φs in those circumstances.[221] This is what Smith calls the 'advice model' as opposed to the 'example model' of the internalist requirement on reasons. An agent's reasons are determined by what a fully rational agent would advise her to do in her actual circumstances. The example model, on the other hand, provides reasons for a fully rational agent, not an actual agent who in all likelihood is irrational to some degree. What it means to be fully rational is for an agent to possess no false beliefs and all relevant true beliefs along with a systematically justified desiderative profile. The work of deliberation is to sort and discard desires such that the end result is a coherent and unified set of motivations. Having a unified and coherent set of motivations exhibits structural rationality, so the desires of the fully rational agent are systematically justified. Starting with the desires an actual agent has, the fully rational deliberator will form evaluative beliefs about particular desires being subsumed under general desires. These general desires justify and explain the more particular desires because they show them to be unified and coherent. The general desire behaves as a principle that governs the particular desires. The fully rational agent will adopt the general desire on the basis of the evaluative belief because it is rationally required. In this way, deliberation adds new desires to a rational agent's motivations, but they also make the original smattering of desires more unified and

[220] (Smith 1995, 110).
[221] For Smith, this is an analytic claim about the concept 'reason'. There are many doubts about whether this really is a conceptual claim, for instance, (Copp 1997; Sayre-McCord 1997). But for our purposes, I grant that Smith has a successful argument for internal reasons.

coherent. In addition, there are desires in the original profile, even ones which originally appeared indispensable, which cannot be subsumed under the more general desires. As Smith says, "perhaps because we can see no way of integrating those desires into the set as a whole they will come to seem *ad hoc* and so unjustifiable to us".[222] Believing that they are unjustified will cause a rational agent to abandon those desires. So, through this process of systematic justification a fully rational agent comes to have a unified and coherent desiderative profile.[223]

At this point, it seems that Smith has just amended Williams' account of deliberation, thoroughly eliminating any role for imagination But, in fact, Smith does not put too much stock into the starting desires of a desiderative profile either because he thinks that rationality ultimately will make our profiles converge. That is, fully rational agents would act in the same way in the same circumstances. This is because the concept of a reason is non-relative. Unlike Williams, 'There is reason to φ in C' does not mean 'There is reason to φ in C for *me*'. This is because the circumstances in which reasons are provided capture all the so-called relativity. The circumstances relevant to a reason are under-specified when it is claimed that reasons are properly relative. Two cases show this clearly. First, agent-relative reasons appear to be relative to the agent. And they are, but in a normatively irrelevant way. For instance, I can have reason to save *my* drowning child where a stranger might not have that same reason. But what this means is that my circumstances are different to strangers in the situation. The fact that *my* child is drowning is a relevant to the circumstances where it is irrelevant to those who do not stand in that relation to my child. The child's mother would have the same reason as I do. Second, and more pertinent to Williams, reasons might be relative due to our psychology. But again, an individual's psychology is just part of

[222] (Smith 1995, 115).
[223] It should be noted that Smith seems to believe that evaluative beliefs are part and parcel of the desiderative profile despite not being desires. As we will see presently, agents might not have a desire or a general desire but the fully rational agent believes that X is desirable and so would desire it. At times, Smith talks about not only the desiderative profile but a coherent and unified evaluative outlook. This is because evaluative beliefs and desires should be "counterpart[s]" of each other (Smith 1995, 115).

her circumstances and those who had a relevantly similar psychological makeup would have the same reasons in the same situations.[224] For example, the fact that I prefer to exercise after work and you prefer to exercise before work does not mean that our reasons about when we exercise are relative—it just means that when these preferences are relevant, they enter into the circumstances which partly constitute any agent's reasons.[225] Any fully rational agent in the *same* circumstances will desire the same thing, she will have the same reasons.

These considerations, along with the fact that we can disagree about reasons, indicating that we are arguing about a common subject matter, lead to Smith's main "Kantian" argument for non-relativity.[226] The crux of this argument is that systematic justification of a desiderative profile does not stop at the boundary of the individual: desires must be assessed no matter where they reside because rationality is not partial. The process of systematic justification is meant to remove the arbitrariness of any desire and confer on it normative significance. Suppose that two agents are in the same circumstances and their desires conflict with respect to those circumstances, a fully rational agent must determine which desire to retain and which to relinquish. This is done by a fully rational agent determining which desire falls best under general desires that yield a more unified and coherent desiderative profile. The reason why a fully rational agent must decide amongst these desires is because when a contradictory consideration appears the normative status of those desires is threatened. One or the other is more justifiable given the desiderative profile. The fully rational agent considers *all* desires, because they all have equal footing, in the process of determining what we have reason to do. The fully rational agent, then, is not limited to systematically justifying *my*

---

[224] Smith ought not to go too far in pushing this point. Uncontrollable irrationalities like the tennis player's anger or a person's preferences are different than irrational desires. So, he should acknowledge that only non-rationally assessable features of our psychology are part of our circumstances. The fact that an agent is unmoved by a reason when they should be moved cannot be part of their circumstances because it is the role of reasons to rationally compel agents.
[225] Williams has an avenue to respond to Smith, namely, that it is very difficult to specify the principle which determines the relevancy of a fact such that it is included in the circumstances. For Williams, the facts which are relevant are one's subjective motivational set and rationality.
[226] (Smith 1995, 118).

desiderative prolife. Smith writes: "Which desires *I* end up with, after engaging in such a process, thus in no way depends on what *my* actual desires are to begin with, because reason itself determines the content of our fully rational desires, not the arbitrary fact that we have the actual desires that we have".[227] Of course, an *individual's* desiderative profile plays a role in systematic justification as constituting part of the source material for fully rational deliberation, but the reasons that result from the process in no way depend solely on them. In fact, none of those desires may survive the process. Internal reasons are non-relative because rationality requires convergence in desiderative profile across agents with respect to how they would act in any given circumstances.

How the advice model, systematic justification, and the non-relativity of reasons fit together is a difficult matter of interpretation. I have stressed that an individual's starting desires do not matter much for Smith's theory because rationality will deliver the objective truth for all agents in the circumstances. The fact that we would all converge on the same desires were we fully rational does eliminate the relativity of internalism, in one sense—we do not have different reasons in the end. But it is difficult to square systematic justification with the advice model because the advice model depends on parts of an agent's conative psychology being unassailable in the determination of an agent's reasons. And yet, no part of an agent's conative psychology was supposed to be unassailable in the process of systematic justification. This suggests that an individual's psychology has a more prominent role in determining reasons, that is, it is the basic input to deliberation and partially limits the results. Deliberation begins with desires and reasons from them to more general desires that will converge with others' fully rational selves in *those* circumstances. Convergence depends on a particular irrationality being taken as part of the circumstances which the fully rational agent must countenance. Even if we in fact converge in those circumstances, *where* we converge matters. That is, the content of our reasons, despite convergence, is still relative to *our* starting

[227] (Smith 1995, 124).

desires, and this may be very far away from the reasons we consider people to have. Take Williams' abusive husband example to illustrate. The husband cannot help but be violent due to his particular irrationalities. What would his fully rational self advise? Even though his fully rational self would not desire to abuse his wife, the husband cannot give up his abusive tendencies. So, the advice might be that he has reason to be minimally abusive, given that he *will* be abusive to his wife.[228] Convergence says that we would all end up at such a reason if we were in his circumstances. But this is still a highly paradoxical result: does he *really* have reason to minimally abuse his wife? Even though convergence is the logical result of rationality, it does not mean that reasons being relative to motivations is not problematic for Smith. Because it is a version of internalism, starting motivations still matter for what reasons agents have and this can lead to some ugly results.

Applied to group reasons, I think it is clear that the non-relativist version of internalism can produce group reasons. Considering two agents with rationally imperfect psychologies, the fully rational agent could advise them what to do in some circumstances. Because the fully rational agent knows all the relevant facts, i.e., the co-agents' desiderative profiles, she can advise them about what they ought to do together. Imagine two overly lusty lovers who happen to be the only people on a section of beach and who generally desire to help people. They have no motivation whatsoever to stop their canoodling, and their lust would override any other course of action that gets in its way. At this time, they hear a call for help and notice two people who appear to be drowning and whom they could easily help out of the water. Also, the location of the lifeguard tower does not have a clear line of sight to the swimmers but it does to the lovers. What do the lovers have reason to do together? Intuitively, they have reason to go and help the two drowning people. But this does not take their lust into consideration—they are in an irrational state and will not help the swimmers.

[228] The advice model might say that he should stay away from his wife, but this would not respect his particular irrationality which is part of his circumstances. Just like the angry tennis player must be angry, the abusive husband must be abusive. The question is what to do under such conditions.

Given that their irrationality is part of the circumstances in which they find themselves, the fully rational agent might advise that they merely get the attention of the lifeguard by waving their arms together and, then, they can return to each other, confident that they did the right thing. Thus, the theory can provide an account of group reasons because a fully informed and rational agent can advise co-agents on what to do together all the while taking into consideration their irrationality.

Even though agents' desires would converge in circumstances C, inclusive of irrationality, if they were fully rational, a non-relativist internalism does not yield a plausible account of group reasons. The problem, like the relativistic version, is not that the theory cannot produce group reasons. But instead, there is no way for either co-agent to spontaneously grasp their group reason. That is, the co-agents will have to engage each other in order to be able to arrive at their reasons to act together. Fully informed and rational agents have no trouble grasping group reasons because it is in their nature to have all of the relevant true beliefs. For actual agents, however, who are not by nature fully informed, they need to gather all of the relevant information. Seeing as Smith builds parts of our psychology into the circumstances that partly determine the reasons agents have, I do not see how an agent could come to have that knowledge without interacting with her co-agents. Perhaps one lover's lust is not so overpowering that she cannot act on a reason to help the swimmers, but the other is too enthralled to stop. How could she know the effect of her lover's passion on her motivations without interaction? The same problems arise for the non-relativist version of internalism as the relativist one. It is not a matter of whether the theory can produce group reasons—it's whether it can produce group reasons in a way that they can play the role of explaining and justifying joint actions. This is not exactly a problem of normative epistemology because there is no serious difficulty in acquiring the information relevant to figuring out what reasons co-agents have. The problem is that that source of knowledge is off-limits when group reasons are required to explain spontaneous joint actions. Because agents' psychologies do matter

for determining the circumstances and the desiderative input for deliberation, agents could not know their group reasons without interacting with others in order to access their desires and irrationalities in that moment. Even though Smith's position is non-relative, it is still beholden to the psychologies of agents.

This problem might make one wonder whether we should follow Smith in adopting the advisory model of the internalist requirement. Because the advisory model depends on an agent's psychology to determine her reasons, perhaps the problem can be avoided by adopting instead what Smith called the 'example model'. He attributes this model to Korsgaard and is described as: an agent has reason to φ in C insofar as her ideally rational self would desire to φ in C.[229] Reasons, on the example model, are for an agent who has no irrationalities whatsoever, the ideally rational self, and it is up to the actual agent to meet the demands of reason. On this model, co-agents would not need to know each other's psychology to grasp their group reasons, they would merely require a sound deliberative route to them. So, if agents do not need to know about their co-agent's psychological profile in order to grasp their group reasons, it appears that co-agents could act together spontaneously. It is worth stating the motivation for internalism again, which is that in some way reasons must be able to motivate agents to act to count as reasons. Recalling the half-furious doubles team, it appears to be true, on the example model, that both players have group reason to shake their opponents' hands. But the point about the one player's fury, that is to say, her insurmountable irrationality, is that that group reason cannot possibly motivate her to act on it. The fact of her irrational state undercuts the motivation for internal reasons on the example model: the irrational agent has reason to act but cannot possibly do so; so, the reason *cannot* motivate her; so,

[229] Korsgaard does not formulate the internalist requirement in this way. She describes it variably as "practical-reason claims, if they are to present us with reasons for action, must be capable of motivating rational persons" and it requires that "rational considerations succeed in motivating us insofar as we are rational" (Korsgaard 1986, 11, 15). One might wonder why Smith smuggles the term 'ideally' into the example model when Korsgaard does not use it. But we should remember that 'ideally' here means a systematically justified desiderative profile that does not include any irrationalities, something Korsgaard seems to require in order for reasons to have authority over an agent.

she does not really have an internal reason. Korsgaard might complain that this motivational impossibility is overstated, after all reasons need not always motivate even though they should do. But this complaint does not take seriously how debilitating irrationality can be for an agent: in the circumstances of the tennis match at that particular time in the player's life, it could not be otherwise for her to be furious and to smash her opponent if she got close to her.[230] And for internal reasons to get a grip on an agent, it must be possible for the agent to act otherwise. If this were not true, we would lose any sense of rational success and failure, that is to say, any sense of rationality and irrationality. Even though the example model does not rely on agent's psychologies like the advice model does, it still has a problem explaining how reasons get a grip on agents when they are insurmountably irrational. The example model of the internalist requirement does not offer a solution to internalism's problem. Thus, the trouble with internalism is that it fails on its own terms, namely, that there must be a connection between reasons and action, but group reasons could never explain spontaneous joint actions because the agents lack the relevant epistemic standing, or motivational possibility in the case of the example model.

Perhaps some do not find this argument persuasive. Let's grant that the argument is not completely successful. It still poses a challenge to internal reasons because the cases in which they can explain spontaneous joint actions will be highly limited. A good theory should aim to explain all the instances, and if another theory can do better, then that is the better theory. Perhaps Williams can rely on imagination in ways that I have not considered such that co-agents can imagine their ways to corresponding group reasons. But this is a matter of luck because it depends on particular agents with determinate imaginative capacities. Considering the Shaq case, which appears to be a spontaneous joint action, will be a spontaneous joint action if the co-agents are so lucky or, if

[230] I want to note that the necessity involved here is rather difficult to specify. Of course, there is a possible world where the agent is not furious. However, indexed to these specific circumstances, it could not be otherwise.

unlucky, it will be an instance of mutually coordinated individual actions. The point here is merely to emphasise that even if internal group reasons can explain *some* spontaneous joint actions, it is an impoverished theory because it will leave out many prima facie cases, and there is a better theory on offer or so I argue below.

There is one form of internalism which I have not considered here because I think it will be limited in similar ways. Imagine a Kantian form of internalism where an individual's practical identity of being human gives a priori reasons in virtue of being human. Given that each of us are human and share this practical identity, why can't I assume that we share group reasons? Since we all share the same reasons and our practical identities would provide reasons a priori, we wouldn't need to interact to find our group reasons. All it would take for us to act together spontaneously is to recognise the reasons that we share. This is a very quick sketch of a position, but it is worth saying why it is inadequate.[231] Let's suppose that this theory does generate group reasons that we all share in virtue of being human. The problem is that the theory is not fine-grained enough: it would only produce agent-neutral reasons, and most of our group reasons appear to be quite relative. It is rare for every human to share a group reason, whereas it is far more common that smaller numbers of people have reason to act together (excluding others). It is hard to see how such a theory could yield agent-relative group reasons because the means by which it produces reasons is shared by all of us. Because of this extensional limit, a Kantian form of internalism such as this will fail to supply an adequate account of group reasons even if it could occasionally explain spontaneous joint actions.

3. Group Reasons in General

Since internalism has failed to provide a satisfactory account of group reasons when coupled with the spontaneity requirement, it is important to characterise what group reasons are in general and to

[231] I do not intend this scant description to represent Korsgaard's constructivism (Korsgaard 2009). To reconstruct how Korsgaard's already difficult theory can produce group reasons is at least a paper in itself. Though, Korsgaard's discussion of publicity will be discussed in the next chapter.

whom they apply. I briefly made some claims about this in the previous chapter. Here, I intend to argue for those claims. The notion of a practical normative reason, as a consideration that counts in favour of an action, is formally construed as a relation, R(consideration, agent, circumstances, action).[232] In natural language, this relation is expressed with the predicate "is a reason for". The first thing to note is that there is some fact, a consideration, that partly constitutes a reason for an agent. For instance, the fact that the flight is long is a reason to bring along a small pillow. Or, to borrow Scanlon's example, the fact that the knife is sharp is a reason not to grab the blade. Because these considerations are facts, they are available to everyone under the right conditions: that the knife is sharp, as long as an agent has those concepts and has good epistemic standing with respect to that fact, is graspable by her. No one in this debate disagrees that normative reasons are facts of some sort because, by definition, they are true.[233] But reasons are not identical to the consideration-facts. Not all reasons normatively engage all agents in the same ways; for instance, the fact that the prime minister's flight is long is not a reason for *you* even if it is for him. For a fact to be a *consideration*, the fact must somehow recommend or disfavour an action. As Tyler Burge writes, "all reasons that thinkers have are *reasons-to,* not merely rational appraisals".[234] Again, the consideration that the flight is long is not merely a standard against which an agent's behaviour is assessed, indeed it rationally compels her to bring a pillow. Part of what it means to be a reason is that it can motivate agents to perform actions, so the recommended action is an essential part of the reason relation. Moreover, just because there is a relation between considerations and actions does not

[232] See (Scanlon 2014, 30–33) for a discussion of this along with footnote 21 where he discusses different formulations of the reason relation. He thinks that his characterisation covers those other varieties. I take it that the internalists would accept this formal relation, but tricky ontological issues may arise. For instance, relativistic internalists might not require an agent variable. The reason is identical to a motivation, actual or hypothetical. However, motivations are always motivations of agents, so it would be redundant to include an agent variable. In that case, the relation would be a different relation. Further, the reason relation is not limited to practical reasons but, for my purposes, I will only discuss them.

[233] Of course, there are a variety of views which deny that there are facts about reasons, such as expressivism or moral error theory.

[234] (Burge 2006, 250).

mean that a fact is always related in the same way to a type of action. This is why the circumstances around the fact matter. For instance, the fact that the flight is long does not give the pilot a reason to bring a pillow. In such a case, the fact of a long flight likely recommends a different action, such as getting a good night's sleep the night before. So, the very same consideration in different circumstances can recommend different actions. Further, a fact in one circumstance may be a consideration where in another it may not be at all. For example, the fact a student needs a good grade on her paper to pass can recommend that she head to the library to study. However, were she in a warzone, needing a good grade to pass does not seem relevant in determining what she should do in those circumstances.[235] This is all to say that when specifying what reasons an agent has it is essential to index them to the circumstances in which she finds herself. The final part of the reason relation is the agent. The fact that reasons are for agents means that they are the ones who ought to act in the recommended way. Part of the concept of a reason is that it can make an agent modify her behaviour, but not necessarily. An agent might not always respond to a reason, but that does not mean it does not make a claim on her to act in a certain way. That is to say, she *ought* to act in a certain way whether she in fact does or does not do so. For an agent to understand her reasons means that she feels their normative force. Agents don't use reasons merely to appraise actions, as when a voyeur might acknowledge that a mugger is doing something wrong without feeling compelled to act. Normative reasons have a necessary grip on an agent's thinking and behaviour such that she will have *failed* as an agent if she neglects her reasons or *succeeded* if she acts on them. The concept of a normative reason is partly constituted by the fact that reasons apply to agents who ought to act on them.

[235] This last point might be disputed. Indeed, what I say in Chapter 5 will cause me to deny it. However, it is plausible, at the conceptual level, that the same fact, in different circumstances, may not be a consideration.

So far, these have all been formal features about reasons. But there is a more substantive question that matters to the inquiry at hand: how must the agent variable in the reason relation be filled? Is there anything in the idea of a reason that indicates how the agent variable ought to be understood? Traditionally, the agent variable is filled by a *singular* agent, usually an individual. A typical example is "the worker has reason to eat lunch". Reasons apply to agents, and agents are individuals. Clearly, individual reasons are the most common and intuitive. However, it is not uncommon to say, for instance, "the Supreme Court has reason to overturn the conviction". Here, the court is a singular entity constituted by a number of Justices and organised by a specific set of rules. As long as a theory of agenthood allows these singular group agents to exist, then it seems that there are reasons that apply to them insofar as they are agents.[236] However, singular reasons for groups do not extend to any particular member of the group as an individual because that would alter the agent of the reason. For instance, if the court has reason to overturn a conviction, no particular justice has reason to overturn the conviction—in fact, it is not in her power to do so.[237] These reasons are still, what I have been calling, 'singular reasons' because they are for singular agents. The group, e.g., the supreme court, is one agent that performs an action on the basis of a singular reason. These kinds of agents, if they exist, should also satisfy the agent variable of the reason relation in addition to individual agents. But now the question is whether there is reason to limit the application of reasons to singular agents? From a conceptual point of view, is it not

[236] Other examples of singular groups might include demographic groups, e.g., Hispanics, regional groups, e.g., Westerners, and institutional groups, e.g., a philosophy department. How these singular groups are constituted is a lively debate, but the prominent view is that they are functionally determined from a subordinate set of individuals. For instance, 'Alberta believes that it ought to secede' is true insofar as the majority of Albertans believe that the Province ought to secede. See (Pettit and Schweikard 2006; List and Pettit 2011; List 2018). It should be noted that a singular group's action can be constituted by a collection of unintentional individual actions. That is, functionally determined groups might act whether or not the constituent members intend to act together. So, there is only a subclass of singular groups to which this discussion applies.

[237] It might be the case that an individual can do what the group ought to do. But this still does not mean that the reason for the group is her reason. If the reason was hers, that is, she filled the agent-variable, then the reason would not be for the group. These are ontologically different reasons due to having different constituents.

possible for a reason to apply to more than one agent? I do not think that there is good reason for such a limitation. Reasons can apply to a unified multitude of agents, which I will call a plural group or a plurality. The concept of plurality still implies unity, which is needed for joint action, without denoting singularity.

Three interconnected conditions must be met for group reasons to apply directly to co-agents: reasons must be able to modify the co-agents' behaviour; they must be able to have a grip on the co-agents; and they must allow descriptions of failure and success with respect to the co-agents' actions. Imagine two co-agents who have a group reason to help up Shaquille O'Neil after he has fallen on the sidewalk directly in front of them. The co-agents, after each acknowledges the group reason, help Shaq up. It is because of the group reason that each co-agent gave Shaq a hand rather than walking by him. In light of the group reason, each modified her behaviour. Likewise, the group reason has a grip on the co-agents because they are rationally compelled to act on its basis. Imagine, though, that one co-agent is in fact rushing to get to an important meeting for which she is already late when she acknowledges the group reason to help Shaq. Insofar as she acknowledges it and, by that token, it has a grip on her, she deliberates about what to do—continue on to the meeting or help Shaq with this other person. The fact that there is a need for deliberation in her mind as a result of acknowledging the group reason is a mark of the grip of reasons. Perhaps she misses the reason entirely and it does *not* get a grip on her. However, were she in a position to acknowledge it, the reason *would* have a rational grip on her. To be able to grip an agent, reasons do not need to be acted on nor acknowledged by the agent, rather were they to be considered, their force would be felt in deliberation.[238] In the group case, it seems that group reasons have a grip on each co-agent directly otherwise they, as individuals, would not consider them in deliberation or feel

[238] It is possible that this deliberation is over so quickly because one reason so overwhelmingly outweighs the other that it barely appears to happen. The claim about the grip of reasons is not an entirely phenomenological claim though. Instead, were the reasoning made explicit, consideration of the reason would be exhibited.

their force at all.[239] Finally, a co-agent can be individually evaluated with reference to her group reasons. It seems entirely appropriate to say that a co-agent has failed to respond to her reasons when she has a reason to φ together (with another) and, for no good reason, is not rationally compelled by it. Imagine now that both co-agents standing before Shaq are only out for leisurely strolls. When one co-agent walks on by the helpless Shaq and the other helpful co-agent, it seems appropriate for the other to criticise or demand that she help—that is, to let her know that she is failing to respond to her reasons. Likewise, it seems appropriate to say that each co-agent has rationally succeeded if she responds to their group reason. The success of their joint action will depend on each co-agent performing her role in the joint action, but her rational success depends on her appropriately responding to her reasons. On the basis of these considerations, there appears to be no impasse to thinking that the agent-variable can be filled by more than one agent. This means that group reasons apply to a plurality or a plural group constituted by co-agents who act together.[240]

It will help to clarify the notion of a plural group a bit more and how they differ from singular groups. Both are often constituted by a multitude of individuals, but there can be a difference in the kinds of actions performed by these groups. Plural groups, when they act, only act *together* intentionally. An action by a singular group, on the other hand, is not necessarily an instance of individuals acting together intentionally. For instance, for Canada to apologise to indigenous peoples, the Prime Minister or some other high-ranking official alone needs to apologise. In this case, only one person's action constitutes the group's action. This is because the group is an entity to which reasons apply over and above the members of the group. The actions of singular groups

[239] If group reasons were singular group reasons that apply to group agents, they would be used to appraise the group agent, even if the appraiser were a member of the group. The singular group reason does not get a grip on group members. For instance, Canada has reason to apologise to indigenous communities. This does not mean that each Canadian ought to apologise on the basis of that reason, *Canada* should, given an adequate procedure of public apology. Nevertheless, this normative state of affairs might give Canadians individual reasons to make sure the country acts on its group reason.

[240] The point of this term is to leave metaphysical space for a collection of individuals who do not form a group over and above the members and who can still legitimately use the collective 'we'.

can be performed by one individual. This is a considerably different notion than that of acting together intentionally. It is different because it is not possible to act together intentionally if only one person performs an action—two or more people did not do anything. Another difference is that (singular) group members can discordantly produce an event for which the singular group is responsible. Consider Smith's discussion of the invisible hand. By each person serving her own interests, they allegedly better humanity—they better humanity collectively. Under this assumption, the members of humanity do not engage in joint agency and the singular group, humanity, can be responsible for bettering humanity's lot. On the other hand, joint actions performed by plural groups require at least more than one member of the group to act with another intentionally.[241] Plural groups are unified by their group reasons such that the co-agents' actions count as acting *together*, but they are also multitudinous in that they are not distinct from the members who constitute them. It is in virtue of the circumstances in which co-agents find themselves that they have group reason to act together and it is in virtue of the group reason that co-agents constitute a plural group at all. Unlike singular group agents, plural groups come and go by group reasons applying to co-agents.[242] When two people intentionally move the piano together, they constitute a plural group that disbands once the action is done. Plural groups are unified by the group reasons which apply to their constituent co-agents. Lastly, group reasons directly apply to the co-agents who ought to act on them. Whereas reasons for singular groups apply directly to the group and then member(s) may carry out the required actions in the name of the singular group. So, plural groups, that is, a unified multitude of co-agents, are held together by the group reasons which directly apply to each co-agent. And this plurality can fill the agent variable in the reason relation.

---

[241] Margaret Gilbert's theory of joint action is sometimes called the 'plural subject' view. The plural subject is that which is picked out by uses of 'we', which includes co-agents. Although the plural subject is constituted by co-agents, it has an existence over and above any one of them. In virtue of this fact, I take the plural group which I now discuss to be distinct.

[242] The same group of individuals may have more than one group reason unifying them. These pluralities would survive so long as they have a reason that applies to them.

The fact that the agent variable can be filled by a plurality of co-agents thrusts an additional question on us: what is the nature of the grip that group reasons have on co-agents? Are group reasons normatively public or private? The publicity of reasons has two meanings that need to be parsed. Publicity as intelligibility only entails that others can understand another's reasons, or, negatively, an agent can only have a reason if another can, under suitable conditions, grasp that the agent has a reason. The intelligibility of reasons says that anyone can recognise what others have reason to do, under the right conditions, though their normative force only applies to the relevant agent. Intelligibility does not carry with it normative force.[243] On the other hand, normative publicity means, very generally, that a reason transmits normative force to more than one person. On this reading, normative force is not a private matter. Others do not merely understand another's reasons and use them to assess or evaluate an agent's rationality, but they are also rationally compelled to behave *in some way* on the basis of the same reasons. This does not mean that all agents are compelled to do the same thing by the same reason—that is clearly false—but that reasons make a claim on more than one agent.[244]

Normative publicity is an important feature of group reasons because each co-agent has to be rationally compelled to act on them and, in virtue of the plurality of agents, their normative force transmits to more than one person. It is not the case that a co-agent interprets the other and then realises that her reasons align with the other's. That would describe two agents who have two matching singular reasons. Rather, it is the very same group reason that the co-agents share, even if they are required to perform different instrumental actions to execute the joint action. At bottom, the importance of publicity to group reasons is that reasons can apply directly to more than one

[243] The more accurate claim here is to say that intelligibility does not *necessarily* entail normative force. If it can be established that reasons are normatively public (in the right way), then understanding a reason would constitute being gripped by it. However, this requires argumentation.
[244] It is false because an agent-neutral reason requires each agent to act in ways that are suitable to their own circumstances.

individual, which group reasons, by definition, do. However, one might object to this construal of normative publicity. After all, the co-agents are picked out by the group reason just as a singular agent is picked out by a singular reason. So, the publicity of group reasons is analogous to regular, old singular reasons: the normative force of the reason does not extend to those who are not addressed by the reason. There is no rational claim on others, and yet they can understand that co-agents have group reasons. However, there is a very simple reason to deny that group reasons are public only in terms of intelligibility. An account of group reasons must, in some sense, be committed to normative publicity because if normative force is a private affair, then group reasons do not exist because they could not apply to a plurality of agents. The general concept of privacy involved in this debate concerns one person alone:[245] if two people share something, that thing is no longer private. If at least two people share a group reason, which they would by definition, then group reasons are public in some sense. To deny the normative publicity of reasons is to deny the possibility of group reasons. The claim that group reasons really do exist needs defense, which I do in the next chapter, but it is a conceptual matter that group reasons are normatively public.

This scepticism about publicity does force a distinction to be made. 'Publicity' defined as 'normative force transmitting to more than one person' has an unrestricted and a restricted interpretation. This definition does not tell us how far normative force extends. The unrestricted interpretation claims that the normative force of a reason transmits, in some way, to all rational agents. Every reason, in some way, makes a rational claim on every agent. A group reason's claim on most people may only be for them to avoid interfering with the co-agents to whom the group reason directly applies, but that is still a rational claim nonetheless. The unrestricted interpretation has many objectors, as we will see in the next chapter. The restricted interpretation, on the other

[245] This debate derives from similar considerations as those concerning the privacy of language. See, (Wittgenstein 2009; Korsgaard 1993; 1996; Wallace 2009).

hand, claims that the normative force of group reasons only transmits to those directly addressed by the reason. Two people, for instance, can have a group reason to act together while others are not rationally compelled by that group reason. This is still a form of publicity because normative force is not private, but it need not transmit to everyone.[246] We'll see in the next chapter that the unrestricted interpretation is defensible but it is not the case that the definition of a group reason settles this matter. The point of making this distinction, though, is that it is possible for group reasons to be public without their normative force necessarily transmitting to all agents. To some this possibility may be more palatable because the restricted interpretation preserves the idea that group reasons are intelligible to those to whom they do not apply. One upshot of this distinction, though, is that we cannot simply carry over a plausible analysis of singular reasons to an analysis of group reasons. It is at least plausible that singular reasons are normatively private whereas group reasons by definition are normatively public even if they do not make a rational claim on those to whom they are not addressed. In the end, there is nothing in the concept of a normative reason that excludes the possibility that group reasons apply directly to co-agents.

4. Realism about Reasons and Group Reasons

Seeing as group reasons apply to more than one individual by definition and they must be able to explain spontaneous joint actions, what theory of normative reasons is available? Sections 1 and 2 showed that group reasons cannot be motivating nor internal reasons. The option that best satisfies these requirements is realism about reasons: reasons exist independently of what we think

[246] One might find this to be an odd distinction if one is coming from Wittgenstein's private language arguments. There, publicity means available for everyone, not available for some and not others. But there is a difference between intelligibility and normative force that explains the disanalogy. An agent's irrationalities are intelligible to a degree but that doesn't mean that the standards constituting those irrationalities have normative force. For instance, an agent who is going to murder someone is intelligible, but whatever she takes to be her reason does not transmit any normative force to me. (I may have reason to interfere with her course of action, but that reason is grounded in something other than that which she takes to be her reason.)

(construed broadly) about them, or what our motivations are.[247] The first thing to note is that, unlike internal reasons, real reasons do not depend on a relevant motive, actual or hypothetical.[248] For instance, if I have reason to call my mother, then that is true irrespective of any motivation that I have or might have. Lacking a relevant motivation or a deliberative connection to it does not falsify a reasons statement.[249] This is to say that realist reasons are, in some sense, mind-independent—their existence as *reasons* does not depend on the mind.[250] Second, judgments about reasons are true in virtue of realist reasons. Because reasons exist independently of an agent's mind, judgments about reasons can be true or false depending on whether they accurately represent normative reality. This entails that realism is committed to cognitivism about judgments about reasons: the nature of our beliefs about reasons is that they aim to truly represent normative reality. Being a cognitivist, on its own, does not make one a realist about reasons. Rather, it only says that our judgments are truth-apt and anti-realist theories, such as moral error theory and internalism, can agree with this claim.[251] Finally, because normative reasons exist independently of us and normative judgments are about them, some normative judgments are true. Plenty of normative judgments that agents make are false, but in principle there are some which are true.[252] If normative judgments are truth-apt and reasons exist (mind-independently), then there must be some true normative

---

[247] This position has also been called externalism or objectivism about reasons. Some examples include, with dispute, (Nagel 1970; McDowell 1978; 1985; Sturgeon 1985; Railton 1986; Scanlon 1998; Parfit 2006; 2011; Enoch 2011).

[248] There may be *some* real reasons that depend on psychology, but this won't be true all of the time. For instance, the fact that I have a false belief is reason to abandon it. In these cases, psychology is treated as part of the world rather than playing the part of gripping the agent.

[249] There is, indeed, an issue with how real reasons can bear on motivation. But the fact that we have reasons does not depend on the mind's stance towards them.

[250] The kind of mind-independence involved in realist theories of reasons varies, but it is not an issue for my purposes. The last chapter will discuss this in a bit more detail.

[251] For instance, see (Mackie 1977, chap. 1). Internalism is also a position that has a cognitivist strand: Williams did argue that reasons statements can be true or false. Some internalists also claim to be realists, for instance, (Railton 1986; Smith 1991), because there are truths about what an agent has reason to do. But these theories are revisionary because they analyse 'reason' in terms of a suitable conception of desire. Revisionary realist theories allow normativity to be held hostage by the mind.

[252] This final feature of realism further distinguishes it from error-theory which believes in cognitivism but denies the possibility of true normative judgments. Error theory, as construed by Mackie, denies the existence of mind-independent values, so there is only false normative judgments. Note that on the realist conception an individual can *happen* to lack any true normative judgment, but it is possible that she has true ones.

judgments. These last two features of realism, cognitivism and true normative judgments, are shared by internalism about reasons. What really distinguishes realism is its commitment to the mind-independence of reasons. But where do these reasons come from if they do not come from *our* values or attitudes? The answer is simple: reasons are a part of reality just like other facts such as the colour of the wall, the Pythagorean Theorem, and the law of non-contradiction.[253] The reasons which exist can be elucidated with reference to the other facts which make up the reason relation. And because they are facts, reasons can be *discovered* as opposed to constructed. Often, though, reasons are immediately available to the agent as are other facts about her surroundings.

It is noteworthy that realism is a much more elegant theory than the others discussed beforehand. However, when a theory is elegant philosophers are typically forced to play defense. And there are many objections to realism. Some of the main ones being: how can reasons motivate if they stand at a distance from us? How are they related to the natural world? And how can we know them?[254] These objections cannot be discussed in any detail here as they are besides the goal of this chapter, which is to determine which theory of normative reasons can supply a suitable account of group reasons for the explanation of spontaneous joint actions, were those reasons to exist. In the following chapter, I will argue for realism about reasons and say something a bit more detailed about each of those objections. The previous chapter should serve as a partial answer to the motivation objection: because desires are judgment-sensitive, we are responsive to our judgments about reasons that can truly represent normative reality, including our judgments about group reasons. Nevertheless, answers to these objections require putting more cards on the table, which, for now, is unnecessary.

---

[253] Even though colours are 'secondary' qualities, they are still real in the relevant sense (McDowell 1985). If we stopped talking about colours because they were not real qua 'primary' quality, we would no longer be able to explain a wide array of experiences that we have. Likewise, even though mathematical and logical truths do not have a natural basis and seemingly are not involved in causal explanations, they are still truths. What is real is not only limited to what is causally explicable or what the sciences study.

[254] Each of these objections, in some form, is raised by Mackie in his classic chapter, (Mackie 1977, chap. 1).

So, what are realist *group* reasons and why do they help to explain spontaneous joint action? First, group reasons, like individual reasons, would exist as part of the world. They would be there for all agents to recognise and to be gripped by insofar as they were in the right epistemic circumstances. This means that realist reasons are essentially public, in one normative sense or another: more than one person can recognise on the basis of the circumstances and the considerations contained in a situation what reasons there are and those reasons are normatively compelling. For instance, everyone can recognise that the fact that the dog is vicious is reason to keep one's distance; it is even a reason to stop another person from getting close. In the group case, the fact that Shaq has fallen in front of you and another person and he can't get up is reason for both of you to help. The fact that group reasons exist as part of the world means that they are public in terms of intelligibility and in terms of normative force.[255] Group reasons' publicity does not entail that reasons are immediately self-evident or recognisable. Instead, it means that agents are never in principle severed from their group reasons and can recognise their authority, under the right conditions. Just as a citizen might be unaware of a law, an agent might be unaware of her reasons—their ignorance though has no bearing on reasons' normative authority. The fact that they are normatively public means that it is possible for more than one person to recognise them under the right conditions, not that more than one person does recognise them. This, though, should not be a problematic feature of realist group reasons because, as cognitivism about normative judgment implies, an agent's judgments about her reasons can be false. The fact that group reasons are discoverable is partly owed to the fact that they are public. If, for realism, group reasons really are out in the world, they should not just be there for one individual alone—by definition, they are reasons for more than one agent. This is not the case, for example, for Williams' internalism: I might be able to interpret B.W.'s (internal) reasons provided that I had access to the relevant

[255] It is not the case that all versions of realism are committed to publicity in both the unrestricted and restricted sense.

information and his style of deliberation. His reasons would be intelligible to me, but they don't have normative force over me unless there is a sound deliberative route from my reason determining features to a reason consistent with B.W.'s. This is to say that a different reason than B.W.'s normatively bears on me. Real group reasons, instead, are partly constituted by public considerations and circumstances in principle available to everyone. The fact that group reasons are normatively public explains how they have normative authority over co-agents. Upon discovering their group reasons, co-agents should be gripped by them without needing to interact with each other. Co-agents can discover their group reasons by investigating the world.

Publicity of group reasons goes some way in helping to explain spontaneous joint actions because co-agents can be gripped and act from their real group reasons without interaction. But what really separates realism from the other theories is the fact that group reasons are not determined by the minds of the agents whom they address. Unlike internalism, real group reasons are solely the product of the factual circumstances in which they find themselves. The other co-agent is important only because the group reason applies to her as well, but she is not, in principle, an epistemic point of access to the group reason. Each co-agent does not need to know the other's psychology in order to recognise the group reasons that they share—the co-agent's psychology does not determine group reasons. One might wonder here about Smith's case of the angry tennis player. Suppose two people just got demolished at doubles and one is so angry that she cannot avoid violence if she gets close to her opponents. Does the irrational state of one co-agent partly determine what they have reason to do together, e.g., shake their opponents hands as a team? This must be denied for realist group reasons to explain spontaneous joint actions. If a co-agent's irrational state partly determined her group reasons, then realism about group reasons would suffer from the same problem as internalism: co-agents would need to interact in order to know a reason-determining circumstance. An angry tennis partner might not show her anger and yet still be unable

to shake hands, so her calm partner cannot trust appearances. Consequently, her partner would need to interact with her in order to realise that her anger dissolves any group reason to shake hands. Spontaneous joint action, in such a case, would be impossible. This would seem to speak against the realist account or, perhaps, at this point, even, the existence of spontaneous joint actions. However, realism should affirm that reasons are determined by considerations and circumstances in the environment because it preserves a very plausible normative framework in the group context. The tennis partner who is unaware of her partner's irrational state will judge, on the basis of environmental features alone, that they have reason to shake hands with their opponents. She walks over to shake hands and realises that her partner stormed off to the changing room. If she judged correctly that they have reason to shake hands, she has grounds to ask questions of her partner and blame her for her behaviour. If she judged incorrectly, because their reasons are partly determined by the irrational state of her partner, then she has no such ground and she was just wrong about her reasons. The right answer though appears to be the former, she does seem to have a case against her partner for not shaking hands. If there were no group reason, the partner would not be justified in asking "why didn't you shake hands?" or saying "you ought to have shaken hands". And it seems that the partner has good reason to ask why she had to fail in doing what *they* had reason to do together because her partner did not do her part. This reason is based on another good reason, the group reason to shake their opponents' hands. The normative framework that I have described can still accommodate the claim that the angry partner is justified in heading to the changing room. That the angry player has a group reason to shake hands and that she has (singular) reason to head to the changing room is not inconsistent insofar as these are both pro tanto reasons that need to be weighed against each other. The consideration that she will harm her opponents might very well *override* the group reason to shake hands. The angry player who acts on this reason is *excused* for not acting on the group reason, but this excuse does not eliminate the normative force of the group

reason for both co-agents. So, the angry partner can be justified in not shaking hands while the calm partner is justified in demanding an explanation for her actions. None of this depends on having epistemic access to the psychology of co-agents. Realist group reasons avoid the epistemic problem of internalism by affirming the mind-independence of reasons.

Perhaps one agent does not need to know the psychology of a co-agent to access their group reasons, but what about her abilities? If I am ignorant of a potential co-agent's abilities, then how do I have access to our reasons? It seems that in order to know what we ought to do, I need to know what we can do. Assuming that 'ought implies can' is true, I do not think this epistemic claim follows from it. In the individual case, whether or not I have a particular reason depends on my relevant abilities, but whether I *know* of my abilities does not effect the ontological status of my reasons. My epistemic position might excuse my failure to recognise my reasons, but it does not follow that those reasons simply do not exist. Likewise in the joint case, the existence of real group reason is not effected by our epistemic standing. But what is challenged by abilities-ignorance is whether we could *access* our real group reasons and act spontaneously on them. Perhaps I need to engage with my fellow co-agent in order to find out if she can do what I think our group reason requires of us. In response, this possibility, I think, looks at realism at the same time as having an eye on internalism. Real reasons, individual or group reasons, present themselves to us as bits of the world, and we may be wrong in our judgments. But epistemic evaluation only happens after the reason presents itself. So, whether or not my co-agent has the relevant abilities may factor into my evaluation of whether or not we *truly* have this group reason. But the possibility of being wrong about our group reasons does not make it impossible for us to act spontaneously on real group reasons that present themselves to us. Not knowing my co-agents' abilities is not necessarily an epistemic barrier to real group reasons, though it may affect how I evaluate the group reasons that

present to me.[256] In the good cases, real group reasons simply present to us and we can act spontaneously on them.

Because group reasons are public and mind-independent, realism can explain spontaneous joint actions, unlike internalism. To illustrate, a real group reason picks out two or more individuals and recommends that they act together in some way. These reasons, being public, are there for the agents to grasp and be gripped by. If the agents are gripped by the group reason, they can act together spontaneously. Even though it is possible that agents are in a position of ignorance about their group reasons, this ignorance does not does pose a problem for the theory to explain spontaneous joint actions. Of course, if agents are unaware of their group reasons, then they will not act together, let alone spontaneously. But once they do grasp the group reason, they have everything they need to act together. They know (within limits) with whom they should be acting and they know what they are doing together. Most importantly, they can know all this without any binding interaction between them. The epistemic problem for internalist theories of group reasons is that the source of knowledge had to be the other co-agent and the means of acquiring that knowledge required interaction between the two co-agents. This interaction counts as binding interaction and conflicts with the spontaneity requirement. Even if these theories can produce group reasons, they cannot *explain* spontaneous joint action. This problem does not exist for realist group reasons because the world provides the reasons for the agents without the need to investigate her co-agents' psychology or interact with them. Because group reasons are mind-independent and, so, in principle, co-agents have access to everything they need to discover their group reasons, each co-agent has epistemic access to her group reasons and, hence, is in proper epistemic position to act

[256] Ignorance of abilities, even in the individual case, seems to be relevant to the correctness of a judgment about reasons rather than to the access of the reason. For instance, I do not know whether I can hike a 1000m trail, but that ignorance does not stop the reason to do so from presenting itself.

together spontaneously. There is no requirement for an agent to interact with her co-agent prior to acting together. Furthermore, even when group reasons are not immediately evident to co-agents and they must discover them by reasoning, just like the internalist will have to deliberate towards her reasons, co-agents can deliberate with the facts of the matter to arrive at their group reasons. They are not required to engage with their co-agents to discover the relevant facts, though that may be the most effective means of discovery. For the simple reason that realist group reasons are public and mind-independent, there is no need for them to interact with each other in order to come to know and act spontaneously on their group reasons.

5. Conclusion

This chapter has not argued that realism about reasons is true; it only argued that realism would be the best account of normative group reasons such that they must be able to explain spontaneous joint actions. I take this conclusion though to lend support to realism's cause in its battle against anti-realism. What good would normative group reasons be if they cannot play some role in rationalising joint actions? The challenge was for theories of normative group reasons to explain co-agents being able to act on their group reasons without any prior interaction. It was argued that group reasons cannot be motivating reasons nor internal reasons, relativistic or non-relativistic. The problem with each theory was the same: because reasons are beholden to the attitudes of agents, the only way for co-agents to know their group reasons would be to interact and discover the relevant motivations of their co-agents. This violates the requirement on spontaneity for joint action. What was needed instead was a theory of group reasons which are not necessarily related to the psychology of the co-agents, and this is exactly what realism provides.

There is a serious epistemic challenge often levelled against realism that would seem to threaten the arguments that I have made.[257] The previous section granted that co-agents could *know*

[257] This objection is again in (Mackie 1977, chap. 1).

their group reasons. Nevertheless, the traditional epistemic objection to realism begins with the claim that normative facts are not reducible to natural facts, a claim that is explicitly endorsed by non-naturalist realists.[258] Now, how could an agent come to know such facts if they are not reducible to natural facts? Most of what we know comes from our engagement with the natural world. It appears that the only available option is some faculty of 'normative intuition', a highly suspect faculty. Non-naturalist realists typically respond by analogy: normative facts are like mathematical facts. We know mathematical facts just by thinking about mathematics. And the truths of mathematics are not problematic, so normative truths are not problematic. It seems that normative and mathematical truths are known via very similar faculties. However, there is a serious disanalogy between mathematical and normative truths: normative facts seem to crucially depend on natural facts in a way that mathematical facts do not. For instance, the reason for Anne not to grab the blade is because it is sharp, and the blade's being sharp is a natural fact. This is why non-naturalists inevitably end up arguing for some kind of supervenience relation between normative and natural facts. This requires a further epistemic capacity of the agent to not only recognise normative truths but to know that those truths connect with the world in a particular way. This is a serious challenge to realism and it is a potentially serious challenge to my position of joint action. But if this epistemic objection holds, then it is only a problem for non-naturalist realism. Even though non-naturalism is the status quo realism, there are other options, in particular, naturalist realism, which claims that normative facts *are* natural facts. There are a variety of naturalist realisms, however, many turn out to be revisionary.[259] By that, I mean that they primarily are concerned with accounting for how normative language works, and this often comes at the price of robust normative *force*. The epistemic objection can be avoided by adopting naturalist realism, but the trick

[258] Paradigmatic examples include (Moore 1903; Parfit 2006; Scanlon 2014).
[259] (Sturgeon 1985; Railton 1986; Smith 1994)

is to be naturalist and non-revisionary. In the next chapter, I present Davidson's version of moral realism and show how it meets the objections discussed in this chapter. The point of the next chapter though is to argue that normative reasons, and by extension, group reasons, exist. We will find that Davidson's argument for moral realism is conceptually tied to joint action, hence, justifying why Davidson has such a prominent role in this dissertation.

CHAPTER 5

The Publicity of Group Reasons

A group reason is a reason that we share, not in some distributive way where we each have reasons that match, but we have the reason jointly. The protesting Algerians do not each have their own independent reasons to protest. Rather, they share *a* reason to protest together. A single group reason applies to multiple individuals and counts in favour of their performing an action together. This description may seem to make group reasons inherently relativistic, that is, relative to the individuals to whom the reason applies. Indeed, it seems that the shared nature of group reasons renders them agent-relative reasons of a special sort. But this would be a mistake because it is possible that there is a group reason for all agents to act together. Perhaps, we all share a group reason to save the Amazon together. Group reasons will most often be partial or selective in some way, but it is not impossible for them to be agent-neutral, insofar as it is a group reason that applies to everyone.[260] The seeming relativity of group reasons is merely a function of the innocuous relativity of the reason relation: all reasons must apply to agents and, so, some agent or agents will always be picked out by them. The fact that a group reason may often be for some agents and not others does not make it relative in an interesting way. But this now raises an interesting question: if it isn't relativity, what makes it possible for a group reason to be shared? After all, a principled sceptic might very well doubt the possibility of sharing reasons in this manner *because* reasons are normatively private. In order to make the idea of sharing a single group reason intelligible and to persuade the principled sceptic, group reasons must be normatively public in some way. The very idea of a group reason depends on the possibility of the transmission of normative force to more than one individual. To make good on this idea and to vindicate the account of spontaneous joint

[260] How this distinction is characterised will affect the truth of these claims. But the intuitive distinction between reasons that apply to everyone and reasons that apply partially does not help us clarify group reasons.

action developed in Chapter 3, two claims must be proven: first, that group reasons in fact exist and, second, that reasons in general are unrestrictedly public.

The last chapter concluded that group reasons, which are needed to explain spontaneous joint actions, are best accounted for by a realist theory of normative reasons. The main idea is that for co-agents to be able to grasp and respond to a group reason without binding interaction, group reasons had to be determined by non-mental, public features of the world. Moreover, group reasons are inherently public because, by definition, they apply to more than one person, that is, the normative force of a group reason is not private. In virtue of its commitment to mind-independence, realism can rather easily accommodate the inherent publicity of group reasons. Consequently for realism, there is no prima facie impasse to their applying to more than one individual. Even though the concept of a group reason entails publicity, which realism accounts for best, a sceptic might doubt the existence of group reasons entirely: just because realism *would* be the best theory of group reasons does not amount to the truth of the ontological claim that any such reasons exist. Scepticism about real reasons is typically guided by finding them metaphysically objectionable and epistemically mysterious. Importantly, such scepticism threatens the account of spontaneous group action developed in Chapter 3 because the account depends on the existence of group reasons. Without real group reasons, an adequate account of spontaneous joint action would be impossible.

This chapter has the further task of clarifying the way in which group reasons are public. Recall that the publicity thesis about group reasons has an unrestricted and restricted interpretation. Are group reasons public merely because, by definition, they are considerations for more than one individual to φ or do they exert normative force on everyone, *in some way*?[261] The former, restricted interpretation is established by the conceptual point about group reasons: a group reason's

[261] 'In some way' is intended to indicate that a consideration can issue different commands to distinct people.

normative force extends to those co-agents who are required to act on it. The restricted interpretation of the publicity of group reasons would be vindicated by the fact that spontaneous joint actions occur and a convincing argument in favour of the existence of reasons is provided, thereby rebutting scepticism about realism. But might group reasons be unrestrictedly public? Mightn't their normative force matter to everyone and not only those who are explicitly picked out by them? A complete account of group reasons must settle this matter. A significant motivation for settling this matter is that unrestricted publicity makes the existence of group reasons more palatable. It might seem that a special kind of reason, being necessarily public, is bootstrapped into existence by the fact that spontaneous joint action happens and reasons are real. This conjunction of claims is not the strongest argument for publicity and group reasons, and a sceptic may still deny the existence of spontaneous joint action because the consideration of normative privacy is just too strong.[262] However, if reasons in general were unrestrictedly public, reasons, and not just group reasons, exert force on more than one person in every case. The difference between singular and group reasons would simply be how many people are required to act and the kind of action to perform. In this case, publicity would not be a special feature of group reasons alone and, as a result, it would be more difficult for a sceptic to deny the existence of group reasons, their public nature, and, by extension, spontaneous joint action.

Conveniently, the argument that I offer in favour of realism also supports the unrestricted interpretation of publicity. However before introducing that argument, I explore other arguments in favour of unrestricted publicity. While doing this, I will shift to discussing singular reasons, partly because that is the focus of those arguments and partly because if it can be established that singular reasons apply in some way to everyone, then there is good reason to accept that group reasons are

[262] Note that the two interpretations of publicity do not arise for singular reasons. Since they only apply to an individual, restricted publicity is not an issue.

unrestrictedly public as well. Seeing as group reasons are public by definition, it would be an odd result if the normative force of singular reasons applied to everyone while group reasons did not. What would account for this difference? Though it is possible that singular reasons are private while group reasons are restrictedly public, this difference would be accounted for by the peculiar nature of group reasons. The question of unrestricted publicity applies to reasons in general and if it can be established that singular reasons apply to everyone, then group reasons do as well.

I will cover two different kinds of argument for publicity.[263] The first argument centres on the agent-neutral/-relative distinction. I call this the 'argument from agent-neutrality'. Despite providing insights into the experience of normativity, this argument ultimately fails because agent-relative reasons appear to be public just as well as agent-neutral ones. The second argument, from Korsgaard, I call the 'private reasons argument'.[264] Analogous to Wittgenstein's famous 'private language' argument, Korsgaard argues that reasons must be normatively public and not merely intelligible. In the end, we will see that this argument fails because it does not take our sociality seriously enough.

Following these attempts to vindicate publicity, I will focus on providing an argument for the publicity of reasons that, at the same time, is an argument for realism and the existence of group reasons. This argument is Davidson's triangulation argument, an argument that was incipient in his earlier work.[265] Its purpose was to show that thought and language in part depend on two individuals interacting with each other and their environment. Unlike Korsgaard's use of Wittgenstein, Robert Myers has shown that the triangulation argument can go beyond intelligibility

[263] A different argument for publicity that I do not consider is (Julius 2023, chap. 7).
[264] (Korsgaard 1996).
[265] Once the triangulation argument is on board, it is very difficult not to see it in the background of, at least, "Radical Interpretation" (Davidson 2001b, chap. 9). It has been argued by Claudine Verheggen that Davidson does not undergo a radical transition in thought, but rather forged deeper into the foundations of his earlier work (Verheggen 2017).

to robust normativity.[266] The basic idea is that agents triangulate on normative properties in their environment, properties which provide reasons. Because those properties are triangulatable, they must be public to the agents. Further, the reasons provided by these properties are real because it is true, independently of what agents think, that they have reason on the basis of the normative property and their circumstances. This achieves half of what is necessary: reasons are real, public, and mind-independent. But the final goal is to show that *group* reasons are real, public, and mind-independent. I argue that the triangulation argument shows this as well, and it is in fact the best way to understand the *act* of triangulation. It is, in the first instance, a joint action, and there has to be a reason why individuals would triangulate in the first place. Why don't they stay in their solitary worlds eating berries and lounging under trees?[267] They have to be motivated to triangulate because, after all, triangulation is an action. And, given that Davidson's account of desire is sensitive to judgments about reasons, the motivation that explains action will depend on an agent responding to reasons. Because triangulation necessarily involves at least two people, that is, an individual cannot do it alone, the reason that is the source of the triangulators motivations must be a group reason. Triangulators, in the first instance, are acting on the group reason to triangulate jointly together.

1. Publicity and the Argument from Agent-Neutrality

The publicity debate begins with Nagel and Korsgaard who defend what I call the argument from agent-neutrality. Despite their differences in argument and approach, their arguments share a common structure and arrive at their conclusion on the basis of the same assumption. The argument from agent-neutrality concludes that reasons are public because agent-neutral reasons unrestrictedly apply to everyone, and they are the only valid non-derivative reasons. A non-

[266] Although language is widely considered to be normative, this typically means that error in the use of words is possible. It does not typically mean that an agent is *required* or categorically ought to use words in some way.
[267] Interestingly, Davidson claims that without triangulation, agents cannot have a *world* because there would be no distance between themselves and the environment (Davidson [1992] 2001, 119).

derivative reason is a reason that does not depend on another reason for its normative force, such as a reason for an instrumental action that depends on the reason for the end. Agent-neutral reasons are incontestably public since they apply to everyone indiscriminately. But the trick is to show that they are the only legitimate non-derivative reasons. Structurally, the solution is rather simple since it depends on the cogency of the agent-neutral/agent-relative distinction as Nagel first described it.[268]

Nagel tells us that the distinction is formal. Reasons are expressed by predicates, which, in the agent-relative case, contain a "free agent-variable" and, in the agent-neutral case, do not. Agent-relative reasons refer to the agent who is to perform the action in the content of the reason. For instance, if TN has reason to publish *his* writings, then he and no one else should make his writings public on the basis of this reason. When agents have agent-relative reasons, the free agent-variable is occupied by agent-referring terms, such as 'his', 'her', and 'my'. His reason and his alone tells him to publish his writings, not some other person's. TN would not be responding to his reasons by publishing anyone else's work but his own. Agent-relative reasons apply to and have normative force on a particular agent such that she performs an action or brings about events that concern her. Further, no one else has reason to get TN to publish his writings on the basis of his agent-relative reason: "the free agent-variable prevents the transmission of derivative influence to any acts of another person".[269] This is to say that agent-relative reasons are normatively private. Despite this limitation of normative force, every agent, in the same circumstances, would have the same agent-relative reasons, that is, they must be universal reasons. A plausible interpretation of the claim 'everyone has reason to preserve her life' is that each individual has reason to preserve her *own* life. It tells no one to do anything that would preserve another's life—her life is her concern. Agent-neutral reasons, on the other hand, lack a free agent-variable and, as a result, their normative force

[268] This distinction has been notably criticised by (Korsgaard 1993), to be discussed shortly, and (Rønnow-Rasmussen 2009).
[269] (Nagel 1970, 93).

does not stop at the boundaries of an individual. The reason statement, 'everyone should publish TN's writing', picks out the agents who are to promote the end of 'publishing TN's writings'—everyone. They are not each required to individually publish TN's writings, as that would likely result in conflict and maybe even frustrate publication. Instead, everyone should do what they can such that TN's writings are published. So, for instance, a colleague should encourage TN to publish; another should provide written comments to him; TN should send his work to a publisher; the factory worker should email TN and demand that he publish (likely, though, the factory worker merely has reason not to interfere). Depending on the circumstances and the end in question, each individual can be required to perform different actions in order to satisfy the demands of an agent-neutral reason. In this way, agent-neutral reasons can produce derivative agent-relative reasons, e.g., a colleague should provide TN with *her* comments. Likewise, a person might have already helped TN but since he still has work to publish, she may still be required to do more. This is to say that the normative force of an agent-neutral reason extends to everyone to promote its end, however their part contributes to the end. For agent-neutral reasons, everyone is responsible for promoting the general end whereas, for agent-relative reasons, each is responsible to promote her own end. Agent-neutral reasons are (unrestrictedly) public because everyone is subject to the normative force of the reason. The whole point of the distinction for Nagel is to explain how some reasons have limited normative transmission and other's do not.

With this distinction in hand, the argument from agent-neutrality can be presented. Agent-neutral reasons are public by definition—they apply to everyone and issue derivative reasons that promote the general end.[270] However, due to the free-agent variable that prevents normative transmission to those not mentioned in the reason, agent-relative reasons are private. No one else

[270] The characterisation of reasons as 'promoting' is worrisome because it construes reasons as being for *events* and not for *agents*, as they are typically thought to be. This seems to conflate the normative and explanatory roles of reasons.

has reason to act on them, save the referred to agent. Now, both Nagel and Korsgaard supply arguments intended to show that *non-derivative* agent-relative reasons lack normative force. For Nagel, all reasons can be represented from the personal and impersonal perspective because of our self-conception that we are one among equally real others. Valid non-derivative reasons must have motivational and justificatory 'content' from both perspectives, that is, the two perspectives must be congruent. Whether an agent thinks of her reasons from the personal or impersonal perspective, her reasons should be able to compel her to act. However, agent-relative reasons are problematic because, depending on the perspective, agent-relative reasons lack either motivational or justificatory content. Agent-relative reasons are not valid non-derivative reasons because they would be inconsistent with our self-conception. Hence, all valid agent-relative reasons must derive their force from an agent-neutral reason.[271] Consequently, all non-derivative reasons are public because all non-derivative reasons are agent-neutral.

For Korsgaard, valid non-derivative reasons are reasons that agents can share and, by definition, agent-relative reasons are not sharable. Sharable reasons are agent-neutral and inherently public.[272] This claim is supported by Korsgaard's account of values as intersubjective. Korsgaard does not think reasons and the values that ground them are objective or 'out there in the world'. Instead, intersubjective values are constructed from our individual values that stand up to the scrutiny of our practical identities.[273] Individual values are filtered through the practical identities of

[271] At the end of the day, Nagel changed his mind about the relationship between agent-relative and agent-neutral reasons. In the postscript to *Possibility of Altruism*, he thought that the argument establishes a different conclusion, that every agent-relative reason has a corresponding agent-neutral reason. This means that agent-relative reasons do not derive their legitimacy from agent-neutral ones. Later still, Nagel argued that there are some agent-relative reasons that have no corresponding agent-neutral ones and that this is genuinely puzzling (Nagel 1986, chap. 9).
[272] Korsgaard's use of 'shared' to mean public should be kept distinct from the sense of 'shared' as I used it to describe group reasons at the start of this chapter. In the final analysis, rationality and morality may be joint endeavours for Korsgaard, though it seems clear that her public reasons are not reasons for joint action.
[273] (Korsgaard 1993, 28)

the agents who are in particular relations, and the values that can be shared between them survive.[274] What this means for agent-relativity is that it has effectively no normative significance because all reasons have their basis in intersubjective values that are necessarily shared. This is to say that *all* reasons are agent-neutral, construed intersubjectively. Agent-relativity is relegated to agent-relative *motives* that respond to agent-neutral reasons.[275] The structure of the argument from agent-neutrality consists in an affirmation of agent-neutral reasons, which are public, as the only kind of reasons and a total denial of agent-relative reasons.[276]

Both Nagel's and Korsgaard's arguments have been effectively criticised. Sturgeon dismisses Nagel's congruency condition on which his argument depended and Korsgaard's account of shared reasons and values has been attacked by many for a variety of reasons.[277] However, rather than pursue these lines of argument, R. Jay Wallace takes a different approach. He questions whether the formal feature of agent-relativity, the free agent-variable, actually commits us to the claim that agent-relative reasons cannot have normative significance to anyone but the referred to agent. Wallace begins his interpretation of publicity by arguing against the teleological justification of actions in order to make space for agent-relativity. The basic idea behind teleological justification is that there is some valuable state of affairs that will be *produced* by and (pro tanto) justifies an agent's action. The trouble with teleological justification, when assuming publicity, is that it appears to eliminate agent-relativity. What is valuable and is the basis of reasons is *produced* by an agent's actions. And since reasons are public, every agent will be justified, to some degree, to promote the valuable state of affairs. It is difficult to see how agent-relativity can survive this combination of teleology and

---

[274] At times, Korsgaard suggests that different people have different reasons relative to their *particular* relationships. I find this to be too relative for Korsgaard's own aims. Wallace is also concerned with the coherence of this claim (Wallace 2009).

[275] (Korsgaard 1993, 38)

[276] It should be noted that Korsgaard views this argument as an argument against the distinction. Korsgaard seeks to preserve agent-relativity by relegating it to the psychological domain.

[277] (Sturgeon 1974; O'Day 1998; Norman 2000; LeBar 2001). Criticisms of Korsgaard's argument often mix together the argument employed here with the one from *Sources of Normativity*, which I discuss in the next section.

publicity when value calls every agent to promote it. For instance, helping a friend because *we are friends* would not be a valid teleological justification because no state of affairs is promoted by helping her. Instead, the agent-relative feature, our friendship, must be described as a value that can be promoted, such as the value of friendship itself, something that is valuable to every agent. To protect agent-relativity from this consequence while preserving the publicity of reasons, Wallace offers a non-teleological model of justification where, at least in agent-relative cases, the relation of an agent to another or a project grounds reasons.[278] To see how this model differs, consider helping a friend move house. I help a friend move because she is my friend—the consideration of our relation of friendship is what provides the reason to help her move. No doubt that there is a valuable state of affairs in helping her move, but it is based in *our* relationship rather than in the state of affairs of having moved house. It seems unlikely that it is valuable to help people move house without a relation to the particular person.[279] The difference with the teleological model is that instead of, first, looking for values that justify courses of action, the non-teleological model starts with reason-giving relations which then confer value on a state of affairs. This model preserves agent-relativity because relation-based reasons will depend on particularities of agents that cannot be generalised across people. The question is how to understand the publicity of agent-relative reasons if everyone does not possess those same reasons.

The answer is rather simple and intuitive: reasons and values come along with a requirement of non-interference.[280] To recognise that another is acting on reasons and doing something valuable entails the recognition of the inappropriateness of interference. That is, the same consideration that

---

[278] Wallace is careful to note that teleological justification can have a place in moral theory, it just cannot be the only model of justification (2009, 474).

[279] There is a question here for Wallace how the non-teleological model gets a grip on the value of helping a random person move. I am not sure how to approach it, but the answer does not bear on the developing discussion.

[280] Agent-neutral reasons also supply this requirement: if such a reason requires another to help someone, others should not interfere with her reasonable action so long as she is helping. Nagel also notes that an agent-neutral reason might require non-interference because that is the best way for an agent to promote the agent-neutral value from where she stands.

grounds an agent's agent-relative reason also grounds others' reasons of non-interference. In order to avoid unwelcome interpretations, interference should be understood as intentional, that is, to interfere consists in the intention to "prevent or make more difficult another person's pursuit of their goals".[281] This interpretation has the benefit of ruling out mere causal interference, such as when people unavoidably and excusably get in each other's way on a busy sidewalk. Interference is entirely different when someone has the aim to do so: it's inexcusable for someone to aim only at blocking another's way. To see how this requirement aids the publicity of agent-relative reasons, consider the moving example. If you have reason to help your friend move because she is your friend, then others have reason not to (intentionally) interfere with your valuable activity. Of course, your reason could be overridden by other more significant reasons that justify another's interference, but this is just to say that there is a pro tanto requirement of non-interference. To interfere with someone's valuable activity requires justification. Wallace nicely summarises this point, "it is antithetical to value to endeavor to destroy or to undermine it, and this fundamental consideration explains the basic normative significance of our reasons for other agents".[282] One thought that needs to be abandoned when thinking about publicity is that a reason requires everyone to perform the *same* action, such as bring about the good. This thought holds publicity hostage to agent-neutrality. With respect to agent-relative reasons, one or some people have reason to act on the basis of their relation to a person or project and that very same consideration (pro tanto) requires others to avoid desecrating the value of that action. So, Wallace has made space for the publicity of agent-relative reasons by proposing a non-teleological model of justification and because agent-relative reasons are reasons for others not to interfere.

[281] (Wallace 2009, 478)
[282] (Wallace 2009, 478)

This though is only an interpretation of publicity applied to agent-relativity that must be supported by argument. Before discussing Wallace's argument, this interpretation of publicity on its own challenges the main claim of the argument from agent-neutrality, that only agent-neutral reasons are public. The rationale there was that since agent-relativity makes reference to a particular individual, there is no way for an agent-relative reason to have normative significance for a person who is not explicitly picked out by the reason. For instance, the value of watching her child's school play is of value to the parent because of their special relationship; that value seemingly does not bear on those who stand in a different relation to the play and its performers. The problem for the neutrality argument is that there is a perfectly intelligible characterisation of public agent-relative reasons that includes a requirement of non-interference by those who are not explicitly referred to by the reason. So, the parent's agent-relative reason to go watch her child's school play is also a reason for others to avoid intentionally interfering with her carrying out the valuable activity. However, Nagel and Korsgaard might worry that the requirement of non-interference and the agent-relative reason are not in fact based in the same consideration, that is, there is a mere correlation between them.[283] It might be the case that the requirement of non-interference is based in some agent-neutral value, like respect, rather than the value that provides the agent-relative reason. This does seem plausible but it misconstrues the crux of Wallace's argument, which is to acknowledge that someone has good reason and that their action is valuable, even if the value is for her alone, and to intentionally interfere with her valuable activity is to destroy or undermine that value. These two judgments cannot be held together. To recognise an activity as valuable is to recognise that it 'ought to exist'.[284] And to act contrary to an ought is to act contrary to reason, meaning that there is some requirement present. The requirement of non-interference is the

[283] Wallace anticipates this objection (2009, 486).
[284] I say 'ought to exist' as a general characterisation of value. I am not committed to this as an analysis but something along these lines must be right.

minimal requirement provided by value. Hence, the argument from agent-neutrality should not be accepted because there is an acceptable interpretation of publicity that accommodates agent-relativity. Interestingly, unlike arguments from agent-neutrality, this interpretation eliminates the need to argue that agent-relative reasons have no non-derivative force, something that grates against intuition. In terms of the intuitiveness of Wallace's interpretation, it should be noted that it fits best with our reason-giving practices, such as advice, blame, dialogue, etc.[285] For instance, if agent-relative reasons were not public, you would not be justified in advising a child to treat her mother better *because she is her mother*. Now, the question is whether Wallace has a good argument for publicity even if he offers the most intuitively appealing interpretation of it.

Wallace's argument for publicity is an "inductive generalisation" and is justified by 'our considered verdicts' about what reasons we have.[286] The induction begins by taking into account our first-order judgments about reasons and then reflecting on them in order to clarify what reasons we have and to determine any patterns that may arise. Starting with substantive judgments ensures that the patterns that reveal themselves actually describe our substantive reasons as opposed to relying on structural requirements, that is, it ensures that the result describes normative experience. Naturally, this process includes not only the agent's reasons but the reasons of others such that when we reflect on reasons we reflect on reasons generally.[287] The purpose of reflecting on reasons in this process is to arrive at the reasons which are endorsed and are considered authoritative in deliberation. Via this process Wallace thinks that, what he calls, the two-sided 'justified interference pattern of practical judgments' becomes explicit and justifies his interpretation of the publicity of reasons. The first side of the pattern concerns legitimate interference: it is necessary to judge that a

[285] To see how these practices depend on publicity, see (Bertea 2017).
[286] (Wallace 2009, 484)
[287] I want to note that it is unclear in the text whether reflection is solely for the individual, whether people are to do it communally, or whether it is a hypothetical process. This is an issue for how exactly to understand 'considered' in 'considered verdicts'. I imagine that Wallace is on the hypothetical side since he is a realist about reasons and the other options retain some element of relativity.

person has no good reason for what they are doing in order to legitimately interfere.[288] This does not justify interference, as it is only a necessary condition; in some cases, interference may still be viewed as illegitimately paternalistic when an agent has no good reason for her actions. The second side is just that when we judge others to have good reason, then we appear to be required to leave them alone. The justified interference pattern vindicates publicity because our considered verdicts about reasons bear witness to the interpretation of publicity that allows for agent-relative reasons. When we recognise value or reasons, we cannot help but include them in our deliberations. By reflecting on reasons and arriving at the justified interference pattern and publicity, we make explicit "the normative self-understanding that is implicit in our reflection about other peoples' reasons and their relation to our own".[289]

I am not convinced that this argument is very satisfying because it heavily trades on intuitions about our normative experience. Even though Wallace does not take this argument to depend on 'philosophical biography', it still assumes that there are no fundamental differences in normative experience. For instance, imagine that an egoist is reflecting on reasons via this method. She has a radically different conception of reasons and sees the world in her egoistical way. She does not believe that there are any reasons or values for others that matter to her. If this difference is not merely theoretical but describes how she actually experiences normativity, then it is difficult to see how she would be convinced that she is required to not interfere with others because she does not recognise their reasons as reasons. Similarly, Korsgaard defends her position on the basis of normative experience, but she does not come to the same conclusion as Wallace about agent-relativity. So the question is how much stock can be put into Wallace's inductive generalisations

[288] Although I will not challenge Wallace on the pattern that he identifies, there appears to be reason to doubt this necessary condition. A person may have good reason for what they are doing and yet interference is still justified by stronger, overriding reasons.
[289] (Wallace 2009, 487)

when such a generalisation requires similar intuitions that may not be available? To my mind, the argument sets up a clash of intuitions that does not settle the matter. Moreover, our considered verdicts may change in light of new information, so there is no guarantee that the interference pattern survives reflection and is in fact true. The method described in Wallace's argument should be open to this possibility since it depends on our actual judgments. If this is the case, it is difficult to see why we should be confident that the generalisations truly describe normative reality and human experience. Now, I believe that Wallace's interpretation of publicity is correct, but his argument for it is not convincing. Perhaps this is the best we can come up with when we think about reasons, but if a direct argument is available that does not invite a clash of intuitions, then it should be pursued.[290] Before moving on, it is worth emphasising why Wallace's interpretation is valuable. First, it does appear to adequately describe normative experience. This is certainly an aim of any philosophical claim, but it alone does not justify accepting the claim. Second, and most importantly, it gives agent-relative reasons non-derivative normative standing. One effect of the argument from agent-neutrality is that agent-neutral reasons swallow up relativity and make it difficult to understand how, despite appearances, any form of agent-relativity is normatively significant. But Wallace's interpretation makes intelligible how agent-relative value has a seat at the table with agent-neutral value. In light of these considerations, an argument for publicity of reasons should aim to accommodate Wallace's interpretation of publicity.

2. Publicity and the 'Private Reasons' Argument

In *The Sources of Normativity*, Korsgaard offers a different argument for the publicity of reasons, one similar to Wittgenstein's 'private language' argument. That argument goes something like this: meaningful thoughts and utterances depend on public standards because private standards of

[290] Wallace, after discussing Korsgaard's argument that I cover in the next section, notes that his strategy may not be the only one to vindicate publicity (2009, 482).

meaning are indeterminate or lead to a vicious regress and, therefore, are no standard at all.[291] Korsgaard's argument is rather difficult to interpret and many object to it because they interpret it as an implication of Wittgenstein's argument: that the publicity of reasons falls out of the publicity of language and thought.[292] That result though would be a baffling conflation of normativity with intelligibility. Despite the fact that language is normative, in the sense that the meaningful use of words has conditions of correctness, the normativity of reasons is different. Correct usage of a term does not obligate in the way a reason does because it partly depends on what an agent wants to say in any given circumstance. Even though Korsgaard does not make it easy for readers, this interpretation is, I think, uncharitable. The better way to interpret Korsgaard's argument is as an analogy to the private language argument, so I will reconstruct the argument so that it bears a likeness to Wittgenstein's while focusing on right kind normativity.[293]

The private reasons argument, as I understand it, has two strands that together are supposed to amount to the publicity of reasons and the impossibility of private reasons. The first strand aims to show that standards are necessary for reasons whereas the second aims to show that those standards are necessarily public. Without standards, reasons would lack their normative status and could not explain the ir/rationality of an agent. And if reasons are a kind of standard, then they must be public because standards are public. Further, if reasons were private, then they could not obligate others, and our reasons do obligate others. The first strand, then, makes a conceptual claim about an agent having reasons and the second shows that reasons are fundamentally public as a

[291] The correct interpretation of Wittgenstein's argument (2009) is a matter of lively debate, but these claims are widely acknowledged. Some influential interpretations are (Kripke 1982; Mcdowell 1984; 1992; Boghossian 1989).
[292] (O'Day 1998; Norman 2000; LeBar 2001; Wallace 2009). Korsgaard has denied this reading (2009, 196), though her argument has received distinct criticism (Hurley 2001; Gert 2002).
[293] A note on Korsgaard's terminology is needed. She, again, aligns the terms 'public' and 'private' with 'agent-neutral' and 'agent-relative' (Korsgaard 1996, 133 n.3). These terms, though, do not play an important role in the argument and it might be possible to understand the argument with Wallace's interpretation of publicity, that all reasons have some normative force over all agents.

result of human sociality.[294] I aim to show that Korsgaard has not taken advantage of human sociality in a way that helps the argument.

Before discussing the first strand, it is worth noting that Korsgaard believes that two views could never explain public reasons. A 'Hobbesian' view of reasons, according to which reasons are grounded in self-interest, can only explain the normative significance of another's reasons to an agent insofar as they bear on her interests. Moreover, a neo-Kantian view, according to which reasons are determined by moral law, can only explain the normative significance of another's reasons insofar as consistency or rationality require it. But both of these views suffer from the same problem according to Korsgaard: they start with private reasons and try to build up to public ones. The Hobbesian is still only guided by her self-interest—if another's reasons do not bear on her interests, she has no reason to consider them. Likewise for the neo-Kantian, another's reasons do not actually compel the agent: she is compelled by rational requirements which require her to act in a way consistent with another's reasons. Korsgaard believes that these views do not sufficiently take into account the thoroughly social nature of normativity.

The first strand of the private reasons argument aims to prove that reasons are normative in an obligatory sense and that this is unintelligible if they are private. For something to count as normative, it must be possible to go wrong or get it right with respect to that thing. Normativity exists in a space where error—not merely discordance—is possible. For instance, if someone bumps into me and I spill a drink, I did not make the *mistake* of spilling the drink. I didn't go wrong because I could not help but do it—I was *caused* to do so. Similarly, if a desire presents itself to me and I have no means of scrutinising it, then I could not possibly go wrong by acting on it. Perhaps I

[294] The main points of these strands, at least as I understand them, have been interpreted as two distinct arguments (Gert 2002). One reason to consider these strands as parts of the same argument is that the first strand makes no mention of other agents, and it is at least a plausible thought that an agent can have reasons on her own even if other agents are not compelled by them.

determine that this desire is good, right, etc. and I should act on it again in the future when it appears. This determination though is no help in the future because there are no means to compare a new desire with the old one whose normative force has extinguished with its satisfaction. So here again, I cannot go wrong in the future because there is no standard against which to compare desires and determine them as the same, that is, as according to a rule. How is it that I can make a mistake then? I must be a reflective, thinking agent—"what obligates me is reflection".[295] If I can reflect on the desires that present themselves, then I can scrutinise them, I can hold them up to a standard. Reflection makes error and, simultaneously, norms and reasons possible. An ability to reflect on desires creates a critical distance between desires presented and the norms used to assess them. Were I to act on a desire that my reflective consciousness could not endorse, I would act irrationally or without reason. Normative relations require a social scene, according to Korsgaard, specifically a legislator who lays down norms and citizens who obey.[296] And this is what a reflective consciousness is: a thinking self who endorses standards and applies them to an acting self who ought to obey those standards. In a reflective consciousness, desires that present themselves are inescapably scrutinised by the standards which ought to guide the agent as a whole. Crucially, an agent's mind is already a social place without the involvement of other agents because of its reflective capacity. The private mind is not a reflective mind and cannot be a source of reasons.

Now, it is not enough to claim that reflection makes space for error. Reflection must have standards or correctness conditions, otherwise error is once again impossible. To go wrong implies a *correct* way of going. Thus, reflective consciousness is intelligible only if it is governed by determinate norms and these are provided by an inescapable practical identity.[297] A practical identity

---

[295] (Korsgaard 1996, 136)
[296] (Korsgaard 1996, 137)
[297] Agents may have many practical identities, but there must be at least one that is not optional in order to preserve the idea of reflective consciousness. For Korsgaard, the inescapable practical identity is humanity. I will avoid discussing practical identities in too much detail because their content does not matter to the argument.

makes the agent who she is: it supplies her values, her self-understanding, and her basis for rationality. Hence, due to the social scene of the mind as reflective consciousness, the thinking and acting selves can recognise the determinate standards provided by the agent's practical identity that make rational (and irrational) behaviour possible. Reasons are public, in some sense, even in the mind of a solitary individual.

Nothing in the first strand of the argument indicates that an agent's reasons bear on other people. After all, the first strand concerns only an individual mind which is perfectly compatible with private reasons. Instead, it argued that reasons are *in some sense* public because the mind is essentially social via reflection. The second strand completes the argument by showing that reasons are public, in that they bear on other agents, because *we* are reflective agents who share practical identities. Korsgaard takes it as self-evident that we are reflective, thinking agents and that, in turn, partly makes up our practical identities as human. The role that a practical identity plays for Korsgaard is to provide a check on the endorsing function of the thinking self. For instance, the acting self proposes a desire as a maxim to the thinking self who then can endorse that maxim and make it a reason. Endorsement, though, must be guided by some reason-making criterion, otherwise, it would have no conditions of correctness and could not make reasons out of desires. Practical identities, and in particular the practical identity of humanity that all reflective, thinking agents share, constrain and guide endorsement. The thinking self reflects on her desires in light of her humanity and ought to endorse only those which are consistent with her practical identity. Public reasons must go beyond one reflective consciousness to count as properly public. Despite Korsgaard's insistence on a social mind, if the normative force of reasons did not extend beyond one mind, they would, in the context of the wider debate, count as merely private.

For reasons to be properly public, reasons must be considerations for all deliberative agents. They cannot be mere noise that can be ignored.[298] Korsgaard tells us that it is nearly impossible for us to ignore the reasons of others, that is, for their reasons not to intrude on our deliberations. The reason for this is that "the reasons of others have something like the same standing with us as our own desires and impulses do".[299] How is this possible? It must be because of the capacity for reflection, since that is the source of obligation. Because an agent cannot ignore the reason of another, she reflects on it in order to see the humanity in it. She puts herself in the other's shoes. The fact that agents share practical identities eliminates any normative gap that appeared to be between them, the shared practical identity is fundamental. When agents exchange reasons, they are allegedly deliberating together to determine a course of action consistent with their identities. They are not offering up reasons that can be ignored or heard as mere noise. When a person offers a reason to another, the other should reflect on it as if she were her and decide on whether to endorse it. The person's reason has normative force for her because she recognises the other as a human being with whom she can swap places and, so, she ought to see them as reasons just like her own. Were she in her shoes, she would endorse that reason too. By acknowledging that another shares her practical identity, she cannot help but hear her claims as reasons to consider. Because they share practical identities, the same normative standards constrain their deliberations, that is to say, the reasons that they have are public. So, reasons are public because agents share the practical identity of being human, which entails the reflective capacity to endorse the desires of any human being as reasons.

[298] Of course, there are issues about weights of reasons. Not all reasons are conclusive for each agent, but they ought to be considered. This is to say that an agent cannot go on 'just as before' she heard the reason. The reason intrudes into deliberation.
[299] (Korsgaard 1996, 140).

I do not think this argument is successful in establishing the publicity reasons because each of its strand has a problem. In the first strand, reasons are public because reasons only appear in reflective consciousness, a social scene where thinking and acting selves are subject to the same standards. I am sceptical that Korsgaard's lawgiver metaphor can establish the fact that the mind is *social*. The fact that the mind ought to be unified because it is subject to rational standards does not make it social. Korsgaard obviously does not think that the first strand establishes publicity strictly speaking because she develops the second strand. Especially in this debate, this is an abuse of the term 'public' because it is consistent with private reasons. Indeed, one might even take the first strand as an argument for the privacy of reasons. So, even though reflection may have this capacity to obligate, it does not yet show why we can be obligated by others and their reasons, or vice versa. For this to happen, the central claim of strand two must be true: that other's reasons must have the same standing as an agent's own desires in deliberation because there is not a normative gap between agents. But Korsgaard's burden is to demonstrate whether this claim follows from the fact that people share practical identities. The fact that two people have reasons that are consistent and perhaps sometimes identical might very well be explained by shared practical identities. But this does not show that the normative force of a person's reasons *necessarily* has a bearing on another.[300] Consider a case where a person has the ambition to become a pilot and her friend cannot understand why. Do their matching practical identities reveal a totally common normative world? The ambitious person explains all of her reasons for wanting to become a pilot and asks the other to put himself in her shoes. The other can acknowledge that there is reason *for her* to pursue her ambition, but why does that matter to *him*? Likewise, he may have an ambition that she does not value that she then comes to see why *he* sees value in it and yet it still does not matter to her. The fact that we can put ourselves in other's shoes when we share practical identities does not show that

[300] What I say here is similar to Joshua Gert's criticism of Korsgaard's argument (Gert 2002, 314–17).

there is no gap at all between people's normative worlds because being able to understand why another values something does not make it the case that one must find it valuable from one's own position. This is to say that we can find other people's claims about their values *intelligible*, especially when stepping into their shoes, without necessarily obligating us in our *own* shoes. The contingent fact that we often share reasons in virtue of our shared practical identities does not show that reasons are necessarily public. When we share values, we can reason together to find consistent reasons but not all of our values are shared. Even with the practical identity of being human, which we all possess, we are not necessarily obligated by the same reasons, even when they are qualitatively identical. The fact that I have *my* reasons in virtue of being human and you have *yours* in virtue of being human does not overcome the normative gap between our reasons. Though our reasons pass through the same filter, that does not entail that we are obligated by the numerically same reasons.

It is important to remember that Korsgaard should not appeal to consistency to make the argument work. An agent is not obliged to consider another's reasons because consistency demands it. This would make the argument another ill-fated neo-Kantian attempt. But without an appeal to consistency via shared practical identities, it is difficult to see why an agent must conform to another's reasons in her own shoes. It must be the case that the other's reasons are consistent with *her own* reasons issued by her *her own* practical identity. Now, Korsgaard thinks that consistency does not do the heavy lifting because the normative gap was always already closed.[301] And the closure depends on our sharing practical identities. Having matching or distributed practical identities does not establish that we are subject to the numerically same reasons even if they are qualitatively alike. For instance, we might each have reasons to house our own children, but these are distinct reasons. Having matching practical identities might provide some public reasons but it cannot guarantee that all reasons are public. If practical identities are like tokens of a type that each of us possess, then it

[301] (Korsgaard 1996, 143)

seems that reflection via practical identities will not always issue the very same reasons. But assuming that practical identities 'match' might beg the question against Korsgaard by introducing a gap. So, imagine instead that we have a numerically identical practical identity, there would be no gap between us since we would share the very same practical identity. But this raises the question: what could make it *my* practical identity, the thing that unifies *my* mind and agency, the thing that makes me *me*? The fact that this question arises seems to dissociate the agent from her inescapable practical identity. If practical identities are not like tokens of a type that make each individual the agent that she is in virtue of possessing them, then it seems that the inescapability of a practical identity is unsettled and cannot issue the same reasons to everyone who shares in that identity. For the gap to have already been closed, I worry that Korsgaard has in fact assumed public reasons instead of providing an argument for publicity.

Moreover, Korsgaard insisted that sociality played a crucial role in the argument, but it is hard to see what is social about practical identities. In one sense, my practical identity is mine, even if others have a matching identity. And if it is not mine in this individual sense and, instead, we share it in some social sense, then it is hard to see what could make it properly *mine*. Once again, the social is not properly take advantage of in the second strand of the argument. Nevertheless, Korsgaard, I believe, is right to point to our social nature as a component in the argument for the publicity of reasons. The problem is that she finds sociality and publicity in an individual and builds up from there, a similar problem she pointed out for the Hobbesian and neo-Kantian.[302] Instead of an individual's capacity for reflection being the source of error and normativity, Korsgaard should have shown how others are necessary for normativity from the first moment and appealing to shared identities does not do that. The private reasons argument may have been better served by

[302] This is O'Day's main complaint against Korsgaard (1998). His solution, a social contract for rationality, does not help either, as Norman observes (2000), since social contracts require rationality to get their results.

taking our actual engagement with others to be the place where error and normativity appear. If others are necessary from the beginning, then the normative gap would be a mirage. If error can be shown to be a social product, then perhaps these private, yet intelligible, reasons do not appear.

3. Publicity and the Triangulation Argument

Near the end of his career, Davidson became more and more concerned about practical thought and the foundations of ethics. He believed that there are grounds for a naturalist *and* non-revisionary version of normative realism. It would be naturalist in the sense that normativity is part of the natural world and non-revisionary in the sense that normativity retains its critical edge without reducing it to the non-normative.[303] Davidson thought that his triangulation argument supplied such a version of normative realism, one that avoids the pitfalls of many versions of realism. My intention is to show that this argument also supplies an account of unrestricted normative publicity and the existence of group reasons. Davidson originally developed the triangulation argument for the objectivity of language and thought but then came to see its implications for practical thought.[304] The goal of the triangulation argument was to provide necessary conditions for an individual attaining thought, a first language, and the concept of objectivity, error, and belief (these concepts are inextricably linked for Davidson). It aims to show that for this to be possible, language and thought must in part depend on a public world. The contents of thought and talk are fixed or determined by two individuals in a triangular relation with a common cause: each recognisably relates to the cause and to the other. With respect to language and thought, the triangulation argument only gets us to publicity as intelligibility. But unlike Korsgaard's argument, the world and the agents' interaction play essential roles in thought and talk in general: we can think something

[303] It is very common that normative naturalism is also revisionary, e.g., (Sturgeon 1985; Railton 1986; Boyd 1988). This is not a necessary compromise as Davidson aims to shows. Neo-Aristotelians also deny the compromise, e.g., (Hursthouse 1999; Foot 2001; Thompson 2008).

[304] This argument is developed in a variety of papers, the most important of which are (Davidson [1992] 2001; [1997] 2001; [1990] 2001; 1994; 2001a).

determinate because a part of the world helps cause *us* to have a concept about it. So, how did we gain our normative concepts? By being influenced by normative properties in triangulation, or so the argument goes. This argument establishes normative publicity because the (normative) world must be shared by agents, so long as the triangulation argument is right.[305] Normative properties bear on each triangulator because each of them is in a position to appreciate the property only if both of them can do so. The structure of this section begins by explaining why triangulation is necessary for thought and talk in general and then it takes practical thought as an instance of the general argument.[306] It then draws implications for group reasons and spontaneous group action.

Let's start by explaining the goals of the triangulation argument, to explain objectivity of thought and determinate content. The idea of objectivity entails that there is a contrast between truth and falsity, seeming to be the case and being the case, correct and incorrect. For a thing to be objective, there is a way for it to be correct or incorrect. If a belief is objective, it is possible for it to be true or false. What this means is that there is a standard to which thoughts, sentences, intentions are accountable in order to count as objective. Objectivity entails the idea of 'trivial' normativity, the idea that a thing is correct or incorrect.[307] Without a standard to determine the correctness of, for example, a belief, we are left with a mere disposition which is neither correct nor incorrect. The disposition just is what it is. Objectivity here is not to be contrasted with subjective in the sense that false beliefs are subjective because they do not conform to reality. For Davidson, a thought's objectivity depends on the fact that it must be either true or false and its truth typically depends on

[305] One might worry about 'conceptual schemes' blocking any account of publicity. One consequence of the triangulation argument is that the possibility of conceptual schemes is undercut.
[306] I am greatly indebted here to Claudine Verheggen's interpretation of the triangulation argument as well as Robert Myers' application of it to practical thought (Myers and Verheggen 2016; Verheggen 2006; 2007). The triangulation argument has received many different interpretations, the best of which is, I believe, Verheggen's (Child 2001; Montminy 2003; Sinclair 2005; Bridges 2006; Amoretti and Preyer 2011). Much of what follows in this section derives from their characterisation of the triangulation argument.
[307] Verheggen distinguishes trivial normativity from hypothetical and categorical normativity which involve 'oughts'. Trivial normativity merely says that there is a standard for judging the correctness of a thing (Myers and Verheggen 2016, chap. 2).

something external to us.[308] We will see that attaining, as Davidson calls it, the concept of objectivity is quite difficult because it depends on attaining thoughts with determinate contents.[309] Thoughts have determinate contents and those contents, *p*, are true or false depending on whether it is the case that *p*. Only when thought has determinate content can it be evaluated as correct or incorrect. Moreover, thought is holistic, so to identify any one thought by its content is to locate it within a pattern of beliefs that use the same concepts.[310] In light of these patterns, the truth conditions of the agent's concepts, that is, when they are correctly or incorrectly applied, is manifest. And once the truth-conditions for an agent's concepts are clear, then the truth conditions for any thought can be given because concepts are the basic units, in addition to logical operators, of (propositional) content. What it means for a thought to have determinate contents is for its contents to have truth conditions, that is, a standard for its correctness. Now, of course agents are not consciously aware of all of these patterns because the patterns are infinite but, in thinking, they avail themselves of these patterns.[311] Further, it is important to note that the content of thought is not extensional. The fact that two people refer to the same thing with words does not show that they *mean* the same thing by their words; thought and meaning are intensional. This is important to the triangulation argument because even if creatures are in a position to refer to the same thing, e.g., by ostension, it does not amount to their *applying concepts* to that thing. How could they tell if one has gone wrong?

To achieve objective thought, then, a triangular situation is necessary. Consider Wittgenstein's case of an individual thinking about a green sample, where the meaning of a thought

---

[308] 'Typically' here leaves open the rare cases of non-observational thoughts such as self-knowledge and practical knowledge. Davidson occasionally notes that not all objective thoughts must *have* truth-conditions because we have not put much thought into what would make it true (Davidson 2004b, pt. Apdx). This doesn't entail that we make the truth up—instead, we don't yet have an interest in figuring out what we *mean*.

[309] One might already be worried about circularity here. This concern will be addressed at the end of the section.

[310] I want to flag now the importance of holism for the normative aspect of the triangulation argument. The fact that concepts have a role in many different sentences will be crucial to the argument.

[311] The patterns are infinite due to the simple logical operators and compositionality. Even though there will be plenty of redundancy, 'X and Y' is a different thought than 'X' and 'Y', as is 'X and Y and X'.

is determined internally.[312] What exactly is she thinking when she thinks about a green sample? Does it represent *green* in general? Maybe the purest *green*? Say she is in fact thinking about the sample's shape. The point is that the agent will be thinking whatever she takes herself to think, there is no right or wrong when thinking about the sample because there is no standard. Say she checks herself a moment later and thinks 'No, I meant that this is the purest green'. The question now is: what did she in fact mean when, at one time, she says that she meant the shape and, at another, a shade? We have no reason to credit her with a determinate thought because there is no way to judge that she is incorrect at any one time. Internally determined thoughts are indeterminate.[313] If an individual cannot determine her thoughts internally, perhaps she can do it externally, with the help of the world and her perceptual capacities. Say that she thinks 'there is a table' because she sees a table. The external world causes her to think a determinate thought. But once again, even though the world appears to offer a standard of correctness, an individual's perception cannot help her fix the contents of her thoughts on its own. As with the internalist, the solitary perceiver will not be able to settle the question about the contents of her thoughts. When she sees a table and thinks 'there is a table', on the basis of what does she think 'there is a table' and not 'there is a table and chairs', 'that's an interesting grain', 'there is a brown varnish', etc.. On her own, what she thinks is what she says that she thinks at any one time, and this is no standard at all. She cannot be wrong, and what appears to be thought turns out to be merely dispositional.[314] Even though perception may provide the individual with the material for thought, she has no way to determine the meaning of her thoughts on her own.

---

[312] (Wittgenstein 2009, sec. 73). Wittgenstein prods through many different accounts of internalism and they all come up short for the same reasons.

[313] This is the general thrust of the argument, but it goes further to suggest that even if somehow the indeterminacy can be eliminated, the thought will lead to an infinite regress. The regress reveals that the determination was an illusion.

[314] For a more detailed account of the argument against solitary perception, see (Verheggen 1997).

What is needed, according to Davidson, is the introduction of a second person who can provide the discrepancy required for the acquisition of the concept of error or objectivity. The problem with individualism in both its internal and external forms was that it could not make sense of thinking or speaking incorrectly. The second person's introduction makes space for agents to recognise that they could be wrong by allowing them to fix a common cause of their thoughts and utterances. This triangular arrangement takes two forms, primitive and linguistic. Primitive triangulation involves non-linguistic individuals, such as lions chasing a gazelle, and linguistic triangulation, naturally, involves individuals with a first language. The point of primitive triangulation for Davidson is to fix the distal cause, which could be done without language. Without another person, an individual cannot determine whether a cause is on the tips of her nerve-endings or the big bang. The problem of fixing the distal cause though has been effectively dismantled. As Jason Bridges points out, our perceptual systems are capable of identifying distal causes without another agent.[315] Nevertheless, primitive triangulation still provides a *necessary,* though not sufficient, condition for the possession of objective thought. Without the second person and the world, there would be no hope of determining the contents of thoughts.

Primitive triangulation, then, begins with two individuals who are being caused to perceive the same object. One individual responds in a way to the cause as does the other. From each individual to the cause, there is a line that constitutes a side of the triangle. Each responds similarly to the cause when she is caused to react to it. For instance, whenever the lion hears the gazelle move through the tall grass, she perks up her ears. As indicated above, an individual's perceptual system allows her to locate medium-sized objects without engaging with another, so each has a good idea where the object is that is causing them to react. But on her own, each individual cannot take advantage of the fact that she responds to similar causes in similar ways such that she categorises the

[315] (Bridges 2006, 307–10). Verheggen acknowledges Bridges' point (2016, 19).

cause as one thing rather than another. The cause merely disposes her to respond in a particular way that is not recognisable as the *same* response to previous occurrences of the cause. Though, things change when the two individuals interact, forming the base of the triangle. Imagine that each is reacting similarly to a cause and now each sees the other reacting similarly to the cause. Each is able to attend to the cause and her triangulator who responds in a similar manner as her. Again and again, the primitive triangulators react similarly to a similar cause and they can see that about each other.

Consequently, there are three different similarity patterns in the triangle. If one of the triangulators does not respond similarly to the cause, the triangular arrangement makes space for the recognition of this discrepancy. However, the space for recognising discrepancies does not mean that primitive triangulators in fact *recognise* the discrepancy or error. Their problem is that they cannot take "cognitive advantage" of their triangular arrangement because there is still an ambiguity that they cannot resolve.[316] Even though the individuals may see that they react similarly to a similar cause, the cause is not yet a *common* cause between them. They cannot fix the relevant aspect of the similar cause to which they are responding. Is it the gazelle's loud chewing or perhaps it is the grass breaking as she walks through it? *We* may very well *think* that the lions are focused on the chewing. But this is because we are thinkers who apply our concepts. What could make it the case that the lions can distinguish which aspect of the gazelle is causing them to respond in similar ways? Having a common cause requires removing the aspectual ambiguity of the cause because an ambiguity does not settle what the common cause is. So, despite the fact that the triangular arrangement allows individuals to see another responding similarly to a similar cause and makes space for recognising discrepancies in response to the cause, it does not allow primitive triangulators to distinguish the

[316] (Davidson [1992] 2001, 120)

aspect that is the cause of their perhaps exactly similar responses. After all, acting in a way that accords with a rule does not show that one is following it.

Being able to fix the aspect to which one is responding is crucial to meeting the conditions of objective thought. If one lion perks up to the sound of breaking grass and another does not, this does not provide the resources for thinking *there's a gazelle* versus *there's an antelope*. Objective thought requires that agents be able to say of a thing that it is one way and not another—this is what it is and not that. When this occurs, error is possible: there is a gap between the way an agent responds and what caused them to respond. The response can be incorrect. The only thing that is fine-grained enough to fix the aspect of the cause is language. Having a language is sufficient for having objective thought and a language. Now, this seems clearly circular, however, having a language depends on triangulation, that necessary arrangement that makes space for agents to possess the concept of objectivity, error, and belief. We already saw how an individual alone will fail to acquire this critical gap between her 'thoughts' and the external world. When a first language is already possessed by the triangulators, then the common cause and the meaning of their thoughts and utterances can be fixed. The argument is non-reductive in that it does not aim to give necessary and sufficient conditions, although it does give necessary conditions and that is why it is not viciously circular. This is what might be expected from language and thought because the vocabulary of language and thought cannot be reduced to an extensional vocabulary. Davidson often remarks that, to *us* who have thought and language, it seems as if we will never be able to fill in the gap between the extensional and the intensional, judgment and disposition.[317]

Fixing the meaning of a thought requires the possession of the concept of objectivity, which in turn requires linguistic triangulation. Now, both triangulators have a language and interact with each other and the world. Why does the addition of language allow us to credit them with the

[317] E.g., (Davidson 2001a, 13).

concept of objectivity, determinate thought, and the awareness of the possibility of error? Language permits the fine-grained determination that was lacking in primitive triangulation. Instead of lions, imagine two individuals who hear a sound in the tall grass. Both respond to the sound claiming that 'there is a gazelle' and both of them see that the other responds similarly. But when triangulators agree in response, in the basic case, they still cannot be credited with the concept of objectivity because they still do not have a standard against which to compare their uniform response. They do not yet have a contrast between what they say and what is the case. However, once one triangulator responds differently that contrast "forces" the concept of objectivity upon them.[318] When one triangulator says 'there is a gazelle' and the other says 'there is an antelope', the divergent responses makes each agent aware that her response may be incorrect. The possibility of error is now on the scene. Error was not available to the primitive triangulators because they could not fix the aspect of their cause. But in linguistic triangulation, the aspect can be fixed and the common cause of their thoughts and utterances discerned. With divergent responses, one triangulator can ask the other why she thinks it is an antelope and vice versa. Because each is a competent language user, each can say what exactly made them think it was one thing rather than another. They can investigate what exactly caused them to respond as they did—'it is walking too heavily to be a gazelle', 'let's go look for tracks', etc.. In the basic case, linguistic triangulators developed concepts for the causes of their responses and, in the presence of a cause, these concepts can be correctly applied. The extension of an agent's concept will depend on her history of triangulating with similar causes. What is crucial is that the words used to refer to a cause, in the basic case, are "endowed with a meaning"—it is correct to use this word for this cause and incorrect for a non-similar cause.[319] With thought on the table, triangulators may both be incorrect about whether it was a gazelle or an antelope, but what is

[318] (Davidson [1975] 2001, 170)
[319] (Davidson 2001a, 14)

important is that they are aware that they can apply their concepts incorrectly—they have taken cognitive advantage of the triangular arrangement. To have the concept of objectivity requires that an agent has language, which in turn requires that her meanings are fixed, which in turn requires that she is aware that she may misapply a concept and be in error, and to be aware of the possibility of error requires that an agent has the concept of objectivity. This is a circular affair, but each of these concepts is needed to make sense of the others and in fact they all depend on the same thing: standards of correctness. How agents are aware of standards of correctness depends on linguistic triangulation: in the process of fixing their meanings through triangulation, agents eventually realise that they may be wrong because of divergent responses and, with that, they understand that the standard that makes their thought and talk correct is independent of what they say it is.[320] What is achieved by triangulation is this idea: what is true is independent of what I think and, because of this, the contents of thought can be determined.

Before moving on to practical thought, it is important to stress that the meanings of our words are not grounded in mere agreements or social sharing between individuals.[321] They could not establish standards of correctness without the world's involvement. Whatever agents in such a scenario would be 'thinking' would be unmoored. We would be left wondering how it is that their words, in a case of understanding, refer to the same things. What agents mean by their words depends on their being caused by properties, objects, and events in order to establish correct literal usage. What there is to know about meaning is publicly accessible: engaging with another and the world is all that is required. There is no introspective private place in the mind, instead agents are intelligible. But this result does not yet reach the conclusion that reasons are normatively public

[320] This may make it seem as if we mechanically adopt meanings, but, agents do have a role to play in what they mean by their words and which words are endowed with which meaning.
[321] Davidson does like to use the phrase 'social sharing' but he does so cautiously because it tends to be understood as agreement without the world's contribution.

because it merely establishes publicity as intelligibility. To do this, we must think more about the holistic character of the mind and how it relates to triangulation. In order to fix the aspect of our causes depends on the triangulators' ability to disambiguate one aspect from another, which in turn depends on their ability to use language to specify what the cause is. For a triangulator to mean *there is a table* by 'there is a table', she must be able to mean, for instance, *not the table and chair set*. The concepts used to mean any one thing are always related to other concepts because concepts are used in an infinite amount of other sentences and thoughts. So, to be able to say anything (determinate) about any one thing is to be able to make an infinite amount of other (determinate) statements. Fixing the causes that provide the standards for our meaningful utterances and thoughts depends on the holism of content, that is, the mental. But it is not the case that beliefs are the only contentful attitudes. The mind is comprised of other attitudes as well, such as intentions, desires, hopes, suppositions, etc. And, further, it is not the case that beliefs have merely descriptive content, they can have normative content too, such as when an agent believes that she ought to be on time for her appointment.[322] 'Ought' has as clear of a meaning as 'time' and 'appointment' in this belief, and it constitutes a portion of an infinite amount of thoughts an agent may have. Just as the meaning of descriptive concepts are holistically determined, so are normative ones. So, how can an agent have the ability to use normative concepts and how do they occupy a place in our holistic conceptual repertoire?

Davidson's answer is: in the same way that every other concept is acquired, by triangulating on public properties. However, in this case, those properties are normative. Imagine a teacher and a learner watching some kids light a cat on fire for fun.[323] The learner does not have the concept

[322] Naturally, not all philosophers, especially non-cognitivists, will agree on this point, but Davidson is careful to attend to their concerns (Davidson 2004b). He does not find the motivation for their position to be very convincing and, indeed, he believes that the triangulation argument avoids non-cognitivists concerns while doing justice to our normative experience.

[323] This is Harman's famous example tailored to a triangular arrangement (Harman 1977, chap. 1).

'wrong' in a normative sense. So, the teacher says to her pointing, 'look how wrong that is, they are setting an innocent cat on fire for fun'. With this indication, the learner can focus on the event of setting the innocent cat on fire for fun rather than the mugging occurring on the next corner over. The teacher is not merely indicating for the learner to focus on the event, but to focus on a property of that event, *wrongness*. And the teacher says again, 'that is wrong', and at some point the learner grasps the concept of wrongness as being applied correctly in this case. From now on, she can use the word correctly or incorrectly, but she is aware of the fact that there is a correct usage. The way agents acquire normative concepts is by triangulating on public normative properties that exist in the natural world—they are present in events that cause agents to respond to them. This argument establishes that the normative property, let's assume that it is a reason, is a real feature of the world, external to the minds of those triangulating on it. Because agents cannot have a concept without being caused to have it in triangular arrangements, normative concepts must be acquired by being caused by real normative properties in triangular arrangements. This is why Davidson is a normative realist: because of the externalist nature of the triangulation argument, reasons and values exist independently of our sensibility and our minds. In summary, once it is conceded that triangulators acquire concepts and are able to have determinate thoughts by triangulation and the meaning of those thoughts is holistically determined, then it must be acknowledged that normative concepts acquired and normative beliefs had are the result of triangulation. The move from descriptive concepts and thoughts to normative ones depends on the holistic character of content because normative concepts are holistically entangled with descriptive ones and vice versa. The holism of the mental forbids cleaving the normative from the descriptive.

As I said at the beginning of this section, the triangulation argument can vindicate a robust, naturalist normative realism. It is a naturalist realism because normative properties are part of the natural world and *cause* us to have normative beliefs under the right conditions. One might worry

that saying that the normative is causal is a conceptual blunder, i.e., the normative is not causal, it is rational. But we must wonder how we have normative beliefs. Non-naturalist epistemology typically appeals to intuition, but we should be suspicious about intuition, in particular, because it is a mystery how we could have a determinate normative thought.[324] It is not dissimilar to Wittgenstein's attacks on semantic internalism. Instead, one might think that normative beliefs get their content by being, in some sense, a rational reflection of an agent's desires.[325] Now, how normative beliefs acquire content depends on the theory of desire. But the most credible account of desire in fact inverts the relation between normative beliefs and desires that these reflection accounts require. Instead of desires being sensitive to descriptive beliefs, the better account claims that they are in fact sensitive to normative beliefs.[326] If desires are in fact sensitive to normative beliefs, then it cannot be the case that normative beliefs depend on desires for their content. Rather, desires get their content from normative beliefs. Even if triangulation provides an adequate account of the epistemology of reasons, there is still the worry about how norms could cause. The idea here is that if norms are causal, then there must be a law of nature that explains their causality, and now it appears that the normative in some way causes us to act rather than telling us that we *ought* to act. But Davidson famously held that the fact that there is a causal relation between two events or states does not imply that the causality appears in every description of the two events and their relation.[327] For instance, the neurologist provides a true non-intensional explanation of why a person raised her hand, i.e., some neurons fired plus physiological effect, and yet there can be a true intensional description of her movement lacking merely causal vocabulary, i.e., a rationalisation of her

[324] Some examples of intuition realists are (Shafer-Landau 2003; Enoch 2011). Scanlon interestingly does not appeal to intuition but instead to reflective equilibrium (Scanlon 2014). But why should we have any confidence in reflective equilibrium if we have not established that any of the input judgments are true? Having a more consistent set of beliefs does not make them true.
[325] For instance, (Smith 1995).
[326] This position has been argued for by Robert Myers (Myers 2012; Myers and Verheggen 2016, chap. 6).
[327] (Davidson 2001d)

intentionally raising her hand. Because the intensional is anomalous does not mean that it is wholly divorced from the nomological because there is more than one way to describe the very same events. In fact, the nomological is merely one other way of describing events. So, we can understand normative properties as being causal and also, under intensional descriptions, as being robustly normative, that is, as making claim on us without causally necessitating us to act. Furthermore, Davidson can champion a non-revisionary account of normativity since the normative is not characterised in non-normative terms. The triangulation argument does not force us to abandon robust normativity in order to reconcile normativity with the natural world.[328]

But what does the triangulation argument offer us in terms of publicity? It is true that it vindicates the publicity of meaning, i.e., as intelligibility. But the goal of this chapter is to vindicate normative publicity, in particular unrestricted normative publicity. There are two interrelated questions that need answering. First, are group reasons unrestrictedly or restrictedly public? Recall that unrestricted publicity is the thesis according to which all reasons apply to everyone in some manner. Whereas restricted publicity is the thesis according to which reasons apply to more than one person but not to all. The answer to this question depends on the second question which is whether the triangulation argument can make sense of public agent-relative reasons. The connection between these two questions is that the distinction between unrestricted and restricted publicity is structurally similar to the differing interpretations of agent-relativity by Nagel/Korsgaard (restricted) and Wallace (unrestricted). There are three possible answers. First, group reasons are unrestrictedly public because there are only agent-neutral reasons. Second, group reasons are restrictedly public because agent-relative reasons only apply to those picked out by the reason. And,

[328] There was a worry in Chapter 3 regarding the reliance on normative reasons. When co-agents are acting on a false normative reason, why should they still be considered as acting together? There, I appealed to background normativity to help with the worry. It should be clear that, as agents, we experience the world as normatively imbued. It is part of what it means to have a mind.

finally, group reasons are unrestrictedly public because the normative force of agent-relative reasons apply to everyone in some way. Let's start with the first option.

I think it is clear that the triangulation argument establishes some form of the publicity of reasons because the normative property that anchors triangulators has to be accessible to both of them *as normative*. They would not be triangulating on a *normative* property if it only had normative force for one of them. If a normative property were only available to one person, then that property would not be able to determine the content of the person's thought. Hence, so long as there are normative properties that determine the contents of triangulators' thoughts, then the normative force of those properties must be public to those who triangulate on them. Further, at this general level, we cannot discriminate between who the triangulators are, whether they have a special relation that produces an agent-relative reason or not. In principle, normative properties are triangulatable by *anyone* coupled with another. For instance, a teacher is teaching a learner that setting the cat on fire for fun is wrong. The wrongness inhering in that event is equally available to anyone, no matter their relation to each other. The fact that normative properties qua causal property are non-exclusive shows that their normative force is for everyone, and this means that reasons are unrestrictedly public.[329] This argument might appear to establish that reasons are agent-neutral, since their normative force extends to everyone. This depends on whether there is a serviceable understanding of public agent-relative reasons. It might be the case that there are no agent-relative reasons because triangulation rules them out. In that case, there would only be agent-neutral reasons that are unrestrictedly public.

Is it possible for there to be agent-relative reasons whose force extends only to those individuals picked out by the reason? In this case, reasons, if they apply to more than one person

[329] Assuming that normative properties are reasons. As we will see shortly, maintaining this becomes complicated. Nevertheless, normative properties are involved in reasons, what they are exactly will be left for another time.

such as parents, would be necessarily restrictedly public because not all reasons would be unrestrictedly public. For instance, parents having developed a special relation to their daughter obliges them in agent-relative ways where it does not oblige others in the same way. This is the route that Robert Myers takes.[330] He notes that realists are often pushed into denying the existence of non-derivative agent relative reasons because it is difficult to understand how agent-relative reasons could exist if we all share normative properties. He suggests that "agent-relative reasons simply grow out of agent-neutral ones as special relationships are forged".[331] The idea is that there are reasons to care about the protection of children generally, but as parents forge a relationship with their daughter their special, agent-relative reasons arise to protect their own child.[332] In this case, the parents share a reason to protect their daughter that is normatively public for them while its force does not extend to others. Myers acknowledges that others can appreciate people's agent-relative reasons based on special relationships that they have: "the special importance that a child has for her parents is a property of the situation that everybody can appreciate".[333] What exactly are others appreciating in such situations? Is it the same property as the parents are recognising? It must be, otherwise it is committed to non-public normative properties and this cannot follow from the triangulation argument.[334] What else could explain the lack of extended normative force? Perhaps it is because there is an indexical involved in the property that excludes others. The normative property is normative for the person(s) to which it is indexed and merely intelligible for the others. But we are left to wonder how triangulation works if this is the case. When triangulating on 'agent-relative' normative properties, the normative property would be normative

[330] (Myers and Verheggen 2016, 179–83).
[331] (Myers and Verheggen 2016, 182).
[332] Agent-relative reasons are not *derived* from agent-neutral ones because their force does not depend on the more basic agent-neutral reasons. Nevertheless, the reason to develop a special relationship that produces agent-relative reasons depends on acknowledging the force of the basic agent-neutral reason.
[333] (Myers and Verheggen 2016, 181).
[334] I don't want to say that these are private normative properties because at least two people have access to it. But they are not public in the sense required by the triangulation argument.

for one person and not normative for the other. Why would the other consider it a normative property if it has a radically different effect on her, that is, a non-normative effect? Even if we assume that the normative property that grounds the agent-neutral value is the very same as the one that grounds the agent-relative value, it seems to get our normative experience wrong. It is not merely that I recognise that the parents are in a better position to protect the child and, hence, that the agent-neutral value is best realised by them. If this were true, then the agent-neutral value could be used to trump the agent-relative value when the agent-neutral value is not being or could be better realised. This is where normative experience tells against this account: agent-relative value does not (always) serve agent-neutral value. Often, agent-relative value carves out an inviolable domain that is protected from the non-partial commands of agent-neutrality. This counts against merely intelligible agent-relative reasons. Can we really draw the line between intelligibility and normativity so easily? Recall that reasons were won by the triangulation argument because we apply these concepts together to the same normative properties. So why would a normative property lose its normativity for anyone who is not specially related to the circumstances of that property? The triangulation argument speaks against restrictedly public agent-relative reasons because they depend on conflating senses of publicity.

This might seem to suggest that agent-relative reasons are doomed, but perhaps unrestrictedly public agent-relative reasons can be made intelligible. The aim here is to make sense of Wallace's interpretation of publicity. The first step is to shift the focus from reasons to considerations in that the consideration is what is primarily normative.[335] Reasons are defined as considerations that count in favour of φing. But the trouble with reasons in this context is that their

[335] There is a problem here of how to understand the identification of reasons with considerations. The same consideration at the same time can count for and against the same action depending on two agents' circumstances. In one sense, they have different reasons, but the very same consideration is conclusive for them. Does the normativity of a consideration necessarily include the 'that counts in favour' clause or is the consideration normative and what it counts in favour of can vary based on the circumstances? My aim is to suggest the latter.

specific command speaks to those who are picked out by them. For instance, the agent-relative reason to protect their child is a reason *for* the parents. Reasons are relations between particular persons, a consideration, the circumstances, and the action to be performed. This relation excludes anyone who is not mentioned in it and she is not commanded to act *in the way* the reason commands. But if considerations are the basic normative entity and not the reason as a whole, the normative force of the consideration can extend to different people in different ways depending on the circumstances. Wallace appears to have this conception in mind: "if considerations C provide me with reason to do X, then they equally provide other people with corresponding reasons for action".[336] This shift towards (normative) considerations should not seem too queer because considerations differ from non-normative facts in that they partly comprise reasons. For instance, the fact that the sky is green is nothing more than an interesting sight *for me,* but to a scientist, it may be a consideration to write a report about pollution. So, a fact's being a consideration already reveals that it is in the normative realm as opposed, if you like, to the descriptive realm. Now, if the normative property on which agents triangulate 'attaches' to the consideration, then it is possible that unrestricted publicity is consistent with triangulation. This is because the consideration taken on its own does not speak in favour or disfavour of anything in particular. More is required for that to be the case, in particular, the other constituents of the reason relation. By taking the triangulators' circumstances into consideration, their reasons come into view. Even though they triangulate on the same normative property and considerations, the reasons provided by them may be different. To give an example: the triangulators notice that a child lacks workbooks for school, however one of them is her parent. This situation clearly provides reason for the parent to provide school materials for her child, but it is also the case that other triangulator may have reason to persuade her partner to buy school materials, not interfere with their purchasing, etc. The consideration that the child

[336] (Wallace 2009, 471)

does not have adequate materials to learn provides them both with reasons and for one of them it happens to be an agent-relative reason.[337] In this way, the requirement of publicity involved in triangulation is respected and room is made for the unrestricted publicity of group reasons. Every consideration is normative for every agent, but how it figures in reasons for each agent depends on their circumstances. Perhaps it is a misnomer to call it the 'publicity of reasons' but, since considerations play an integral role in the constitution of reasons, it is not a stretch to think that considerations bear normative properties instead of reasons. The triangulation argument vindicates unrestricted publicity because it makes sense of Wallace's interpretation of agent-relativity.

4. Triangulation, Group Reasons, and Spontaneity

One might worry that some heavy machinery, located far afield from group action, was produced to explain spontaneous group actions. But there is a close connection between triangulation and the theory of group action proposed in Chapter 3, namely, that triangulation only makes sense on the supposition that there are group reasons to triangulate. One question worth asking about triangulation is: why do people engage in triangulation at all? One way to answer this question is to explain their motivation for the action. Without appropriate motivation, agents would not undertake the action of triangulation. So, how do they become motivated to triangulate? Davidson has a particular account of desire that results from the triangulation argument, that is, desires are judgment-sensitive attitudes. In Chapter 3, I discussed the nature of Davidsonian desires. To reiterate, agents form normative beliefs which then condition the content of their desires. It is in the nature of desire to adopt the content of an agent's normative beliefs and it is in the nature of normative belief to condition desire. What allows desire to fulfill its nature is the fact that it is guided by a systemic aim. Desires are guided by the systemic aim to get normative matters right (in action). Because desires are guided by the aim to act on the good in light of normative truth, they

[337] Note that the strength of the reason for each person may vary, but the point is that it is the very same reason.

are sensitive to the reasons agents believe that they have. When an agent discovers that her normative beliefs are wrong, her desires *ought* to change as well. Her desires are under normative pressure in virtue of their systemic aim to conform to her normative beliefs. In order to rationalise the action of triangulation, agents must be appropriately motivated.

But it should be clear from this account of desire that agents must have normative beliefs about what they ought to do in order to have desires with the appropriate content. In order to have a desire to triangulate, agents must judge that they have reason to triangulate. But this reason is not a reason for each of them alone, since triangulation cannot occur without another equally suited to the task of triangulating. An agent cannot triangulate alone and in virtue of that she cannot respond to a reason to triangulate on her own. This simple fact shows that, at bottom, triangulation depends on a *group* reason to triangulate. If triangulation requires another agent in order to perform the action, that is, triangulation is an action that must be performed together with another, then they must be triangulating on a group reason to have the desire to triangulate. To rationalise the act of triangulation requires that triangulators respond to a group reason to triangulate, otherwise their motivations would not have the content that they do. At this very basic level, for their desires to have these fixed contents requires that they in fact respond to a group reason, that is, they are caused to think that they ought to triangulate. This is an interesting conclusion in its own right. But seeing as group reasons are the focal point of the rationalisation of triangulation, it is not ad hoc to introduce this Davidsonian framework in order to illuminate spontaneous group actions.

Nevertheless, one might wonder whether it is legitimate to employ the triangulation argument in the service of explaining spontaneous group actions. After all, triangulation depends on the *interaction* of the triangulators with the world, and spontaneous group action by definition lacks interaction on the parts of the co-agents. Something appears to have gone very wrong if the explanandum requires the very thing it is supposed to lack. At first glance, this appears to be a

serious and damning charge. Yet, I think that the non-reductive nature of triangulation provides an answer. Non-reductive theories aim to show how their essential concepts depend on each other and to illuminate their natures in light of each other. In the basic case of triangulation, how it happens at all, the intensional feature of the triangulator's attitudes is not explained by the causal, non-intensional element of their interaction with each other and the environment. Such casual relations cannot explain intentionality, and recall that meaning must already be on the scene to explain the other essential elements of triangulation, such as judgment, truth-conditions, etc.. Even if there is some form of (non-intensional) interaction between triangulators, what is important is whether that interaction is what legitimates occupying the co-agential perspective. But the only thing that could legitimate occupying that perspective is a group reason. The simple fact of looking at another and the environment, although an interaction, does not warrant the co-agential perspective. This would conflate normative, intensional vocabularies with a non-intensional vocabulary. The trouble with this objection is that it demands that the triangulation argument goes reductive when it is inherently non-reductive. In the very basic case of rationalising triangulation, what explains the action is the triangulators' response to a group reason—their interacting does not explain anything in particular.

5. Conclusion

The triangulation argument in one cumbersome, fell swoop achieves both goals of this chapter: to defend the realism and the (unrestricted) publicity of reasons. It does this showing that objective normative thought depends on the existence of considerations that are equally normative and public for everyone. It was argued that the particular reason provided by a public consideration depends on the circumstances of the individuals, and this move permits an understanding of public agent-relative reasons. What this amounts to in the group context is that group reasons are unrestrictedly public, that is, their grounding consideration provides reason to all others. Agent-relative reasons and group reasons do not merely command those to which they refer. Moreover, the triangulation

argument provides some support to the existence of group reasons and not merely singular reasons because the rationalisation of triangulation depends on the group reason to triangulate.

CONCLUSION

The aim of this dissertation was to provide an account of spontaneous joint action, which as it turns out is a requirement for any theory of joint agency. It is required because spontaneity makes good on the other essential conceptions involved in joint agency, e.g., togetherness, rational agency, being partnered. If we fail to meet the spontaneity requirement, one of these essential conceptions must be revised. This web of ideas implicated by joint agency resulted in the challenge of spontaneity: how can agents both act together and do so spontaneously? To meet the challenge, I argued for the reasons account of joint action, which appeals to normative reasons for joint action, or as I called them group reasons. Group reasons are reasons for more than one agent to act together and they are best construed as *real* reasons, which are mind-independent and normatively public. Their mind-independence allows co-agents to circumvent the exploitation of their fellow co-agents as epistemic resources that justify joint thinking and acting. Real reasons warrant the occupation of the co-agential perspective. Moreover, the fact that group reasons are normatively public and tell agents to act together explains how agents are bound together. For agents to act together spontaneously according to the reasons account, they must respond to the group reasons that they share. Responding to their group reasons allows co-agents to think and act together—they think and act as partners when their agency is structured by their group reason. The appeal to real group reasons was met with scepticism, so it had to be argued that reasons exist, that *group* reasons exist, and that reasons are normatively public, in an unrestricted sense. This was done by drawing out some implications of Davidson's triangulation argument. In order to make sense of normative thought and talk, agents must triangulate on normative properties in the world. This establishes that reasons exist. Moreover, triangulation is an action that agents perform together, so, at least, in the first instance, agents must be responding to a group reason to triangulate together. This establishes that group reasons exist. Finally, in order to triangulate on normative properties, normative properties

must appear equally to agents *as* normative, otherwise they would not triangulate on the same bit of the world. This establishes that reasons are unrestrictedly normatively public. The task of explaining spontaneous joint action has taken us seemingly quite far from our starting point of joint agency. Indeed, it brought us to some of the most difficult questions in philosophy. I think that this path shows how important joint agency is to our philosophical thinking in general. It is not just some interesting side issue with minimal application, joint agency is at the very heart of thinking and acting.

I believe it is common in philosophy these days to try to make every phenomenon a social phenomenon, or to explain everything by means of sociality. This is not what I think this dissertation does. What it does, however, is show how a particular form of sociality, the joint agency in triangulation, is a necessary condition for *individual* thought and talk. The human experience of individuality crucially depends on our joint capacities, but we should not reduce our individuality to our sociality. One of the philosophical difficulties in social philosophy is how to strike a balance between the individual and social perspectives we so obviously occupy. In this dissertation, I have tried to find that balance. I believe that I have made good on Baier's suggestion that joint agency is not a construction of individual perspectives by showing how the co-agential perspective is necessary to joint, rational agency. But the ability to occupy this perspective does not eradicate the individual perspective. Indeed, the ability for these perspectives to co-exist is helped by the non-reductive nature of the Davidsonian framework used throughout the dissertation. The tendency to reduce what is clearly irreducible and to revise concepts in light of difficulty is not amenable to the Davidsonian spirit. And, so, the fact that we occupy two very different perspectives, the individual and the joint, is not so problematic so long as we accept that we should not be constructing one out of the other. What is needed, rather, is an explanation of why it is possible for us to have these two perspectives, and the triangulation argument provides the basis for an argument of this kind. The

group reasons that triangulators respond to explain why we can occupy the co-agential perspective. To see how the individual perspective is drawn from this situation, we need to point back to one of the original purposes of the triangulation argument: to account for possession of the concept of error. As soon as a triangulator recognises that the other is not behaving as she does, she is thrown into her individual perspective. When error is on the scene, it warrants thinking of the other third-personally, as opposed to one member of the 'we'. How these two very different perspectives are unified is one avenue for future research. As I have just indicated, I believe that the Davidsonian framework within which I am working has the resources to explain this phenomenon.

There are three other paths for further research that the dissertation opens up. First, what exactly is the normative property that triangulators triangulate on? Is it the reason, the consideration, or some other property, like requiredness? 'Normative property' is just vague enough to make the argument run intelligibly: whatever the normative property in fact is agents triangulate on it. In the service of a systematic account of the world, it would be nice to know what exactly grounds an agent's normative judgments. There are two troubles in answering this question, I believe. First, the normative property cannot be a natural, non-normative property, like pleasure. But then, what are we left with? It seems like we do not have language, at the moment, to characterise what exactly the normative property is because moral philosophy for the past century has been concerned, if at all, with making space for the normative in a scientific world. The fight for the category of the normative has left us without much to say about what actually falls under that category. And the triangulation argument puts that question front and centre. The second trouble is whether this question is intelligible. Does it make sense to search for a *particular* property that is the normative property? Isn't characterising it as normative enough for our purposes? Indeed, what purpose would it serve to have a precise answer to this question? But this is the interesting thing about philosophical questions: we want to know what we do not know, and we often have no idea

how the answer will help us or bear on other philosophical matters. Because the question is not immediately unintelligible, we can only know if it is an intelligible question based on the results of the inquiry. However, I think that the consideration that we have been so focused on—preserving normativity in a non-normative world—supports inquiring more deeply into the normative.

The second path worth pursuing, I believe, is to ask to what other purposes we might put group reasons. No matter the success of the reasons account of joint action, I think that there is independent reason to think that group reasons exist and, so, it's worth thinking about which problems they can help us solve. There is no doubt that it is common in the joint action literature to apply theories to the prisoner's dilemma. Once we take on board a joint perspective, then it seems like the prisoner's dilemma dissolves because it takes for granted individuals reasoning in isolation. I also believe that the reasons account and group reasons can provide an interesting account of the prisoner's dilemma. However, what I want to focus on instead is whether group reasons help other interpersonal paradoxes, in particular, the learner's paradox from the *Meno*. The impasses that produce many interpersonal paradoxes often result from assuming the individual's perspective as primary. The learner's paradox roughly says that despite seemingly learning new things, learning is impossible because if we know, then we do not need to learn and if we do not know, then we cannot know what to look for. This impasse turns Plato towards the theory of recollection, which is tied up with some other questionable theses such as Platonic forms and eternal souls: we don't learn, we simply recollect what is already in the soul. The learner's paradox suggests that we cannot identify a teacher because we do not know what it is that we are supposed to learn, and if the learner's paradox were right, there wouldn't be any teachers. However, I think that group reasons and the co-agential perspective could be used to resolve the learner's paradox. Ignoring the problem of how anyone gains any bit of knowledge in the first place and assuming that there are people in a position to teach, then there could be group reasons for teacher and learner to engage in education.

The group reason would tell the student to search out the teacher with whom she shares a reason. The group reason traverses the individual impasse by indicating where to search because if she has access to her group reason, then she will know that there is someone who could teach her. The trouble with the learner's paradox is that it takes the individual's ignorant perspective to be the only perspective available. But, by occupying the co-agential perspective on the basis of a group reason, an individual can be shown what it is that is worth learning. Similarly, there are problems with the rationality of transformative experience that I think also assume the individual's perspective. Group reasons could help here as well. The basic idea is that we do not blindly submit to the discipline of a teacher, who hammers the fact or value into our heads, when we act from a group reason to engage with them. The group reason makes the learner's activity rational even when she doesn't know what exactly she pursues. Ultimately, there are a number of interpersonal paradoxes that might find resolution once we acknowledge that the individual perspective is not the only perspective from which we can reason. Group reasons and the co-agential perspective can shed light on these paradoxes.

The final path worth pursuing concerns ethics from the co-agential perspective. It seems that there might be a special kind of harm that can be done to others from within the co-agential perspective, or the perspective might help characterise a special kind of harm. In Chapter 1, I described the practical failure that an agent sometimes experiences when she occupies the co-agential perspective with another: without excusing circumstances, my co-agent fails *me* when she doesn't do her part of what we have reason to do. It isn't simply the case that we failed to perform our action but that *we* failed to do what we had reason to do because *she* failed us. Still, because we failed, *I* failed—not as an individual, but as part of a group. There seems to be a kind of agential harm that comes from situations like this. Perhaps it is a form of disrespect. Though she doesn't negate my agency, she certainly frustrates my agency by limiting my possibilities. In Kantian terms,

she hasn't stripped me of my autonomy and yet I *qua* group member cannot do what we ought to do. In a way, I never have freedom to do what *we* ought to do, but there seems to be an interesting difference between two co-agents trying to act on their group reason and two co-agents who recognise their group reason and one willfully ignores it. In the first case, we are both trying to do what we have reason to do and yet the world may stop us. But in the second case, one co-agent has shut down the possibility of our joint activity—she has unilaterally put me in a position where I am compelled by a group reason and yet *cannot* succeed.[338] This co-agential harm can feel especially vicious since we are not simply other people to each other, we are supposed to be co-agents, we share a special bond. There is much more to say on the ethics of the co-agential perspective, but I still need to find the words.

In the end, the co-agential perspective, group reasons, and spontaneous joint agency are rich resources for inquiring about our social world. This framework, I believe, offers a different perspective that hopefully provides some clarity on our social lives. We are not trapped in our personal theatres looking out at others. Rather, we can think, feel, and act together.

[338] I do not think that this situation eliminates the group reason because we want the group to survive such that I can convince my co-agent to act with me. If the group reason is eliminated by one person ignoring it, then group reasons cannot perform one of their main functions, to compel agents to be motivated by it.

BIBLIOGRAPHY

Alonso, Facundo M. 2009. "Shared Intention, Reliance, and Interpersonal Obligations." *Ethics* 119 (3): 444–75. https://doi.org/10.1086/599984.

Alvarez, Maria. 2013. *Kinds of Reasons*. Oxford: Oxford University Press.

———. 2018. "Reasons for Action, Acting for Reasons, and Rationality." *Synthese* 195 (8): 3293–3310.

Amoretti, Maria Cristina, and Gerhard Preyer, eds. 2011. *Triangulation: From an Epistemological Point of View*. Frankfurt: Ontos Verlag.

Anscombe, G. E. M. 2000. *Intention*. 2nd ed. Cambridge, Mass: Harvard University Press.

Aristotle. 1941. "Politics." In *The Basic Works of Aristotle*, edited by Richard McKeon, translated by W.D. Ross, 1127–1324. New York: Random House.

Bacharach, Michael. 1999. "Interactive Team Reasoning: A Contribution to the Theory of Cooperation." *Research in Economics* 53: 117–47.

Baier, Annette C. 1970. "Act and Intent." *The Journal of Philosophy* 67 (19): 648–58. https://doi.org/10.2307/2024585.

———. 1997. "Doing Things with Others: The Mental Commons." In *Commonality and Particularity in Ethics*, edited by Lilli Alanen, Sara Heinamaa, and Thomas Wallgren, 15–44. Houndmills, UK: MacMillan.

Baltzer, Ulrich. 2002. "Joint Action of Large Groups." In *Social Facts and Collective Intentionality*, 1–17. Frankfurt: Dr. Hansel-Hohenhausen AG.

Bertea, Stefano. 2017. "Why One's Practical Reasons Are Not Just One's Own Private Affair." *International Philosophical Inquiry* 41 (1): 63–85.

Binmore, Ken, and Larry Samuelson. 2001. "Coordinated Action in the Electronic Mail Game." *Games and Economic Behavior* 35 (1–2): 6–30. https://doi.org/10.1006/game.1998.0749.

Bittner, Rudiger. 2002. "An Action for Two." In *Social Facts and Collective Intentionality*, edited by Georg Meggle, 35–42. Frankfurt: Dr. Hansel-hohenhausen.

Blomberg, Olle. 2016a. "Shared Intention and the Doxastic Single End Condition." *Philosophical Studies* 173 (2): 351–72. https://doi.org/10.1007/s11098-015-0496-z.

———. 2016b. "Common Knowledge and Reductionism about Shared Agency." *Australasian Journal of Philosophy* 94 (2): 315–26.

———. 2018. "Practical Knowledge and Acting Together." In *Socially Extended Epistemology*, edited by J. Adam Carter, Andy Clark, Jesper Kallestrup, S. Orestis Palermos, and Duncan Pritchard, 87–111. Oxford: Oxford University Press.

Boghossian, Paul A. 1989. "The Rule Following Considerations." *Mind* 98: 507–49.

Boyd, Richard. 1988. "How to Be a Moral Realist." In *Moral Realism*, edited by Geoffrey Sayre-McCord, 187–228. Ithaca, NY: Cornell University Press.

Bratman, Michael E. 1987. *Intentions, Plans, and Practical Reason*. Cambridge, Mass.: Harvard UP.

———. 1993. "Shared Intention." *Ethics* 104 (1): 97–113.

———. 1999a. *Faces of Intention: Selected Essays on Intention and Agency*. Cambridge: Cambridge University Press.

———. 1999b. "I Intend That We J." In *Faces of Intention*, 142–62. Cambridge: Cambridge University Press.

———. 1999c. "Shared Intention and Mutual Obligation." In *Faces of Intention: Selected Essays on Intention and Agency*, 130–41. Cambridge: Cambridge University Press.

———. 2014. *Shared Agency*. Oxford: Oxford UP.

Bridges, Jason. 2006. "Davidson's Transcendental Externalism." *Philosophy and Phenomenological Research* 73 (2): 290–315. https://doi.org/10.1111/j.1933-1592.2006.tb00619.x.

Burge, Tyler. 2006. "Reason and the First Person." In *Knowing Our Own Minds*, edited by Crispin Wright, Repr, 243–71. Mind Association Occasional Series. Oxford: Clarendon Press.
Chant, Sara Rachel, and Zachary Ernst. 2008. "Epistemic Conditions for Collective Action." *Mind* 117 (467): 549–73.
Child, William. 2001. "Triangulation: Davidson, Realism and Natural Kinds." *Dialectica* 55 (1): 29–49.
Copp, David. 1997. "Belief, Reason, and Motivation: Michael Smith's 'The Moral Problem.'" *Ethics* 108 (1): 33–54.
Dancy, Jonathan. 1993. *Moral Reasons*. Oxford: Basil Blackwell.
———. 2002. *Practical Reality*. Oxford: Oxford University Press.
Davidson, Donald. 1994. "The Social Aspect of Langauge." In *Truth, Language, and History*, 109–26. Oxford: Oxford University Press.
———. (1963) 2001. "Actions, Reasons, and Causes." In *Essays on Actions and Events*, 2nd ed., 3–20. Oxford: Oxford University Press.
———. (1990) 2001. "Epistemology Externalized." In *Subjective, Intersubjective, Objective*, 193–204. Oxford: Oxford University Press.
———. 2001a. "Externalisms." In *Interpreting Davidson*, edited by Petr Kotatko, Peter Pagin, and Gabriel Segal, 1–16. Stanford: CLSI Publications.
———. 2001b. *Inquiries into Truth and Interpretation*. Oxford: Clarendon ; Oxford University Press.
———. 2001c. "Intending." In *Essays on Actions and Events*, 2nd ed., 83–102. Oxford: Oxford University Press.
———. 2001d. "Mental Events." In *Essays on Actions and Events*, 2nd ed., 207–24. Oxford: Oxford University Press.
———. (1997) 2001. "The Emergence of Thought." In *Subjective, Intersubjective, Objective*, 123–34. Oxford: Oxford University Press.
———. (1992) 2001. "The Second Person." In *Subjective, Intersubjective, Objective*, 107–21. Oxford: Oxford University Press.
———. (1975) 2001. "Thought and Talk." In *Inquiries into Truth and Interpretation*, 155–70. Oxford: Oxford University Press.
———. 2004a. *Problems of Rationality*. Oxford: Clarendon ; Oxford University Press.
———. 2004b. "The Objectivity of Values." In *Problems of Rationality*, 39–58. Oxford: Clarendon ; Oxford University Press.
———. 2004c. "The Problem of Objectivity." In *Problems of Rationality*, 3–18. Oxford: Clarendon ; Oxford University Press.
Enoch, David. 2011. *Taking Morality Seriously: A Defense of Robust Realism*. Oxford: Oxford University Press.
Ferrero, Luca. 2013. "Can I Only Intend My Own Actions?" In *Oxford Studies in Agency and Responsibility Volume 1*, edited by David Shoemaker, 70–94. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199694853.003.0004.
Foot, Philippa. 2001. *Natural Goodness*. Oxford: Oxford University Press; Clarendon.
Gert, Joshua. 2002. "Korsgaard's Private-Reasons Argument." *Philosophy and Phenomenological Research* 64 (2): 303–24. https://doi.org/10.1111/j.1933-1592.2002.tb00003.x.
Gilbert, Margaret. 1989. *On Social Facts*. London: Routledge.
———. 1990. "Walking Together: A Paradigmatic Social Phenomenon." *Midwest Studies In Philosophy* 15 (1): 1–14.
———. (1997) 2000. "What Is It for Us to Intend?" In *Sociality and Responsibility*, 14–36. Lanham, Md: Rowman & Littlefield.

———. 2006. *A Theory of Political Obligation: Membership, Commitment, and the Bonds of Society*. Oxford: Oxford University Press; Clarendon.
———. 2008. "Two Approaches to Shared Intention: An Essay in the Philosophy of Social Phenomena." *Analyse & Kritik* 30 (2): 483-514. https://doi.org/10.1515/auk-2008-0208.
———. 2009a. "Shared Intention and Personal Intentions." *Philosophical Studies* 144 (1): 167–87. https://doi.org/10.1007/s11098-009-9372-z.
———. 2009b. "Pro Patria: An Essay on Patriotism." *The Journal of Ethics* 13 (4): 319–46. https://doi.org/10.1007/s10892-009-9062-2.
———. 2014a. "A Real Unity of Them All." In *Joint Commitment: How We Make the Social World*, 341–56. Oxford: Oxford University Press.
———. 2014b. "Acting Together." In *Joint Commitment: How We Make the Social World*, 23–36. Oxford: Oxford University Press.
———. 2014c. "Consdierations on Joint Commitment." In *Joint Commitment: How We Make the Social World*, 37–57. Oxford: Oxford University Press.
———. 2014d. *Joint Commitment: How We Make the Social World*. Oxford: Oxford University Press.
Gold, Natalie. 2018. "Team Reasoning and Spontaneous Collective Intentions." *Revue d'économie Politique* 128 (3): 333–53. https://doi.org/10.3917/redp.283.0333.
Gold, Natalie, and Robert Sugden. 2007. "Collective Intentions and Team Agency." *The Journal of Philosophy* 104 (3): 109–37.
Haase, Matthias. 2012. "Three Forms of the First Person Plural." In *Rethinking Epistemology*, edited by Günter Abel and James Conant, 229–56. Berlin: De Gruyter.
Halpern, Joseph Y. 1986. "Reasoning about Knowledge: An Overview." In *Reasoning about Knowledge: Proceding of the 1986 Conference*, 1–18. San Francisco, CA: Morgan Kaufmann.
Harman, Gilbert. 1977. *The Nature of Morality*. Oxford: Oxford University Press.
Heath, Joseph. 2020. "Methodological Individualism." In *Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Summer 2020 Edition. https://plato.stanford.edu/archives/sum2020/entries/methodological-individualism/.
Hegel, G. W. F. 1977. *Hegel's Phenomenology of Spirit*. Translated by A. V. Miller. Oxford: Oxford University Press.
Hurley, Paul. 2001. "A Kantian Rationale for Desire-Based Justification." *Philosopher's Imprint* 1 (2): 1–16.
Hursthouse, Rosalind. 1999. *On Virtue Ethics*. Oxford: Oxford University Press; Clarendon.
Julius, A.J. 2013. "The Possibility of Exchange." *Politics, Philosophy & Economics* 12 (4): 361–74. https://doi.org/10.1177/1470594X13496068.
———. 2023. *Reconstruction*. Princeton, NJ: Princeton University Press.
Kopec, Matthew, and Suemas Miller. 2018. "Shared Intention Is Not Joint Commitment." *Journal of Ethics and Social Philosophy* 13 (2): 179-189. https://doi.org/10.26556/jesp.v13i2.250.
Korsgaard, Christine M. 1986. "Skepticism about Practical Reason." *The Journal of Philosophy* 83 (1): 5–25. https://doi.org/10.2307/2026464.
———. 1993. "The Reasons We Can Share." *Social Philosophy and Policy* 10 (1): 24–51.
———. 1996. *The Sources of Normativity*. Cambridge: Cambridge University Press.
———. 2009. *Self-Constitution: Agency, Identity, and Integrity*. Oxford ; New York: Oxford University Press.
Kripke, Saul. 1982. *Wittgenstein on Rules and Private Language*. Oxford: Blackwell.
Kutz, Christopher. 2000a. *Complicity*. New York: Cambridge University Press.
———. 2000b. "Acting Together." *Philosophy and Phenomenological Research* 61 (1): 1–31. https://doi.org/10.2307/2653401.

Laurence, Ben. 2011. "An Anscombian Approach to Collective Action." In *Essays on Anscombe's "Intention,"* edited by Anton Ford, Jennifer Hornsby, and Frederick Stoutland, 270-294. Cambridge, MA and London, England: Harvard University Press. https://doi.org/10.4159/harvard.9780674060913.c11.
LeBar, Mark. 2001. "Korsgaard, Wittgenstein, and the Mafioso." *The Southern Journal of Philosophy* 39 (2): 261–71. https://doi.org/10.1111/j.2041-6962.2001.tb01817.x.
Lewis, David. 1969. *Convention*. Cambridge, Mass.: Harvard University Press.
List, Christian. 2018. "What Is It Like to Be a Group Agent?: What Is It Like to Be a Group Agent?" *Noûs* 52 (2): 295–319.
List, Christian, and Philip Pettit. 2011. *Group Agency: The Possibility, Design, and Status of Corporate Agents*. Oxford: Oxford University Press.
Ludwig, Kirk. 2014. "The Ontology of Collective Action." In *From Individual to Collective Intentionality*, 112–33. Oxford: Oxford University Press.
———. 2016. *From Individual to Plural Agency: Collective Action 1*. Oxford: Oxford University Press.
Mackie, J.L. 1977. *Ethics: Inventing Wrong and Right*. London: Pelican.
Marx, Karl. 1994. *Selected Writings*. Edited by Lawrence H. Simon. Indianapolis: Hackett.
Mathiesen, Kay. 2002. "Searle, Collective Intentions, and Individualism." In *Social Facts and Collective Intentionality*, edited by Georg Meggle, 185–204. Frankfurt: Dr. Hansel-hohenhausen.
McDowell, John. 1978. "Are Moral Requirements Hypothetical Imperatives?" *Proceedings of the Aristotelian Society, Supplementary Volumes* 5: 13-29.
———. 1979. "Virtue and Reason." *Monist* 62: 331–50.
———. 1984. "Wittgenstein on Following a Rule." *Synthese* 58: 325–63.
———. 1985. "Values and Secondary Qualities." In *Morality and Objectivity*, edited by Ted Honderich, 110–29. London: Routledge.
———. 1992. "Meaning and Intentionality in Wittgenstein's Later Philosophy." *Midwest Studies in Philosophy* 17 (1): 40–52. https://doi.org/10.1111/j.1475-4975.1992.tb00141.x.
———. 2010. "What Is the Content of an Intention in Action?" *Ratio* 23 (4): 415-432.
Meijers, Anthoniew M. 2003. "Can Collective Intentionality Be Individualized?" *American Journal of Economics and Sociology* 62 (1): 167–83. https://doi.org/10.1111/1536-7150.t01-1-00006.
Miller, Seumas. 2001. *Social Action*. Cambridge: Cambridge University Press.
Montminy, Martin. 2003. "Triangulation, Objectivity, and the Ambiguity Problem." *Critica* 35 (105): 25–48.
Moore, G. E. 1903. *Principia Ethica*. Cambridge: Cambridge University Press.
Myers, Robert H. 2012. "Desires and Normative Truths: A Holist's Response to the Sceptics." *Mind* 121 (482): 375–406. https://doi.org/10.1093/mind/fzs068.
———. 2017. "Holism in Action." In *Wittgenstein and Davidson on Language, Thought, and Action*, edited by Claudine Verheggen, 8–27. Cambridge: Cambridge University Press.
Myers, Robert H., and Claudine Verheggen. 2016. *Davidson's Triangulation Argument: A Philosophical Inquiry*. New York: Routledge.
Nagel, Thomas. 1970. *The Possibility of Altruism*. Princeton, NJ: Princeton University Press.
———. 1986. *The View from Nowhere*. Oxford: Oxford University Press.
Norman, Richard. 2000. "Public Reasons and the `Private Language' Argument." *Philosophical Investigations* 23 (4): 292-314.
O'Day, Ken. 1998. "Normativity and Interpersonal Reasons." *Ethical Theory and Moral Practice* 1 (1): 61–87. https://doi.org/10.1023/A:1009908512139.
Pacherie, Elisabeth. 2011. "Framing Joint Action." *Review of Philosophy and Psychology* 2 (2): 173–92. https://doi.org/10.1007/s13164-011-0052-5.

———. 2013. "Intentional Joint Agency: Shared Intention Lite." *Synthese* 190 (10): 1817–39. https://doi.org/10.1007/s11229-013-0263-7.
Parfit, Derek. 2006. "Normativity." *Oxford Studies in Metaethics* 1: 325–80.
———. 2011. *On What Matters*. Vol. 2. Oxford: Oxford University Press.
Petersson, Björn. 2007. "Collectivity and Circularity." *The Journal of Philosophy* 104 (3): 138–56.
Pettit, Philip, and David Schweikard. 2006. "Joint Actions and Group Agents." *Philosophy of the Social Sciences* 36 (1): 18–39.
Railton, Peter. 1986. "Moral Realism." *Philosophical Review* 95: 163–207.
Robbins, Michael H. 2002. "Joint Commitment and Circularity." In *Social Facts and Collective Intentionality*, edited by Georg Meggle, 300–321. Frankfurt: Dr. Hansel-hohenhausen.
Rödl, Sebastian. 2014. "Intentional Transaction." *Philosophical Explorations* 17 (3): 304–16. https://doi.org/10.1080/13869795.2014.941909.
Roelofs, Luke. 2017. "Rational Agency without Self-Knowledge: Could 'We' Replace 'I'?: Rational Agency without Self-Knowledge." *Dialectica* 71 (1): 3–33. https://doi.org/10.1111/1746-8361.12169.
Roessler, Johannes. 2020. "Plural Practical Knowledge." *Inquiry*, July, 1–20. https://doi.org/10.1080/0020174X.2020.1787221.
Rønnow-Rasmussen, Toni. 2009. "Normative Reasons and the Agent-Neutral/Relative Dichotomy." *Philosophia* 37 (2): 227–43. https://doi.org/10.1007/s11406-008-9164-2.
Roth, Abraham Sesshu. 2003. "Practical Intersubjectivity." In *Socializing Metaphysics*, edited by Frederick F. Schmitt, 65–92. Lanham, Md: Rowman & Littlefield.
———. 2004. "Shared Agency and Contralateral Commitments." *Philosophical Review* 113 (3): 359–410.
Rovane, Carol. 1998. *The Bounds of Agency*. Princeton, NJ: Princeton University Press.
Rubenstein, Ariel. 1989. "The Electronic Mail Game: Strategic Behaviour under 'Almost Common Knowledge.'" *American Economic Review* 79 (3): 385–91.
Salice, Alessandro. 2015. "There Are No Primitive We-Intentions." *Review of Philosophy and Psychology* 6: 695–715.
Satne, Glenda Lucila. 2020. "Practical Knowledge and Shared Agency: Pluralizing the Anscombean View." *Inquiry*, November, 1–28. https://doi.org/10.1080/0020174X.2020.1837236.
Sayre-McCord, Geoffrey. 1997. "The Metaethical Problem." *Ethics* 108 (1): 55–83.
Scanlon, T. M. 1998. *What We Owe to Each Other*. Cambridge, Mass: Belknap Press of Harvard University Press.
———. 2014. *Being Realistic about Reasons*. Oxford: Oxford University Press.
Schmid, Hans Bernhard. 2003. "Can Brains in Vats Think as a Team?" *Philosophical Explorations* 6 (3): 201–17. https://doi.org/10.1080/10002003098538750.
———. 2009. *Plural Action*. Dordrecht: Springer.
———. 2014. "Plural Self-Awareness." *Phenomenology and the Cognitive Sciences* 13: 7–24.
———. 2016. "On Knowing What We Are Doing Together." In *The Epistemic Life of Groups*, edited by Miranda Fricker and Michael S. Brady, 51–72. Oxford: Oxford University Press.
———. 2017. "What Kind of Mode Is the We-Mode?" In *Social Ontology and Collective Intentionality. Critical Essays on the Philosophy of Raimo Tuomela with His Responses*, edited by Gerhard Preyer and Georg Peter, 79–93. Cham, Switzerland: Springer.
———. 2018. "The Subject of 'We Intend.'" *Phenomenology and the Cognitive Sciences* 17 (2): 231–43. https://doi.org/10.1007/s11097-017-9501-7.
Schmitt, Frederick F. 2003a. "Joint Action: From Individualism to Supraindividualism." In *Socializing Metaphysics*, 129–66. Lanham, Md: Rowman & Littlefield.

———. 2003b. "Socializing Metaphysics: An Introduction." In *Socializing Metaphysics*, 1–38. Lanham, Md: Rowman & Littlefield.
Schönherr, Julius. 2019. "Lucky Joint Action." *Philosophical Psychology* 32 (1): 123–42.
Schweikard, David P., and Hans Bernard Schmid. 2013. "Collective Intentionality." *Stanford Encyclopedia of Philosophy*.
Searle, John R. 1990. "Collective Intentions and Actions." In *Intentions in Communication*, edited by P. Cohen, J. Morgan, and M.E. Pollack, 401–15. Cambridge, Mass.: Bradford-MIT Press.
———. 1995. *The Construction of Social Reality*. New York: Free Press.
Seemann, Axel. 2009. "Why We Did It: An Anscombian Account of Collective Action." *International Journal of Philosophical Studies* 17 (5): 637–55. https://doi.org/10.1080/09672550903301580.
Sellars, Wilfrid. 1980. "On Reasoning about Values." *American Philosophical Quarterly* 17 (2): 81–101.
Shafer-Landau, Russ. 2003. *Moral Realism: A Defence*. Oxford: Clarendon ; Oxford University Press.
Shapiro, Scott J. 2014. "Massively Shared Agency." In *Rational and Social Agency: The Philosophy of Michael Bratman*, edited by Manuel Vargas and Gideon Yaffe, 257–93. Oxford: Oxford University Press.
Simmel, Georg. 1971a. *George Simmel: On Individuality and Social Forms*. Edited by D. N. Levine. Chicago: University of Chicago Press.
———. 1971b. "How Is Society Possible?" In *Georg Simmel: On Individuality and Social Forms*, edited by D. N. Levine, 6–22. Chicago: University of Chicago Press.
Sinclair, Robert. 2005. "The Philosophical Significance of Triangulation." *Metaphilosophy* 36 (5): 708–28.
Smith, Michael. 1987. "The Humean Theory of Motivation." *Mind, New Series* 96 (381): 36–61.
———. 1991. "Realism." In *A Companion to Ethics*, edited by Peter Singer, 399–410. Oxford: Blackwell.
———. 1994. *The Moral Problem*. Oxford: Blackwell.
———. 1995. "Internal Reasons." *Philosophy and Phenomenological Research* 55 (1): 109–31.
Stalnaker, Robert C. 2002. "Common Ground." *Linguistics and Philosophy* 25 (5–6): 701–21.
Stoutland, Frederick. 1997. "Why Are Philosophers of Action so Anti-Social?" In *Commonality and Particularity in Ethics*, edited by Lilli Alanen, Sara Heinamaa, and Thomas Wallgren, 45–74. New York: St. Martin's Press.
Strawson, Peter Frederick. (1962) 2008. "Freedom and Resentment." In *Freedom and Resentment and Other Essays*, 1–28. London: Routledge.
Street, Sharon. 2010. "What Is Constructivism in Ethics and Metaethics?: Constructivism in Ethics and Metaethics." *Philosophy Compass* 5 (5): 363–84. https://doi.org/10.1111/j.1747-9991.2009.00280.x.
———. 2012. "Coming to Terms with Contingency: Humean Constructivism about Practical Reason." In *Constructivism in Practical Philosophy*, edited by James Lenman and Yonatan Shemmer, 40–59. Oxford University Press.
Sturgeon, Nicholas L. 1974. "Altruism, Solipsism, and the Objectivity of Reasons." *The Philosophical Review* 83 (3): 374–402. https://doi.org/10.2307/2183699.
———. 1985. "Moral Explanations." In *Morality, Reason, and Truth*, edited by David Copp and David Zimmerman, 49–78. Totowa, NJ: Rowman and Allanheld.
Sugden, Robert. 1993. "Thinking as a Team: Toward an Explanation of Nonselfish Behavior." *Social Philosophy and Policy* 10: 69–89.
———. 2003. "The Logic of Team Reasoning." *Philosophical Explorations* 6: 165–81.
Taylor, Charles. (1980) 1985. "Theories of Meaning." In *Human Agency and Language*, 248–92. Cambridge: Cambridge University Press.

Thompson, Michael. 2008. *Life and Action: Elementary Structures of Practice and Practical Thought*. Cambridge, Mass: Harvard University Press.
Tollefsen, Deborah. 2005. "Let's Pretend!: Children and Joint Action." *Philosophy of the Social Sciences* 35 (1): 75–97. https://doi.org/10.1177/0048393104271925.
Tuomela, Raimo. 1991. "We Will Do It: An Analysis of Group-Intentions." *Philosophy and Phenomenological Research* 51 (2): 249. https://doi.org/10.2307/2108127.
———. 2002. *The Philosophy of Social Practices*. Cambridge: Cambridge University Press.
———. 2003. "The We-Mode and the I-Mode." In *Socializing Metaphysics*, edited by Frederick F. Schmitt, 93–128. Lanham, Md: Rowman & Littlefield.
———. 2005. "We-Intentions Revisited." *Philosophical Studies* 125 (3): 327–69. https://doi.org/10.1007/s11098-005-7781-1.
———. 2007. *The Philosophy of Sociality*. Oxford: Oxford University Press.
———. 2009. "Collective Intentions and Game Theory." *Journal of Philosophy* 106 (5): 292–300.
Tuomela, Raimo, and Kaarlo Miller. 1988. "We-Intentions." *Philosophical Studies* 53 (3): 367–89.
Velleman, J. David. 1989. *Practical Reflection*. Princeton: Princeton University Press.
———. 1992. "The Guise of the Good." *Nous* 26 (1): 3–26.
———. 1997. "How To Share An Intention." *Philosophy and Phenomenological Research* 57 (1): 29.
Verheggen, Claudine. 1997. "Davidson's Second Person." *The Philosophical Quarterly* 47 (188): 361–69. https://doi.org/10.1111/1467-9213.00065.
———. 2006. "How Social Must Language Be." *Journal for the Theory of Social Behaviour* 36 (2): 203–19.
———. 2007. "Triangulating with Davidson." *Philosophical Quarterly* 57 (226): 96–105.
———. 2017. "Davidson's Semantic Externalism: From Radical Interpretation to Triangulation." *Argumenta* 3 (1): 145–61.
Wallace, R Jay. 2009. "The Publicity of Reasons." *Philosophical Perspectives* 23: 471–97.
White, Stephen J. 2018. "Intention and Prediction in Means-End Reasoning." *American Philosophical Quarterly* 55 (3): 251–66.
Williams, Bernard. 1981. "Internal and External Reasons." In *Moral Luck*, 101–13. Cambridge: Cambridge University Press.
———. 1995. *Making Sense of Humanity*. Cambridge: Cambridge University Press.
———. 2001. "Postscript: Some Further Notes on Internal and External Reasons." In *Varieties of Practical Reasoning*, edited by Elijah Millgram, 91–97. Cambridge, MA and London, England: Bradford-MIT Press.
Wittgenstein, Ludwig. 2009. *Philosophical Investigations*. Translated by G.E.M. Anscombe, P.M.S. Hacker, and Joachim Schulte. 4th ed. Oxford: Wiley-Blackwell.
Zahavi, Dan. 2008. "Collective Intentionality and Plural Pre-Reflective Self-Awareness." *Journal of Social Philosophy* 49 (1): 61–75.
Zaibert, L.A. 2003. "Collective Intentions and Collective Intentionality." *American Journal of Economics and Sociology* 62 (1): 209–32. https://doi.org/10.1111/1536-7150.t01-1-00008.