

On Collective Action:
*underpinning the plural subject with a
model of planning agency*

Joseph Kisolo-Ssonko

Thesis submitted for the degree of
Doctor of Philosophy

Department of Philosophy
University of Sheffield
March 2014

Abstract

This thesis aims to give an account of collective action. It starts with a detailed presentation of its underlying phenomenology. It is argued that in order to understand this phenomenology, we must move beyond the framework of individual agency; thus rejecting Michael Bratman's Shared Cooperative Activity Account. Doing so opens up a space for Margaret Gilbert's Plural Subject Theory. Plural Subject Theory is presented as capturing this phenomenology by allowing that we can act as collective agents. However, it also creates a puzzle centring on the relation between individual autonomy and constraint by the collective will. The solution to this puzzle, this thesis argues, is to apply Bratman's planning theory of agency to the collective agent. In doing so, Gilbert's theory is improved, such that it is better able to capture the sense in which living social lives entangles our sense of individual agentive identity with our sense of collective agentive identity.

Acknowledgements

I would like to thank the University of Sheffield Philosophy Department for funding me to complete a research training MA and encouraging me to take up the PhD; the AHRC for providing me with the funding to allow me to undertake the PhD; and Steve Deegan for proofreading and recommending many additional commas.

I have benefited greatly from the guidance and support of my supervisors, Paul Faulkner and Rosanna Keefe. In particular, their reigning me back from overambitious claims and making me concentrate hard on getting the fundamentals sorted, and Paul's encouragement to go running to de-stress.

Parts of this thesis began as, or formed the basis for, papers presented at the Department's Postgraduate Seminar, and benefited from the comments received on those occasions. In particular, my fellow postgraduate students helped me to get a grasp on the variety of intuitions regarding questions of autonomy. Similarly, the comments from audience members at the talks I gave at the 6th and 7th Conferences on Collective Intentionality, including those of Michel Bratman and Margaret Gilbert, helped me hone my ideas. Feedback from anonymous referees, on my two published papers, also helped greatly.

The assistance of Jeff Lancaster, in helping me to knock the final draft into what is hopefully a good shape, and shaking out any sneaky bugs, has been much appreciated. I am eternally grateful to my mum, Susan Pascoe, for raising me in a way that allowed me, against the troubles I faced, to reach a place where I could complete this study.

Finally, I would like to thank my partner Kim Harrison, for bucketloads of help with spelling and grammar (particularly for battling through some of the early, very rough, drafts) and for providing me with the emotional support without which it is certain that I would never have been able to get this far.

Contents

0.0 – Introduction.....	p.1
1.0 – Our social lives.....	p.4
1.1 – Capturing the phenomenology of the phenomenon.....	p.5
1.1.a – Unity	
1.1.b – Collective intentionality	
1.1.c – Detachment	
1.1.d – Constraint	
1.2 – The phenomenological state of play.....	p.20
1.3 – Why should we care?.....	p.22
1.4 – Chapter summary.....	p.27
2.0 – Mere sums.....	p.29
2.1 – The framework of individual agency.....	p.31
2.2 – Mere sums -vs- the experience of unity.....	p.37
2.3 – Mere sums -vs- the feeling of collective intentionality.....	p.40
2.3.a – The problem of control	
2.3.b – Bratman's solution to the problem of control	
2.3.c – Bratman's Shared Cooperative Activity Account	
2.4 – Mere sums -vs- detachment.....	p.54
2.5 – Mere sums -vs- constraint.....	p.59
2.6 – Chapter summary.....	p.63
3.0 – Plural subjects.....	p.64
3.1 – The possibility of social unity.....	p.66
3.1.a – Gilbert's Plural Subject Theory	
3.2 – The we -vs- the feeling of unity.....	p.74
3.3 – The we -vs- the feeling of collective intentionality.....	p.79
3.4 – The we -vs- the feeling of detachment.....	p.83
3.5 – The we -vs- the feeling of constraint.....	p.85
3.5.a – The complication of hierarchy and standing conventions	
3.6 – Chapter summary.....	p.90
4.0 – Harmony, discord and autonomy.....	p.91
4.1 – Harmony and discord of wills.....	p.93
4.2 – Harmony of commitments of the will.....	p.96
4.3 – Discord between commitments of the will.....	p.103
4.3.a – The puzzle of intersubjectivity	
4.4 – Chapter summary.....	p.114
5.0 – Constraint by the will.....	p.115
5a.0 – Constraint by the individual will.....	p.118
5a.1 – Intentions as rational scaffolding	
5a.2 – Why practical utility is not enough	
5a.3 – Self-governance and having a where-I-stand	
5b.0 – Constraint by the collective Will.....	p.137
5b.1 – Constructing the plural agent	
5b.2 – Identification	
5b.3 – Inauthenticity and the transcendental argument	
5c.1 – Chapter summary.....	p.153
6.0 – Modified PST.....	p.155
6.1 – Replacing pooling with entanglement.....	p.156
6.2 – PST* as better capturing the phenomenology.....	p.164
6.2.a – Romantic couples	
6.2.b – Political obligation	
6.3 – Chapter summary.....	p.174
7.0 – Conclusion.....	p.175
8.0 – Bibliography.....	p.183

“... when we came down from the mountain, there was much foolish talk about who got there first ... All this was nonsense. And in Kathmandu, to put a stop to such talk, Hillary and I signed a statement in which we said 'we reached the summit together'”¹

Tenzing Norgay, on
reaching the summit of Everest
together with Edmund Hillary

Introduction: Social Life

Even a mundane day is likely to involve social interaction. On the occasional mornings when I am not caring for my children, I like to buy a newspaper from the corner shop. Doing so requires talking to the owner, and giving her my money; money I had to get others to give to me. Moreover, even when I get back home and I am alone reading it, my interpretation of the symbols on its pages involves their meaning having been defined by my linguistic community. The point is that, understanding these social interactions will clearly play a vital part in understanding our actions. This thesis takes up a particular aspect

¹ Ullman, 1956, p.263.

of this task. Our actions can be social not only in that they may involve cooperation, as in the cornershop interaction, nor only in the fact that they may depend on the actions of others. Additionally, our actions can be social in the sense that *we feel ourselves to be able to do certain things together*. In this way, I will be examining what it means to say that we are capable of *collective action*.

Because the individual has had such primacy in many inquiries into the nature of agency, we might think that collective action must be a mere illusion; an illusion born of the complexity of the way our individual wills interact. However, in examining the best example of such an individualist strategy, Michael Bratman's *Shared Cooperative Activity Account*, we shall see how even this account fails to capture the essential *collectivity* of collective action. Instead, I will argue that we should accept Margaret Gilbert's claim that we have the ability to act as *plural subjects*.² For Gilbert, forming plural subjects involves, in a real sense, the *pooling* of individual wills to create a collective will. This is the core of her *Plural Subject Theory*. It is the necessity of such an outlook that will form the bedrock of this thesis.

Gilbert supports her theory by telling a compelling story of the phenomenology of our social experience. The most important elements captured by her theoretical set are the sense in which we experience direct normative pressure to act in line with collective intentions and the fact that we feel this normative pressure to be hard to escape from. For Gilbert, when *we* commit to act as a collective *we* form a plural subject. This means that I, as a part of *we*, will find myself to be rationally bound to play my part unless *we*, together, change our commitment.

Despite my general allegiance to the broad picture Gilbert paints, I will argue that, as it stands, her theory is deficient. This will become apparent when we turn our attention towards the relationship between individual autonomy and our being bound by the collective will. Untangling this issue will motivate us to find a robust way to underpin the nature of plural agency. I will make the case that this can be achieved by applying Michael

2 Made most systematically in her book *On Social Facts* (1992).

Bratman's *Planning Theory of Agency* (which he uses to explain how individual wills can bind individuals) to collective action.³

Ultimately, I shall conclude that a proper understanding of collective action requires that we give an adjusted form of Gilbert's Plural Subject Theory. We must augment it by allowing that we need not voluntarily pool our wills. Rather, our wills can become *entangled* merely through our living social lives. Such a modified Plural Subject Theory will be shown to fit better with our actual experiences of being bound by the collective will; in particular, it will be shown to capture the variation we feel in the costliness to our sense of identity of defection from different groups. Plural agency, as we shall see, is a real phenomenon but it is also a complex and messy affair.

³ Bratman, 2007. Note that this is not a use that Bratman would support, as he favours a more individualist account of collective action (or as he calls it shared cooperative activity). However, as we shall see, it is possible to separate his planning theory of agency from his stance on collective action.

Chapter 1 – Our Collective Actions

Some of the things that we do we do alone; just now I am sitting here alone typing this sentence. However, we are social beings and many of the things that we do, we do together with others; this evening some friends and I plan to meet up and play football together in the park. The former action is mine alone; the latter, however, will be one that I will *do together* with these friends. The sitting is merely *my* sitting, the playing in contrast will be *our* playing. This observation might be thought so obvious as to be mundane, but, as with many issues that appear straightforward, much philosophical complexity and controversy lurks not too far beneath the surface. This thesis will make the case that there is a sense of doing things together which is distinct from a mere multiplicity of related individual acts – a sense I shall refer to as collective action. In exploring this phenomenon, I will be chiefly concerned with the work of the two philosophers who have explored it with most focus, Margaret Gilbert and Michael Bratman. Their opposing theories are the most developed attempts at understanding this phenomenon.

As this thesis progresses, I will argue that Bratman's Shared Cooperative Action Account (SCA) is unable to give a fully satisfactory account of collective action and that such an understanding is possible only if we accept the key aspect of Margaret Gilbert's Plural Subject Theory (PST). That is, we must understand the individuals involved as in a *real* sense pooling their agency such that together they form a plural agent. Accepting this does not require that we conjure up any metaphysically suspect separate social realm; however, it does require that we move beyond the notion that we can only understand agency in terms of individual separate agents. Such an approach, I will propose, can best deal with the phenomenology of collective action.

Before we can begin to criticise and construct theory, we must have a clear view of the target. This first chapter explores the particular characteristics of what it feels like to us to act together with others. I will claim that the essential elements of this experience are the feelings of *unity*, *collective intentionality*, *detachment* and *constraint*. This exploration of the

phenomenology of collective action forms the bedrock for the theoretical work that will follow. At the end of this chapter, I will also briefly situate the problem of understanding collective action within the wider philosophical issues that surround it, clarifying this investigation's underlying philosophical assumptions.

1.1 – Capturing the phenomenology of the phenomenon

Generating examples of things that we *do together*, in the sense we are here interested in, is not a difficult task. Michael Bratman, for example, provides the following list, “You and I might sing a duet together, paint a house together, take a trip together, build something together, or run a give-and-go together in a basketball game”.⁴ Similarly, Margaret Gilbert picks out, “... picking mushrooms together, going for a walk together and travelling together”.⁵ I shall be seeking to understand in this thesis the phenomenon that is common in both these lists and, as indicated above, I will use the term 'collective action' to refer to it – the appropriateness of this term will become apparent in what follows.⁶

The ease with which both Bratman and Gilbert are able to construct their lists arises from the everyday ubiquity of our expressing ourselves in ways that imply the existence of collective action. Of course, as Donald Davidson rightly notes, “[o]ur ordinary talk is studded with metaphor, ellipsis, easily recognised irony, and hyperbole, not to mention slips of the tongue, jokes and malapropisms.”⁷ Given this it would be wrong to overly boldly assert, without pause, that we *really* mean to evoke a distinct type of action. Such hesitation is, however, easily overcome, for it is not just the formal structure of such sentences that points to this phenomenon, but also the *actual experience*, the *phenomenology*, of using such expressions and being part of the situations they describe. It is clear that when we use an expression such as “we painted a house”, most of the time we do not feel

4 Bratman, 1999, p.93.

5 Gilbert, 1992, p.155.

6 The choice of this term should not be taken as implying any particular theoretical outlook. There is not agreement across the literature on one standard term and others, such as 'joint action', 'shared action' or 'plural action', may also be used, generally in the discussion of the work of others. Unless otherwise indicated, in this thesis all these terms should be all so be taken as theory neutral ways of referring to the same phenomenon.

7 Davidson, 2005, p.15.

ourselves to be making a joke; nor when we say such things as “we walked to the top of the hill” do we generally want them to be taken as mere metaphor. Rather, in most circumstances, we feel ourselves to be using these phrases in straightforward ways. *We really do* mean that *we* painted the house and *we really do* mean that *we* walked to the top of the hill.

Just as generating such lists is easy, grasping that the examples in them are instances of a common phenomenon needs no special training; it is an expression of the phenomenology of our everyday social lives. That is, it is an expression of the way things naturally appear to us to be. This is why Gilbert claims that “[i]n dealing with such lists the reader is ... not supposed to baulk at an 'and so on' at the end”, and that given this, the reader is thought to have “... grasped a concept or an intuitive principle of some kind linking all the examples mentioned.”⁸ In this case, with the lists of collective actions above, this does indeed seem to be the case. To acknowledge that such grasping takes place does not imply that it is easy to explicitly express the exact necessary and sufficient conditions that govern inclusion in such lists. As is the case with *all* such everyday ways of 'carving up' the world, there will be some examples that we are not sure of and some examples that we feel to be on the fringes of exemplifying the phenomenon. However, this does not detract from the fact that it is easy to see that some examples clearly fit and others clearly do not.

We can focus further in on our folk concept of this phenomenon by considering examples of kinds of actions that do not fit; seeing what they lack will move us towards appreciating what the examples in question share. So let us first take my solitary action of sitting here at this desk, typing this sentence. This clearly is not, at least not in the context in play, an action of the same type as painting a house with another person or jointly going for a walk. It might be thought that the reason such action feels to be of a different type is obvious; the examples of doing things together with other people involve other people, in contrast, the example of my lonely typing involves just me. Doing things together, it could then be suggested, is easily characterised: it is just any action that involves others. However, the involvement of others, whilst clearly a key aspect of collective action, is not – on its own

8 Gilbert, 1992, p.8.

– enough to distinguish collective action from non-collective action. There *is* a sense in which even my solitary typing *does* involve others. That is, it depends for its existence on the actions of others. I would not be able to sit here typing on this computer if the society in which I live had not nurtured its invention, manufacture and distribution. Likewise, if I was not in receipt of some means of sustaining myself, of obtaining food grown by others and such like, I would not be able to spend time pondering philosophical issues. One might think that my activity only has the meaning it has given that I am utilising a language created and sustained by my society. Further, one might think that the fact that I intend for others to read this thesis, and engage with its content, is an integral part of my activity. We might say then that all (or at the very least a great many) of our apparently non-social activities in this sense involve others; they *depend* on the actions of others existing as part of what we might call their background.⁹

So, dependence on others is not enough to characterise how it feels to act together with others. Perhaps this is not surprising as it is a very weak sense of 'others being involved in our actions'. A better candidate for the relevant sense of 'others being involved' is perhaps the stronger condition of *interdependence* between acts. That is, the requirement that all the acts in question are dependent on each other, rather than just that the act in focus is dependent on the preceding acts in its background. In this sense, our action would be an interaction and the distinguishing feature of the experience of collective action would be the existence of *interplay between the actions of multiple agents*.

Again this is taking us in the right direction, i.e. away from the isolated individual towards the socially integrated one. However, again, it cannot quite be the full story, for it fails to exclude examples that do not naturally fit into the lists given above. Take, for example, the drive to the park for the aforementioned football game that I plan to make

⁹ The notion of dependence here is a wide one, and it is possible to give stronger and weaker versions of the idea of individual actions depending on the acts of others. The notion of what we might call 'causal' dependence (e.g. I could not type if food had not sustained me) seems far more straightforward than that of what we might call 'conceptual' dependence (e.g. my action of typing only makes sense because of existence of a community of speakers of a common language). The causal claim is unproblematic but, below in section three of this chapter, I will suggest that on some readings the stronger conceptual claim is at odds with the project in hand.

tonight. This act depends on others' actions as a background condition; I could not drive my car if someone had not built it and, the fact that we are playing the game which we play, only makes sense given the existence of a community that sets its rules. However, my driving is also social in the stronger sense of involving interaction with others: to complete my driving action I will have to stop at the traffic lights to wait for other agents to cross and their actions of crossing will likewise be dependent on my stopping. In the same fashion, I will have to ensure that I drive at a safe speed relative to the driver in front of me and they will (hopefully) also be keeping an eye on my driving and adjusting what they do appropriately. This action, then, is social in a stronger sense than the sense in which my solitary typing is social. While my solitary writing merely depends on others (in the two senses set out above), the driving involves others in the sense of my interacting with them. And yet my act of driving is *not* like playing football, it is not acting *together* with others in a way that would fit into the above lists.¹⁰

What we can say so far is that both Bratman's and Gilbert's lists of activities capture a phenomenon that feels to be, both in use and contemplation of, distinct from merely depending on others, and also distinct from actions that merely involve interaction with others. Both of these types of actions might in some useful way be called social actions. However, collective action is social in a distinct and special sense. This strong sociality involves the additional element of feeling to be collective. The *feel-of-collectivity* has multiple elements and I shall explore these below. One overarching characteristic is that this phenomenon feels not just to involve the experience of seeing others as externally related agents, as in the mere interaction discussed above; rather, it appears to involve seeing *other agents as internal to the action*. What I mean by this is that the relevant agents are not experienced as merely performing separate, though mutually interdependent, individual acts. Instead, it feels to the relevant agents that they are *all together* the performers of a *singular action*. So when *we* sing a duet together, it is not that one person sings and the others are involved in that person's doing so, nor even just that there are many related acts of singing; rather, it is a *we* that is performing an act, a *we* that is singing. Likewise, when *we* paint a house it is not that we are each involved in assisting the others

¹⁰ Bratman makes a similar point with an example of two soldiers in battle (1999, p.95).

painting (though if we are successful then this will be true) but rather that *we* can rightly say that it was *us* that painted the house. This is why it makes sense to distinguish, as I did at the start of this chapter, the act of playing football from the act of sitting alone at a desk, by stating that while the act of sitting is mine alone, the playing football will be *our* playing. We may say, then, that doing things together, in the sense embodied in the examples above, feels to us to be doing things as a singular unit of sorts. This is what makes it appropriate to talk of doing things together as collective action, i.e. it presents itself to us as being *the action of collections of people considered together as a whole*.¹¹ This is the basic and foundational claim of this chapter. The ability to make sense of this, what we might call the *collectivity* of acting together, will be the key test of theories that purport to describe this phenomenon.

In what follows, I will proceed to break down this experience of the feel of collectivity into the following four elements: unity, collective intentionality, detachment and constraint. These, I will argue, are transparently part of our experience. While it is my claim that these features can be found in the way we talk about the things we do together with others, it is true that there are no neat phrases, least not in the English language, that precisely and unambiguously carve out these phenomena. This need not lead us to abandon the contention that understanding these features is part of our everyday conceptual toolbox, for, as Margaret Gilbert notes, “[w]e could possess a given concept without possessing some neat phrase or single term to express it.”¹² Hence, though I take the names I give to my four elements to reflect the common usage of those terms, it should be understood that I am using them here explicitly with the senses which I shall set out. In order to further overcome the ambiguity in everyday modes of expression I will also, on occasion, borrow from Raimo Tuomela the terms 'I-mode' and '*we*-mode'.¹³ Use of the term 'I-mode' will be

11 As Gilbert notes (2008, p.101, footnote. 10), those who work on rational choice/game theory also use the term 'collective action'. However, they (rather oddly) use the term to refer to *any* combination of individual actions, even including combinations of acts that need not even be the same type of action. It seems fair to treat such an unintuitive usage as a peculiar technical term and separate its use in that realm from my distinct use here. Gilbert discusses the connections between these two ways of using the term at more length in Gilbert, 2007.

12 Gilbert, 1992, p.11.

13 I take my usage to be akin with Tuomela's, though I do not wish to be wedded to the framework he constructs around his terminology (See Tuomela, 2007, p.46-64). I also adopt the practice of italicising '*we*' (where it is used to refer to a collective) throughout the thesis to emphasise that it is not being used

a shorthand way of saying that the subject in question is a singular individual, or where there are multiple subjects, where these are each individually being referred to as separate agents. Use of the term '*we-mode*' will be a shorthand indication that the subject in question is multiple individuals being understood together as a collective. Thus, 'I-mode actions' refers to those actions that are seen as being performed by individuals and '*we-mode actions*' refers to actions that are seen as being performed by collectives. Where examples are ambiguous, I will use these terms to distinguish the sense I am interested in.

Let me now propose the four examples that will act as my canonical cases for understanding the collectivity of collective action. I purposely choose a diverse set as examination of the differences between them will play an important role in clarifying our theoretical models as the thesis progresses. My examples will be a romantic couple feeling bound, a climbing duo ascending a mountain, a rugby team pushing a bus and a residents' association campaigning about a green space. I will take each example in turn as each highlighting one of the features of our experience of collective action given above.

1.1.a – Unity

There is a wide sense in which the phenomenology of collective action involves an experience of unity inasmuch as collective action feels to be attributable to one singular thing, i.e. the collective. Unity in this wide sense exists in all circumstances where we can think of sets of objects as having attributes that apply to them only taken together *as a unit*. For example, it is true to say that the set of those with red hair in Sheffield would fill over half of a medium-sized football stadium.¹⁴ This is an attribute of this set of people and is not true of any single member of that set. In this weak sense we experience all-the-red-heads-in-Sheffield as constituting a unit – that is as constituting something that we can attribute properties to as a whole. However, there is also a narrower, and stronger, sense of unity: the *feeling of being united, as in sharing a bond*. Such a feeling may or may

distributively.

¹⁴ Estimate based on percentage of redheads in the UK (6%), the population of Sheffield (551,800) and the average size of a medium capacity football stadium (50,000)

not exist in sets of people such as all the red heads in Sheffield¹⁵ but it will be my contention that it is a necessary element of being part of a collective that is capable of collective action.

The kind of unity that I will claim is apparent in all examples of collective action can be seen very clearly in romantic partnerships, and hence here I set out the example of a united couple as an illustration of this phenomenon. Romantic partnerships are often described as unions, particularly (though not necessarily) when accompanied by formal decrees such as that of marriage.¹⁶ When we think about being in a properly functioning romantic couple, it is clear that we think of the individuals involved as experiencing their relationship in terms of unity. This is a feeling of unity that is very distinct from the mere knowledge that they share common characteristics by which they could be grouped together and referred to as one. It is characterised, as Gilbert has noted, by general comfort in use of the plural pronoun '*we*', that is, by a general feeling of being united in such a way that plural reference is natural and ongoing.¹⁷ Of course, the reason we want, in cases of being in love, to be united with another person may be different to any other collective action. At least when love begins there is an overwhelming desire to be in each other's company, however, as love matures, the sense of the unity we have in mind here comes to the fore; it is the unity of wanting to count yourselves as bound together.¹⁸ As Andrea Westlund puts it "... lovers form a *we* in jointly constituting the subject of a variety of activities, projects, and goals, one of which is often the quite general goal of sharing a life".¹⁹

15 See next chapter for further discussion of the issue of the unity that is generated by mere set membership.

16 See Westlund, 2008, for an illuminating discussion of love as requiring union. I explore Westlund's claims about love and collective action, in particular in relation to her rejection of key elements of Gilbert's Plural Subject Theory (PST), in the article 'Love, Plural Subjects & Normative Constraint' (Kisolo-Ssonko, 2012). I will return to the example of love in the final chapter of this thesis (Chapter Six), where I will claim that we can better understand elements of the phenomenology of love using the modified version of PST that I will develop in Chapter Five.

17 See Gilbert, 1996d.

18 Robert Nozick makes the stronger claim that: "Love, romantic love, is wanting to form a *we* with that particular person to be the right one for you to form a *we* with, and also wanting the other person to feel the same about you" (1995, p.234). It is unclear to me that such a claim captures the multiplicity of meaning we attach to the term 'romantic love'; however, it does seem clear that the actual structure of the committed romantic relationships that we commonly form involves becoming the kind of unit that is set up to engage in collective action.

19 Westlund, 2006, p.5.

For Gilbert, the ease with which a romantic couple can refer to themselves as a *we* and with which they can share in an action, in addition to the contrasting difficulty of two strangers doing so, marks something important. It marks the feeling of unity which couples share as that which is existent throughout the experience of collective action. Gilbert stresses this very point by having us imagine a post-conference dinner where most people are newly acquainted with the exception of two, Tony and Celia, who are engaged to be married. After the main course is finished “...Tony asks Celia 'Shall we share a pastry?' Celia nods agreement. Then one of the other men, Bernard, turns to Celia, who is sitting on his right, whom he hardly knows, and asks 'Shall *we* share a pastry?’”²⁰ Gilbert, and I think rightly so, takes it that the invitation to share a pastry involves here more than merely two people each individually eating half of one pastry, rather it is a call to *do something together*, namely to collectively share in the pastry. In this context there seems to be something particularly presumptuous of Bernard's invitation as compared to Tony's. It is not the intimacy of the act itself, for we can easily imagine the pastry split onto two plates and thus the eating of each piece not involving any close physical contact. Rather, the issue is that the collective act of sharing a pastry implies a level of unity that Celia and Tony are, being lovers, pre-set up to be ready to engage in, whereas Celia and Bernard, mere acquaintances, are not. This does not mean that Celia is not free to take up Bernard's offer. She can choose to perform a collective act with him, and accept the unity this involves, without having had any prior relationship with him. Rather, it is just that without preamble, Bernard's offer seems rather forward.

It would be wrong to claim that all collective action requires a feeling of unity that is linked to an emotional relationship – as it is with a romantic couple. Perhaps in arduous collective acts, such as the act of the mountain climbers in the example that I will explore below, such a bond is likely to become this emotionally strong. However, in the cases such as playing football together the bond need not be emotional at all, and this seems even more likely when we look at examples of collective acts such as picking mushrooms or painting a house together with strangers. Rather, the point is that the romantic couples

²⁰ Gilbert, 1989, p. 175.

wear on their sleeve, their bond of unity; the bond that examination reveals to be a typification of the bond that exists for all collective actions. Love involves a feeling of unity that makes it an ideal set up for us being ready to share in collective actions. This reflects the more general point that for any set of individuals, if there is no feeling of bond at all between them, then it seems odd to describe any activity they are involved in as a collective action performed by them.

1.1.b – Collective intentionality

The second experience I will examine is that of collective intentionality. The following example illustrates the necessary inclusion of the experience of intentionality. By this I mean the necessary inclusion of purposefulness as an element of our experience of collective action. Further, it illustrates that this purposefulness must be directed towards the collective activity, and in this sense must be collective intentionality.²¹ Let us use a real-world example from 1953, the first ascent of the world's highest mountain, Everest, by Edmund Hillary and Tenzing Norgay. In describing the final moment of their epic journey Tenzing said, “*We* stepped up. *We* were there. The dream had come true.”²² Note that he does not refer merely to his personal goal being achieved, but rather he refers to the dream that was shared by them both; reaching the summit is the realisation of this shared goal. That Tenzing should be understood as truly experiencing their summiting the mountain as a collective act is reinforced by the fact, noted by Philip Ebert and Simon Robertson in an article on the philosophy of climbing,²³ that when they were asked to say who reached the top first, both climbers repeatedly rejected the question because they saw summiting the mountain as something that they had done *together*.²⁴

21 I do not mean use of the phrase 'collective intentionality' to imply that it *must* be the collective that holds the purposeful attitudes, i.e. that there is a collective agent. This will be one of the points of contention in the following chapters.

22 Ullman, 1956, p. 265 [*emphasis mine*]. It is possible that the 'we' in this statement could be a distributive I-mode reference to multiple individuals rather than a collective *we*-mode reference to the group. As noted above, expressions in the English language are often ambiguous, particularly when considered in isolation. However, if considering his statement in its context, it appears clear that he does mean to make the stronger *we*-mode claim.

23 Ebert & Robertson, 2010, p.102.

24 The quote which opens this thesis, at the beginning of the introduction (see p. 1), illustrates this (Ullman, 1956, p.263). The idea, which appears to be being expressed in that quote, that understanding the act as collective can be more fundamental than understanding its individual components, is something that I

The element of the example above that I am interested in directing our attention to at this juncture is that getting to the top of Everest is a clear illustration of an intentional act.²⁵ It is obviously not something that is likely to have *merely* happened – unlike say, a stone rolling down a hill or a stream flowing to the sea. Nor is it a simply mechanistic reaction to some external stimuli, such as the triggering of the air bags in a crashing car, or the blinking of my eye when there is a gust of wind. In contrast to these examples, the climbing of Everest was something that is done *with a purpose*.²⁶ Now the general question of exactly how we should understand intentional action, the necessary and sufficient conditions on the mental states of the agent and their causal relationship to the event in question, is the subject of much debate. Indeed, it has its own subject area: action theory. These questions do not need to be settled here. However, while it is hard to exactly characterise the nature of intentionality and its role in action, it is relatively uncontroversial that, as Alfred Mele and Paul Moser put it, “[r]emove intentionality altogether from intentional action, and you have mere behaviour: brute bodily motion not unlike the movement of wind-swept sand on the shores of Lake Michigan.”²⁷ This point seems equally well to hold for collective action, just as it does for individual action. Of course, there may be borderline cases where it is hard to discern if we really have true action, such as when I scratch my head without thought. However, we need not know where exactly to draw the line to see that getting to the top of Everest was not only clearly an act that felt to be

will return to in chapters Five and Six.

25 Confusingly, in action theory the term 'intentional' has a number of meanings. For the main I will be using it in what seems to me the most natural sense which we can roughly think of as 'with purpose or aim'. 'Intentional' can also be used to mean being about or representing something. There is a further term that is pronounced the same but spelled differently: 'Intensional' (with an s). This term is a function of having the truth of usage of terms represent things dependent on meaning and not just reference. Confusingly, all three are sometimes all relevant to single discussions for being intentional (with purpose) seems to involve intentional (aboutness) mental states and these seem to have the feature of being intensional (non-replaceable with referentially equivalent terms). For the purposes of this thesis, I can be taken to be using 'intentional' in the first meaning unless otherwise stated. Note that in adopting this preference I restrict 'collective intentionality' to a narrower meaning than philosophers such as Searle (1996, 2010) who takes collective intentionality to refer more broadly to any shared attitudes that are intentional in the aboutness sense (e.g. shared belief, shared emotion and shared intentions {in the narrower sense of intentions}).

26 We might equally well speak of it being done with a 'goal', or with an 'aim', or with an 'intention'. Such terms may well have subtly different implications and are certainly used to play different technical roles in certain theories. However, I will follow Gilbert (see 2006, p.122) in thinking that nothing important hangs on the difference between these terms, at least not at this point in the discussion.

27 Mele & Moser, 1997, p.223.

purposely willed but, given the extreme arduousness of the feat, it will have been one where the required high strength of will must have been very evident indeed. Though other collective acts – such as painting a house together, picking mushrooms together and playing football together – may not require such high commitment, what makes these events feel to be actions, as opposed to mere happenings, is that *they are experienced as having been purposeful*.

So, given the above, it seems that collective action must *feel* to be intentional. There is a further question about the subject towards which this intentionality must be directed. This can be seen by thinking about the absence of collective action in the following situation: in modern times the ascent of Everest has become a popular hobby for hyper-wealthy tourists, leading to crowds of people all independently trying to reach the summit. Each person is intentionally trying to reach the summit; each is doing what they are doing (plodding along, sucking on oxygen and occasionally digging in their ice axes) in order to achieve this intention. Their behaviours are all I-mode intentional. This leads to it being the case that, as a set, they all move upwards towards the highest point. However, this is not enough to be able to say that *together* they are summiting the mountain. There is something missing, something that would make them, as Gilbert puts it, “... partners and not just participants in the act of travelling together.”²⁸ Part of what is missing is that the intentionality in question is wrongly directed. It is directed towards the acts of each individual rather than towards a possible collective act involving them all together. In contrast, Tenzing and Hillary, whilst each of course likewise performing their own individual intentional acts, seem to embody something more; they seem to be *together* intending to reach the summit.

A last specification we can add to the character of the intentionality, on top of it being pointed towards the collective, is the question of where this intentionality must be seen to be realised. As we shall see, it also seems that it must – in some sense – be *the group* that is the holder of the relevant intentional attitudes. So much can be seen in considering the following: if I were to feel that there was a purpose in a road being blocked

28 Bratman, 1999b, p.94.

unknowingly by a crowd exiting a theatre, say that I was glad to be made late for work, this would not mean that they were thus performing a collective act of doing so. Just as, if I saw a purpose in some single stranger accidentally slipping over, then, this would not turn that individual's accident into an action. Collective intentionality must exist internally to the collective action; it must be the intentionality *of those agents who are together acting*. Whether this means that each of the group members must individually hold the same aim for the collective, or if there are other ways for them to together count as having the relevant collective aim, is a question that must be left until the following chapters. For now, before I move to the construction of a theory of collective action, we can say that, the intentionality feels like it must be collective intentionality inasmuch as it must be such that we can say of the participants that they *together* assign purpose to that which they are *together doing*.

1.1.c – Detachment

The notion of collective action, that is in fact the most robust in the phenomenology of our everyday experience, is of acting together within formal frameworks such as clubs, companies and even nation states.²⁹ We naturally say such things as: Manchester City won the premier league; Coca Cola sanctioned the killing of Trade Unionists in South America; and that the UK invaded Iraq. If anything it is even clearer in these cases than in small scale cases (such as two people going for a walk) that we are asserting collections of people considered together to be the performers of these actions. Whilst this fact is in some way acknowledged by most philosophers of collective action, its implications tend to be seen as peripheral, i.e. as an extension from the norm rather than as the strongest case. For example, Bratman says that in order “[t]o keep things simple”, he focuses on “... activities that involve only a pair of participating agents and are not the activities of complex institutions with structures of authority.”³⁰ Gilbert takes a similar line saying in *Walking Together: A Paradigmatic Social Phenomenon* that “... we can discover the nature of social groups in general by investigating such small scale temporary phenomena as

29 Though the nature of our feeling regarding nation states tends to vary according to our political outlook, a point which will be explored in the final two chapters.

30 Bratman, 1999b, p.94.

going for a walk together.”³¹ That is not to say that either of them deny the possibility of any large and complex groups. For example, Bratman states that “[s]uch shared cooperative activities can involve large numbers of participating agents and can take place within a complex institutional framework”.³²

Large and complex groups are the clearest expression of the phenomenon I will now be exploring, that of detachment. However, while the phenomenology of collectivity becomes stronger when we focus on macro examples (i.e. the experiences of unity, collective intentionality, detachment and constraint are even more clearly felt), the problem is that as we scale up the groups there is a greater theoretical temptation to dismiss the reality behind these experiences – as Hans Bernard Schmid has rather memorably phrased it, to “... cry ontological bullshit”.³³ We can leave to one side the truly massive groups for now – I will return to them in later chapters – as, though most keenly felt in very large groups, the experience of detachment exists in all those collectives capable of collective action. So, let us take the following smaller scale example: in my local area there was a proposal by the council to build a football facility on half of the local green and many local residents objected to this, coming together to form a group the ‘Greenbank Residents’ Association’ (henceforth shortened to ‘GRA’) to campaign against this. While each of us who came to form this collective were united in our opposition to the development, each of our reasons for doing so were not identical. Some focused on the increase in traffic that the development would bring, whereas others were most concerned about the loss of purely non-commercial space. We all agreed that the building of a clubhouse and car park would destroy green space that we felt should remain. In order to produce a united and effective campaign we concurred that whatever our personal differences we should agree on a stance to be taken by us collectively as the GRA. This we did by meeting together, discussing our various opinions, forming a rough consensus and endorsing by majority vote this consensus as *the stance of GRA* as to the ills of the development. We also endorsed ‘stopping the

31 Gilbert, 1996b, p.178.

32 Bratman, 1999, p.93.

33 A term used by Schmid in his talk on corporate entities at the 8th *Conference on Collective Intentionality* (the abstract for his talk can be found on CIVIII, 2012). Interestingly, in his talk Schmid discussed empirical data that suggests that the likelihood of rejecting the agency of large organisations varies between cultures.

development' as the main goal of our group. Though each of us felt we were part of the collective actions that followed, and we all endorsed the collective perspective that we had agreed as collective justification for this action, none of us felt that the group perspective was the same thing as our own perspective. In this way, we experienced the group perspective as something other and removed from us as individuals. We can say then that we saw our perspective and the group perspective as detached from each other. As Gilbert puts it, collective goals "... exist at a different 'level' to personal goals."³⁴

The separation between the perspectives of the individuals and those of the collectives of which they are members can be seen most clearly when we think of cases of compromise. For example, situations where each individual puts forward their personal perspective for consideration but – through a process of discussion – a group position is arrived at and endorsed by a majority vote that amalgamates these perspectives but mirrors none of them. Further, it can be seen even more clearly in cases where the majority elect a small 'sub-group' to deliberate separately from the mass of members but have authority to set the perspective for the group as a whole. While such cases make the phenomenon of detachment most vivid, even in cases where the collective perspective that underlies the collective intentionality of a collective action is in complete harmony with all of the perspectives of the individuals involved, they will not be experienced as *necessarily* identical. Rather, the possibility of them coming apart will always be there. This is what I understand as the experience of detachment; it includes both the experience of *actual* difference between perspectives and also the experience of the mere *potentiality* of such difference arising.³⁵

1.1.d – Constraint

The final key aspect of our experience I am interested in is a phenomenon which makes entering into the unity of collective action something that is not always comfortable.

³⁴ Gilbert, 2006, p.123.

³⁵ This line of argument takes its inspiration from Gilbert's various examples where she raises the possibility of difference between the sum of individual perspectives and the group perspective. See for example (Gilbert, 1987, p.190) where she discusses a poetry group forming a collective position on the aesthetic judgement of a poem. I expand on use of the term harmony in Chapter 4 of this thesis.

It is the fact that we experience collectivity as constraining what we – as individuals – can rightfully do. The possibility of constraint, as we shall see, flows from the fact, discussed above, that the collective intentionality of collective action is experienced as being detached from each individual's personal perspective. This means that there is a possibility that they can be involved in collective actions in which what they – as individuals – intend to happen does not align with what everyone together – as a collective – intends to happen. In situations such as these, the purpose of collective action, and that which must be done by the individual in order to play their proper part in the collective achievement of this, appears to that individual as an *external burden*; i.e. something that they do *against their individual will*.

Take the following example: the East Midlands rugby club, the 'Leicester Tigers', are on their way to a match against their northern rivals, the 'Newcastle Falcons' and their coach breaks down at the bottom of a hill. Knowing that there is a garage at the top of the hill, the Leicester Tigers form the collective aim to together push their bus up the hill. Here, it seems that each player is obliged to put their back into it and participate in getting their bus to the top of the hill. Further, this obligation does appear to stem from their team membership. So much can be seen by imagining the following example: suppose that a bystander, Sue, sees the collective endeavour that the rugby team are engaged in, and offers to help. Imagine that, Sue starts pushing the bus but after half an hour she says that she has to leave (perhaps because she is late for another engagement). Suppose that, in response the flanker Bob also stop pushing and says, "If Sue's stopping then I don't see why I should continue!" Such behaviour might elicit all kinds of responses from the team captain, some too rude for print. However, it would be unsurprising if the content of his response, whatever its tone, included the fact that Sue is not part of the team and thus is at liberty to make her own mind up when to continue or stop. Bob, in contrast, does not have this liberty; he ought to do what they agreed to do together. It feels to be important that we make sense of our criticism of Bob by reference to the collective aims of the group to which he belongs – Bob faces pressure to play his part in what the team intends to do, regardless of his personal preference, because he is a team member. In this sense he feels to be constrained by the collective perspective.

The scenario just painted might strike those who have ever taken part in an arduous collective task (and I imagine pushing a bus up a hill is an arduous task indeed) as far too harmonious. In the picture painted above the participants accept the collective goal and try to achieve it. However, complexities arise when we consider what happens if there are competing visions of what the team should do. Imagine, for example, that after the endeavour has gone on for half an hour the team captain, Jim, calls a stop to the activity. “This is far too risky to our health”, he might say, “we should stop!” The team manager might take a different view and a row might break out. The dynamics of such conflicts, and our experience of them, will form an important part of my examination of competing visions of the nature of collective action in coming chapters. For now we need just note that the possibility of such conflicts does not point to the non-existence of collective constraint, but rather it is an affirmation of its existence; if each individual experienced no such constraint, they would merely experience themselves as going along or not going along. They experience themselves as rebelling and their co-participants experience themselves as having the standing to rebuke them for such rebellion because they feel themselves to be grappling with the constraint of the collective perspective.

1.2. – The phenomenological state of play

So a summary of where we have come so far: we can easily generate examples of acting together in a strong sense, and for the sake of clarity we can refer to this as collective action. Our phenomenal experience of our social lives presents collective action as a distinct kind of social action, distinct in its sociality from merely relying on others or interacting with them. Further, it presents collective action as involving:

- Unity* – the feeling of being bound as one
- Collective Intentionality* – the feeling of purposefulness at the level of the collective
- Detachment* – the feeling of separation between the individual perspective and the collective perspective
- Constraint* – the feeling of being under the command of the collective perspective

A note of caution concerning the strategy of using the phenomenology described above as a test which a successful theory of collective intentionality must pass: it is not always easy to disentangle one's assessment of how things appear from the theoretical positions that one holds about the subject in question. Thus, I accept that there is no such thing as a completely plain presentation of the phenomenology. With this in mind, it may be best not to see the four elements of the experience of collective action, set out above, as being set in stone tests for theories to pass, but rather as signs pointing towards the kind of theory that would be adequate. On this picture, the final test of a theory will not be whether it necessarily generates these phenomena, rather it will be whether it can make sense of them within a broad theoretical framework that is both coherent and fits with our wider understanding.³⁶ In introducing the terms 'collective action' and 'collective intentionality', it might be thought that I have already ruled out individualist readings of the phenomena of doing things together. I am claiming that we must say that our experience presents the social world to us as seeming to involve collective acts.³⁷ However this does not rule out the individualist possibility that we can ultimately reduce the process that produces this experience to a complex interaction between individual acts – i.e. it does not rule out that the collective element is a kind of illusion. That said, I do take it that I have ruled out the following approach: it might be thought that a commitment to the primacy of individual agency means that we should look to combinations of individual actions as the *starting point* for investigation. The problem with doing this is that it risks losing sight of the actual data that we should be explaining; it risks losing sight of our *actual experience* of the social world.

In the following chapter, I will look at what kind of account of collective action we need to be able to meet the challenge of explaining these experiences. In particular, I will

36 This will be why, though I will use the phenomenology presented here to motivate my rejection of Bratman's SCA account (Chapter Two) and my preference for Gilbert's PST (chapters Three and Four), I will ultimately propose that the stronger support for PST over SCA comes from the possibility of aligning Gilbert's theory with a wider account of agency (chapters Five and Six).

37 I thus stand in opposition to philosophers such as Michael Keeley who, in his paper 'Organizations as Non-Persons', presents the idea of 'collective action' as a kind of mystical philosophical thinking invented by philosophers to try to make sense of corporate responsibility, but, in fact, counter to the common sense view (Keeley, 1981).

put Bratman's shared cooperative action account to the test. Before I begin this, however, a survey of the context surrounding the philosophical investigation of collective action will help us further get our bearings as to this investigation's starting point. I will give some reasons why we might think that collective action, as I have characterised it here, is an important part of our social life.

1.3. – Why should we care?

So the phenomenology of our everyday social lives presents collective action to us as involving a real sense of being united together as a *we*. That's all well and good, but why should we care about this? Well, we might think that understanding a phenomenon that is embodied in our everyday experience is an important enough philosophical end in itself. Putting this to one side, there are wider philosophical issues, topics that philosophers already care about, that are affected by our understanding of the nature of collective action. It is not possible to give an exhaustive list of possibly related issues here, especially before we have a full theoretical account of the phenomenon in question. However, in this section I will outline some of the wider questions about society and our social lives that are related to the question of the character of collective action. I will be returning to some of these topics at the end of this thesis, to ask what the particular conception of collective action I develop has to say about them, and I will return to the others as future avenues with which to continue my research.

Arguably the idea that there is an important sense in which we do things together, and that we should care about understanding this, can be found lurking beneath the surface of much historical philosophy. For example, when Rousseau in *The Social Contract* talks of such things as men “...uniting their separate powers” such that they “... are directed by a single motive and act in concert”,³⁸ it is possible to read him as advancing a theory of *what it means* for individuals to act together, as well as a political theory about the moral legitimacy of such unions. No doubt a historical survey of the importance of collective action in different philosophical movements would throw up many such examples and

38 Rousseau, 2004, p.14.

would be a valuable endeavour. Unfortunately, while it is possible to interpret many works as making reference to collective action, these references are mostly deeply buried and uncovering them would require exegetical work beyond the scope of this thesis. Given this, this thesis will mostly be directed at the modern move towards a focused and analytic analysis of collective action. While there is some work in the 1970s that matches this criterion, such as Rolf Gruner's 1976 piece 'On The Action of Social Groups', in which he noted that philosophical work on "... what kinds of entities can and what kind cannot be said to act has hardly been considered",³⁹ most of the developments have taken place within the very recent emergence of the subject area that has been dubbed 'Social Ontology'.⁴⁰ As stated above, I will be most closely interested in the work of Margaret Gilbert and Michael Bratman. However, in terms of the question of the wider implications of an understanding of collective action, it is important to acknowledge the role of John Searle. It was Searle's 2001 plea for "... a branch of philosophy that in English speaking countries does not yet exist ... centring especially around questions of social ontology"⁴¹ that can be thought of as setting the tone of much of the debate and as popularising referring to it as social ontology.⁴²

Searle's main interest in social ontology is in seeking to understand the sense in which social facts are about real things; he thinks both that they are and that we can understand this by seeing them as being constructed by collective acts. Searle believes that if we look at the social world we find "... a class of entities that are objective, such as money and nation states." They are objective because "[i]t is not just a matter of my opinion, for example, that this piece of paper is a twenty-dollar bill; it is a matter of objective fact". However the puzzle is that "... these institutional facts exist only because of our subjective

39 Gruner, 1976, p.443 Also of note as an early example of the kind of focused work on collectives that I will be focusing on is Anthony Quinton's "Social Objects" (1975/1976) and David Copp's "Collective Actions and Secondary Actions" (1979).

40 Though note that this area is still sufficiently neglected for Christian List and Philip Petit to remark, in a similar fashion to Gruner, that "Despite their foundational place, however, the questions have received surprisingly little attention in recent philosophy and the methodology of the social sciences." (2011, p.2).

41 Searle, 2001, p.15. In further work Searle makes the bold claim that "It is an odd fact of intellectual history that the great philosophers of the past had little or nothing to say about social ontology [e.g.] such figures as Frege, Russell, and Wittgenstein, as well as Quine, Carnap, Strawson, and Austin" (p.6, 2010)

42 Gilbert gives a useful survey of the initial emergence in the late 1980's and 1990's of this "important new turn" (as she calls it) in her *Two Approaches to Shared Intention: An Essay in the Philosophy of Social Phenomena* (2008).

attitudes”.⁴³ How can we understand this apparent objective/subjective duality of social facts? Well, for Searle, the solution is that the social world presents itself to us, as individual agents, as being objective because it is not the product of the subjective attitudes of individuals, but rather of collective action. Because of this, facts about it are experienced, by each member of the society, as being what Searle calls 'epistemically objective' – that is their truth is not up to any singular agent to decide – but they remain 'ontologically subjective' – that is, their truth is up to us together.⁴⁴ So what does it mean for something to be 'up to us'? The story Searle tells is that it is a matter of us *doing something together*, namely of us collectively assigning status functions. So, for example, that David Cameron is the Prime Minister is a matter of us together assigning the status function of Prime Minister to his person. It is a matter, so to say, of a *collective act* of assigning that status function to him. The apparent objective/subjective duality of social facts arises then because the social world is a product of collective action. Clearly then, if we are to understand the social world in general, we are going to have to first understand the nature of collective action.

We can say more; the assumption is that social life is not just created by us in some accidental way as the mere result of our behaviour, in the way a valley is created by the mindless passing of water along a river. Rather the social world is thought to be created, or at least in large part, by what is going on in the minds of the people who are its members. In this sense, social ontology is what Gilbert calls an *intentionalist project*.⁴⁵ That is, as Schmitt puts it, social ontology assumes that “... collectivity phenomenon like groups, joint actions, and joint attitudes must be characterised in part in terms of intentional attitudes and their contents.”⁴⁶

43 Searle, 2010, p.18.

44 Searle, 2010.

45 Gilbert, 1992, p. 128.

46 Schmitt, 2003, p.22. It is important not to confuse the project of examining the construction of social facts with the, superficially similar, philosophical project of *social constructivism*. Searle devotes a whole chapter to dismissing social constructivism (1996), as does Gilbert (1992). Social constructivism makes the claim, as Goldman puts it, that “... truths or facts are not in or of the world; they are not ‘out there’ to be discovered but are mere social fabrications or constructions” (Goldman, 2010, p.3). As Schmitt notes this theory is “... incompatible with the claim that intentional attitudes and their contents depend conceptually on collectives. For the two together give rise to a circularity in the characterisation of collective phenomena.” (2003, p.22). Denial of social constructivism is not to deny the obvious truth that our actions (both individual and collective) often are dependent on the prior actions of other individuals, as explored above with the example of my solitary typing. However, unless we can make sense of the formation of individual intentional attitudes separate from already understanding collective attitudes we

So, understanding collective action looks like it is going to be important for understanding the ontological status of the social world, that is, for understanding the structure of the social world. A further way that it may well be important is in understanding the interactions between agents that occur within this structure. This is most clear in the case of considering the rationality of the actions that take place within social settings. Famously, Hume described the following dilemma involving two farmers:

“Your corn is ripe today; mine will be so tomorrow 'Tis profitable for us both that I shou'd labour with you today, and that you should labour with me tomorrow. I have no kindness for you, and know that you have little for me. I will not, therefore, take any pains on your account and shou'd I labour with you on my account, I know that I shou'd be disappointed, and that I shou'd in vain depend on your gratitude. Here then I leave you to labour alone: You treat me in the same manner. The seasons change; and both of us lose our harvests for want of mutual confidence and security.”⁴⁷

The problem is that to secure the best outcome for both of them each farmer must trust the other. However, whoever goes first will know that the most rational thing for the other to then do will be to renege on the deal and refuse to do their part. The generalised version of this problem is known as the *dilemma of rational cooperation*: in its most broad form it asks, “how can we make rational sense of a cooperative society if rationality would always have the agent defect when it is in their interest to do so?” Now, there have been many attempts to solve this problem. Some of these do so by changing what has been called the 'pay off structure' of the different options. This means changing the examples such that the value that each agent places on the different outcomes makes it rational for them to cooperate, for example, by introducing moral feelings that make it so that any agent who fails to cooperate will feel so bad at having harmed the other that this outweighs any benefit they might gain from failing to play their part. However, one might feel dissatisfied with a solution that requires changing the individual pay-off structures, as it feels as if there ought to be a way for agents to count as rational in cooperating *even* with the pay-offs as they are. One interesting solution, from the perspective of our current investigation, which seems to

might fear that we will end up with a vicious circle.

47 Hume, 1739, book III, part II, Section 5.

work without simply changing the pay-offs, is presented by Martin Hollis in his book '*Trust Within Reason*' – he calls it the “team solution to the problem of trust”.⁴⁸ His solution is, on face value, brilliantly simple: “The farmers both get their harvests in if they can trust one another to *play as a team*”.⁴⁹ If we imagine the farmers seeing themselves as each participating in performing the collective action of harvesting all of the corn, then it is clear that from the collective perspective what is rational will be each agent 'playing their part', for that is what is needed to complete the collective goal. The viability of this solution, and the question of the relation between the individuals and their team can only fully be addressed if we can answer the question of the nature of collective action.⁵⁰

Trust is clearly an important aspect of social interaction. This is also true when we think about epistemic interaction, that is, when we think about gaining knowledge from the testimony of others. When we learn things by being told them by other agents, there seems to be something importantly different going on compared to when we acquire knowledge from the information gleaned from non-agentive mechanisms, such as a reliable clock or thermometer. Trusting an agent appears to have a distinct epistemic character from taking a non-agentive device to be a reliable indicator. The relevance of the investigation of collective action to this area of epistemology is that, as Miranda Fricker has pointed out, we do not just trust in the testimony of other people but also in that of collectives.⁵¹ That is, we treat the apparent pronouncements of groups, particularly institutions, not just as signs pointing thermometer-like towards possible truths but we treat such pronouncements as *collective acts of telling* – to be believed (or not), rather than simply as being relied upon (or not). The fact that we understand collective testimony thus is arguably an important part of our taking ourselves to have the kind of socially mediated knowledge that we generally take ourselves to have, for example, that I take myself to have come to know that Iran sent a monkey into space through trusting the BBC's report that this was the case.⁵² It may also, as Fricker claims, thus be an important part of the

48 Hollis, 1998, p.137. Team solutions are also explored by Sugden, 1993 and Bacharach, M. 2006.

49 Hollis, 1998, p.137 [*emphasis mine*].

50 As Hollis notes, one particularly thorny question is: “Does success mean that the farmers have chosen rationally or the team has?” (1998, p.142).

51 Fricker, 2012.

52 BBC, 2013.

functioning of a democratic society.⁵³ However, all this is predicated on the possibility of making sense of collective testimony. If we can understand collective action, then we will be able to get a handle on the possibility of the *collective act of telling*. The understanding of collective action we come to have will affect our understanding of the testimony and trustworthiness of collectives and the relation of these to the testimony and trustworthiness of the individuals who are part of such institutions.

The most ambitious use of understanding of collective action, which can be seen as an extension of the team solution to the problem of trust perhaps, is Gilbert's attempt in *On Political Obligation* to try to make sense of the political obligations we have, in terms of being part of collective acts at the state level. She also extends this to the possibility of our feeling collective guilt over the acts of the states we belong to.⁵⁴ In doing this, we can see her as trying to complete Rousseau's project of understanding the foundations of political society. However, while Rousseau starts with political analysis, Gilbert starts from an understanding of what it is for agents to act together. I think that Gilbert is right to believe that her account of collective action ought to be able to be extended to the scale of state actions. However, some of the things she says seem to go against our general experience of political obligation, and I will later argue that our experience of political obligation is better explained by the modified Plural Subject Theory that I will attempt to construct in chapters five and six.

1.4 – Chapter Summary

In this chapter I have tried to capture the phenomenal experience of doing things together, in the sense that is implied by sentences such as, “*We* painted the house”, “*We* went on a walk”, “*We* summited Everest” and “*We* booked the holiday”. I have argued that the things described in these sentences are best referred to as 'collective actions' because

53 She says that, “[w]ithout [the] possibility of institutional testimony, and the second-personal relations of trust that are required for it, the democratic ideal of how institutions make themselves accountable to the citizens they serve, and the collective understanding of what is at stake in institutional truthfulness, would be very much diminished.” (2012, p.28).

54 Gilbert, 2006. This extended discussion of political obligation is foregrounded by her earlier short sketch of the same position. (1996e)

they present themselves as being performed by collections of people considered together as a unit; furthermore, that they involve a feeling of being bound as one, the feeling of purposefulness at the level of the collective, the feeling of separation between the individual perspective and the collective perspective and the feeling of being constrained by the collective perspective. The question I will move onto in the next chapter is that of what theoretical framework we can use to make sense of this phenomenology.

Chapter 2 – Mere Sums

The term 'collective action' is sometimes used in a narrower sense than I have been using it here, to refer particularly to actions such as industrial disputes that seek to defend the interests of large groups of workers.⁵⁵ However, as we have seen, there is a wider sense of the term that refers to any act which is performed *jointly with others*. In the last chapter, I established that collective actions of this type share the fact that they are experienced as involving the following phenomenological elements: a feeling of being bound as one (*unity*); a feeling of purposefulness at the level of the collective (*collective intentionality*); a feeling of separation between the individual perspective and the collective perspective (*detachment*); and a feeling of being under the command of the collective perspective (*constraint*). I take these four experiences to delineate the sense in which collective action feels to be strongly *collective*. That is, the sense in which it is experienced as a *we-mode*, rather than an *I-mode*, phenomena.

Let us imagine that it is a typical sunny English summer's day and two walking companions are setting out to walk together to the top of Scafell Pike.⁵⁶ According to the picture I have painted, these two ramblers will feel themselves to be united as a pair with the collective aim of getting to the summit. At the same time, however, each will also feel that this collective act is something they are individually detached from and something that can potentially constrain them. Of course, noting how things are experienced as being does not necessarily tell us how they actually are. In this chapter, I will shift from setting out the phenomenology of collective action to examining its substantive ontology; that is, to asking how we might best understand the actual structure of collective action that underlies its appearance. I will begin by setting out some theoretical considerations that might incline us towards what I will call the *framework of individual agency*. This framework restricts us to individualist accounts; that is, it restricts us to accounts that seek to reduce the apparent *we-modeness* of collective action to a sum of individuals' *I-mode* characteristics. Such an

55 Additionally, as noted in the last chapter (see footnote 11), it is also used in a completely different sense by game theorists.

56 A modified form of Gilbert's example (1996, pp. 177–94).

approach faces a number of challenges that motivate modification of the simple individualist account. This modification leads to a more complex version of summativism, one that allows the activity component of collective action to be joint, but still claims that the intentional character of collective action is given by the sum of the intentions held by individuals. Further modifications will be needed to deal with what I will call the *problem of control*. This problem centres around the question of how joint activity can be jointly under the control of multiple individuals. I will explore why we might think that such control is a necessary part of collective action; why it could be problematic to conceive of multiple individuals as having such control; and I will consider Bratman's *Shared Cooperative Activity* (SCA) account as a solution to this problem. I will also explore the fact that for Bratman the nature of the 'interdependence of purpose' (that is at the heart of collective action on his view) implies a commitment to the joint activity and to mutual support in achieving it.

On Bratman's SCA account, collective action consists of joint activity governed by the compatible and interdependent plans of the individual group members.⁵⁷ In addition, on this account this interdependence must be common knowledge to these group members.⁵⁸ So, for example, for us to jointly walk to the top of a hill is for us to satisfy the following conditions: we must engage in the activity of jointly walking up the hill; this joint walking must be led by our individual plans to so walk; and our following of our own plans must be both in accordance with and because of the (commonly known and compatible) plans *held by both of us* to so walk. By requiring that our plans are compatible, Bratman does not mean to say that they must match. Rather, as he puts it, just that they must mesh. That is, that they fit together not only at surface level but also at the level of the subsidiary plans.⁵⁹ For example, if one of our walkers intends that they together walk up the hill using the high path and the other has the aim that they use the low path, then there can be no collective act, for such activities are not co-possible. However, it is unproblematic if one

57 This account is explored throughout Bratman's work, but see in particular 1999b, 1999c, 1999d & *forthcoming*.

58 In this context, the term 'common knowledge' has a meaning that is more specific than, but related to, its common use. Its exact use here is explored below.

59 By subsidiary plans, it is meant the additional plans we must make to fulfil our main goal.

intends that they walk slowly and the other intends that they admire the flora along the way, as these plans are not in conflict.

The conclusion of these investigations will be that, by requiring complex ties of intentional cooperation and inter-reliance, Bratman's SCA manages, within an individualist framework, to give partial explanation for aspects of the feeling of collectivity. However, I will claim that, hampered by this framework, it fails to explain the richly collective nature of our experience of collective action. In particular, it struggles to give an account of our experiences of detachment and of constraint. In highlighting the limitations of Bratman's account, my goal is not to provide a knock-down argument against any possible summative account. Rather, my goal is to expose the weaknesses inherent in such accounts, in order that I will be in a position to argue that a non-summative account can do better. In doing this, I seek to open up the space for an account that takes our experience of collectivity seriously, namely Margaret Gilbert's Plural Subject Theory – a task I will take up in the following chapter (Chapter Three).

2.1 – The framework of individual agency

A seemingly straightforward explanation, for the collective aspects of our experience of collective action, is the claim that it arises from such action being the action of a distinct and separate *collective agent*.⁶⁰ The problem with such a claim is that it is immediately vulnerable, as Thomas Smith says, to “... an accusation of metaphysical extravagance”.⁶¹ Do we really want to invoke, as John Searle puts it, “... some Hegelian world spirit, a collective consciousness, or something equally implausible”?⁶² If we resist any literal notion of otherworldly collective spirits, then an alternative way to read such a proposal is in terms of what we might call the *organic thesis*. According to this thesis, just as individual human beings are organisms that arise from the combination of molecules, social entities are

⁶⁰ Margaret Gilbert's Plural Subject Theory, which I will be examining in the following chapter, also claims that there is a collective agent. However, as we shall later see, it differs from the account in question, as it does not claim that this agent is a separate social organism. The rejection of the thesis that follows (the organic thesis) is thus not a rejection of Gilbert's proposals.

⁶¹ Smith, 2005, p.76.

⁶² Searle, 1996, p.25. When presenting this account of collectivism to my peers, those who study Hegel tend to reject that he held any crude notion of a separate world spirit.

organisms that arise from the combination of individual people. Along these lines, when we talk, as in one of the examples from the last chapter, of the Leicester Tigers' pushing of their bus up the hill, we are invoking a singular action of a singular creature – the action of an organism called 'The-Leicester-Tigers'.

Such an explanation has some minimal plausibility, in particular when looking at the actions of formal organisations, such as sports clubs. However, as Andrew Vincent notes, “[i]t is difficult to find theorists who actually take the organic thesis with complete seriousness, in the sense of actually identifying groups as organisms.”⁶³ While it might be arguable that groups can count as having some of the kinds of properties attributed to organisms – such as being responsive to external stimuli, being capable of reproduction and growth, and maintenance of homeostasis⁶⁴ – at best this might mean that a case could be made that they mirror simple creatures, such as bacteria. It seems doubtful that they might be the kind of organisms that can perform *complex intentional actions*.⁶⁵ Actions such as planning holidays, climbing the world's highest mountain, moving buses to places where they can be fixed, and opposing development of prized green land, all seem to require such a complexity of mental representation that it is hard to imagine could be independently realised by the structure of a group considered as a separate organism.

Even if there were no difficulty in identifying groups as the kinds of complex organisms independently capable of complex intentional action, the organic thesis goes against the spirit of our understanding of the relation between individuals and collectives, as set out in the preceding chapter. The notion of a separate social organism is at odds with the underlying assumption of intentionalism that underpins the current study. Intentionalism, as we can recall from the previous chapter,⁶⁶ is the idea that social facts are

63 Vincent, 1989, p.698.

64 It is worth noting that the idea that a collective could be an organism is not ruled out *merely* by the fact that it would be an organism composed of parts that themselves count as organisms. Indeed, the human body contains trillions of micro-organisms and they outnumber human cells by 10 to 1, though being much smaller they only take up 1 to 3 percent of the body's mass. (National Institute of Health, 2012).

65 Whether something is complex is, of course, relative. My action of taking the milk from the fridge to put in my tea might not be considered complex compared to the act of playing the violin, say. However, in comparison to an earthworm's moving away from the light, it is indeed complex. I use the phrase 'complex intentional action' with the second kind of comparison in mind.

66 See Chapter 1, section 1.3.

generated from the intentional attitudes and behaviour of individuals. Collective action – we can also recall from the last chapter – appears to be the foundation of social fact. That is, it appears to be the mechanism through which social facts are generated. If we were to explain collective action by evoking a separate agent, then it would follow that social facts are *not* generated by the intentional attitudes and behaviour of individuals, but rather that they are generated by the distinct intentional attitudes and behaviour of this mysterious collective organism. Further, if we adopted the organic thesis, we would have a puzzle as to how individual agents would relate to this separate organism. Why, for example, if individuals were mere parts of some greater organism with mental attributes of its own, would individuals experience such a separate entity as constraining upon them, in the sense set out in the last chapter? While such worries do not rule out the organic thesis, paired with the concerns about complexity expressed above, they certainly make it an unattractive research avenue.

If the organic thesis is unattractive, then what is the alternative? In the next chapter (Chapter Three) I will claim that, in Margaret Gilbert's writings, we can find an account of a collective agent that does not require the existence of a separate social organism. In this chapter, however, I want to explore another possibility; that is, the claim that the *apparent* collectivity of collective action is generated by the *mere sum of individuals and their individual attributes*. Consider a group of leaves caught in the wind; as they fly around they appear to have unity, to move as one. There is an apparent collectivity to their movement; however, clearly this apparent collectivity is an illusion. The movement of each leaf is independent from the movement of the other leaves. They merely follow similar courses driven by the invisible wind. The simplest individualist account of collective action is one that makes its explanation akin to the explanation of the unity of these wind swept leaves. It would state that, whilst there is an appearance of action occurring in the *we*-mode, all there *really* is, is a set of I-mode actions related to each other, in such a way as to give rise to the appearance of collectivity.

This analogy is obviously only partial. There is much that is different between our experience of the apparent collectivity of the mass of leaves and the experience of the

apparent collectivity of collective action. As the detailed exploration of the phenomenology of collective action in the last chapter shows, the collectivity of collective action is multifaceted. While the leaves are externally perceived as belonging to a singular moving unit, true collective actions involve the agents internal to them having robust experiences of unity, collective intentionality, detachment and constraint. Further, while the apparent unity of the leaves can be explained easily, the apparent collectivity of the romantic couple, the mountain climbers, the residents' association and the rugby team is not so easy to dismiss. All of which is not to say that a reductive strategy is impossible, rather just that it will not be straightforward. However, there are some reasons why we might think that at least attempting such a reduction is worthwhile,⁶⁷ and I will explore these in what follows.

An account that seeks to reduce *we*-mode action to a sum of I-mode actions is a simpler kind of account because it posits fewer types of actions than one that takes the *we*-mode at face value. Rather than allowing for two types of action – individual and collective – such an account says that there is only one real type of action – individual. That a simpler account is to be preferred follows from an ontological version of Occam's razor; a theory that posits fewer kinds of things is, all else being equal, to be preferred to one that posits more. Of course, whether all else *is* equal is exactly what is up for dispute; however, the principle suggests that we would do well to at least start with the simplest account possible. As it stands, this is not motivation enough to follow a reductive strategy. This is because complexity can be measured in many different ways and there is a sense in which collectivist explanations are simpler. As Christian List and Philip Petit say, such explanations have “... greater descriptive and explanatory parsimony”;⁶⁸ that is, they straightforwardly explain apparent collectivity with no need for reduction.

If an appeal to simplicity is insufficient, we might look to the following for further justification. What motivates many attempts at reducing collective action to individual action is the idea that having rejected notions of social spirits and emergent social organisms the *only thing left* is the action of individuals. A reason to think that this is the

⁶⁷ And hence why, in terms of the overarching structure of this thesis, it is worthwhile examining Bratman's account before moving onto Gilbert's.

⁶⁸ List & Pettit, 2011, p.1.

case can be found in consideration of the fact that action requires intentionality. As we saw in the former chapter, this is what distinguishes it from mere happenings. However, intentionality appears to be a mental phenomenon and, having rejected the separate social entity thesis, it appears to follow that only individuals have minds. Thus, as Searle puts it, it must be the case that “... all intentionality takes place in individual brains.”⁶⁹ From this starting point, it is easy to think that we can move to the conjecture; if action requires intentionality and intentionality only takes place in the minds of individuals, it appears to follow that only individuals can perform actions. For want of a name, let us call these claims together the *framework of individual agency* (FIA).⁷⁰

In this form the FIA implies that there are *no* actual collective actions: there are just sums of individual actions. I follow Gilbert in using the term '*summative*' to describe reductive theories of this type.⁷¹ We can define a summative account as any that takes the properties of a group to be no more than the sum or aggregate of the individual properties of the members of the group. Conversely, a *non-summative* account is any that takes the properties of a group to be something more (whatever that 'more' may be) than the sum or aggregate of the individual properties. The organic thesis, discussed above, is a non-summative account, in that it claims that there is something – the emergent organic properties of the group as a whole – that is more than just a description of the sum of the properties of the relevant individuals.

With all the above in mind, we might construct what we can call a *simple summative account* of collective action. According to such an account, to say that collective action *a* was performed would just be to say that all members of the collective in question performed an action of the same type as *a*. Such an account is a useful starting point for constructing an explanation that is compatible with the framework of individual agency.

69 Searle, 1996, p.24.

70 My usage of this term follows Gilbert's (1989, p.12). In other places Gilbert has also referred to this as the *singularist assumption* (see, for example, Gilbert, 2008).

71 Gilbert takes the term 'summative' from Quinton, who claims that in some cases, such an account of social attributes is obvious; for example he says that, “To say that the industrial working class is determined to resist anti-trade Union laws is [obviously] to say that all or most industrial workers are so minded” (Quinton, 1975/1976, p.17).

However, as we shall see, as it stands, it is clearly inadequate. The question of what we need to add to it, to try to make it adequate, is that which will drive us forward, towards a more complex account.

According to this account, that which is said collectively of the group can be reduced to something that can be said distributively about every member of the group.⁷² At first, it appears that there are times when this account explains collective action. Unfortunately, examining these instances reveals them not to be instances of the kind of phenomena that is our target, i.e. collective actions. For example, if we were to say, “All of Manchester went shopping on Saturday”, we are likely to feel ourselves to be expressing nothing more than just the notion that each Manchester resident (or at least a great many of them) independently performed the action of the same kind – i.e. that they each performed an action of going shopping on Saturday. However, rather than showing the strength of a simple summative account, such cases highlight its failings. The appropriateness of the distributive interpretation in this case, and others like it, is down to the fact that the expressions in question are never meant to convey collective action. We can see as much by imagining a conversation that might follow. On hearing me talk of the whole of Manchester going shopping you reply, “That's amazing, the whole of Manchester went shopping *together*?” To which I might perfectly reasonably reply, “Oh no, I just meant that they each went shopping at the same time.”⁷³ As Gilbert notes, “[t]he sentence 'X and Y are doing A together' is susceptible to a weaker interpretation in which sharing in action ... is not involved”.⁷⁴ Instances where we give a straightforwardly distributive account of expressions may appear to have the form of attributing actions to social groups. However, though their form is similar to true expressions of collective action, they are not, in fact, meant to attribute collective action in the sense we are interested in here.

To construct a more plausible account of collective action, in what follows I will examine how we might adjust and augment the minimal summative account above in line

72 See Rolf Gruner for discussion of the distinction between 'distributive' and 'collective' reference (1976, p.445).

73 Example modelled on Gilbert, 1992, p.154.

74 Gilbert, 1992, p.154.

with the four elements of our experience of the collectivity of collective action, whilst remaining within the spirit of the FIA. Doing so will bring us to Michael Bratman's *Shared Cooperative Activity* (SCA) account, which, as I stated above, I take to be the most promising summative strategy to explain the apparent collectivity of collective action. However, it will also take us past Bratman's account, as such an account will be seen to be unable to sufficiently explain the experiences of detachment and of constraint.

2.2 – Mere sums -vs- the experience of unity

Let us start the task of augmenting the simple summative account by examining what we might add to explain the apparent unity that is experienced as part of collective action. Recall that I used the example of the lovers to characterise this experience; the unity they feel is that of being bound together. This appears to be the kind of experience of unity required for all collective acts. Acting together seems to involve the experience of being bound together in a sense which makes it appropriate to think of yourself together with those with whom you are bound as a *we*. The simple summative account allows for situations where there clearly is not such felt unity (such as the Mancunian shoppers). The question, then, is how we might adjust the account such that it does explain the existence of the feeling of unity?

One possible claim would be that for a sum of individual actions to count as collective action, the members of the collective must themselves share some uniting characteristic. As it stands, however, such a condition would be unclear. As already discussed (in the previous chapter) the mere fact of the relevant individuals being objectively groupable by virtue of some shared characteristic is not enough to generate a feeling of unity; e.g. the individuals who make up the set composed of all the red-heads in Sheffield may well feel no unity at all. Perhaps, though, while a sharing of a single characteristic is not enough, what we need is just more commonality. This, though, is not the case, as can be seen in the case of mountaineers summiting Everest. As noted previously the ascent of Everest has become something of a hobby for well-to-do tourists, with a (comparatively) large number of people all independently trying to reach the summit. On

such occasions, each person is intentionally trying to reach the summit and they make up a set of individuals with lots of shared characteristics; the climbers are all rich, they are all fit, they are all dressed in modern mountaineering clothing, and so on. Yet we still have no collective action; we just have a bunch of individuals swarming over the mountain. Further, adding a requirement for shared characteristics is in danger not just of failing to rule out non-collective acts, but also, wrongly excluding some acts that we *do* think are collective, for it does not seem to be the case that collective acts cannot take place when there is a huge diversity of individual characteristics; for example, the team that climbs Everest *as a team* might have diverse characteristics – be of different genders, different nationalities, different heights etc. None of this appears to block the possibility of their performing a collective action.⁷⁵

One thing that is clearly absent from the example of the individualist mountaineers is cooperation. Each is out for themselves and is not interested in helping the others to achieve their goals. Indeed, in a notorious case in 2006, an injured climber (David Sharp) was left to die by others who did not want to sacrifice their own achievement to try to help him survive.⁷⁶ The contrast to this strident individualism would be a situation where everyone was *cooperating* to reach the summit. Such considerations suggest an account of collective action where each individual is cooperating towards achieving their action. Now, in the last chapter I noted that mere interaction does not seem to be enough to generate the kind of felt unity in question, and thus that it is not enough to turn multiple actions into a collective action. There I used the example of driving down the high street, noting that this involves interaction and cooperation with others⁷⁷ but does not appear to be a collective act.

Perhaps though, what blocks the driving example from being a case of collective action is merely that the individual acts that are being achieved are all of different types. We

⁷⁵ Of course they must all share the characteristic of being part of the team that is climbing Everest, but this cannot be the characteristic we appeal to as constitutive of their unity, for this would be circular.

⁷⁶ Edmund Hillary said of the affair “I think the whole attitude towards climbing Mt. Everest has become rather horrifying. The people just want to get to the top. They don't give a damn for anybody else who may be in distress.” (McKinlay, 2006).

⁷⁷ In driving I have to stop at the traffic lights to wait for other agents to cross and their actions of crossing are likewise dependent on my stopping. I have to ensure that I drive at a safe speed relative to the driver in front of me and they must also be keeping an eye on my driving, adjusting what they do appropriately.

might adopt a summative account that stipulates that there must be cooperation between individuals that are all performing acts of the same type. However, this still does not provide a strong enough criterion to capture the collectivity of collective action. This can be seen by noting that a modified version of the selfish mountain climbers example above can fit this criteria, but still fail to be collective action. Imagine that, rather than being selfish, we make our climbers friendly and helpful: if they see that another is in need, they will throw them a rope; if they are informed on passing another that they are short on food, they will share their food; etc. Now this kind of cooperation is the type of thing that we would expect if they were all collectively summiting a mountain. However, we commit the fallacy of affirming the consequent if we think that such cooperation is thus *in itself* enough to mean that there is a collective action. Indeed, we can well imagine that though the climbers are transformed from selfish to cooperative they remain primarily focused on their individual achievements; they remain driven by their individual goals to each reach the top of the mountain themselves. While they are happy to help the others achieve their goals, which happen to be of the same type, it is only their individual achievements that define their actions. They would be helpful and unselfish soloists, but soloists they would remain.

What then is missing from the helpful soloists example? Whilst they are cooperating, they still have their own individual goals as the focus, the situation merely having been modified such that they are willing to help others achieve their own individual goals, and be helped to achieve their own in return. It seems that we need a stronger form of cooperation. We need cooperation towards the achievement of a *single joint activity*. This is the route taken by Michael Bratman. He believes that collective actions⁷⁸ have cooperation at their core, but that this cooperation must have a particular aim. For Bratman, what we need is that the participants are cooperating to share in the creation of a *joint activity*.

The notion of a joint activity is the notion of something that we are all doing together. In bringing it into play we lose something of the simplicity of the simple

⁷⁸ Note that, Bratman prefers not to use the term 'collective action', as it may be taken to imply a non-reductive analysis. However, in the sense I use it here, the phenomenon of collective action is the same phenomenon that is the apparent target of his investigations.

summative account, for we admit that there is something that must be considered at the level of the collective – i.e. the activity – however, as I will show below, we need not be thought to be abandoning the framework of individual autonomy altogether.

2.3 – Mere sums -vs- the feeling of collective intentionality

Having significantly altered the simple summative account, in order to allow it to explain unity, we come now to test our augmented theory against the feeling of collective intentionality. That is, test it against the feeling that collective action has purpose and that this purpose is directed towards the collective activity. It is in exploring a possible explanation of the nature of this collective intentionality that we come to see how we might (along the lines suggested by Michael Bratman) gloss the notion of cooperation, such that it generates activity that is distinctly collective.

So, above, we have already moved some way away from the simple summative account; from consideration of what it is that each individual is doing towards consideration of what everyone *together* is doing. If we examine the kinds of things that we think of ourselves as being able to do collectively, such a refocus becomes obviously correct. There are many things that we would claim are collective actions that *could not* be individual actions. For example, we might say that the members of the residents' association – from the example raised in the last chapter – together blocked a road as a form of direct action to try to save their park. Such an action would transparently be resistant to a simple summativist reduction as it cannot be that each individual is themselves blocking the road, for each is far too small to be doing so. The way forward is to recognise that trying to stay within the framework of individual agency rules out the idea that the action, in its totality, is understandable only on the collective level. However, it need not rule out the idea that we can consider the *activity as joint*, while keeping the *intentionality summative*.

To understand what it might mean to make activity joint while keeping intentionality summative, we must recall that action is more than just activity. Exactly what more is a matter of some very fine-grained argument. Nevertheless, as I noted previously, we can broadly say that it involves intentionality. To act is not just to *do something*, rather, it is *to do something with a purpose*. This is the difference between the acts, such as that of walking into a room, and something like the accidental activity of tripping over a rug.⁷⁹ As Frankfurt says, there is a "... contrast between what an agent does and what merely happens to him, or between the bodily movements that he makes, and those that occur without him making them."⁸⁰ Of course, from the perspective of the activity that takes place, an action can be identical to a mere unintentional happening. For example, the activity performed by a health and safety rep' who, as an illustration of the dangers of rugs, purposefully makes herself trip over a rug, can be identical to that of a clumsy man who does so by accident.

Just as we can separate the *doing* from the *intending* in individual action, we can do so in collective action. This allows us the option of admitting that one element of collective action must be considered as irreducibly collective, whilst keeping within the spirit of the conceptual framework of individual agency. We allow that the activity is collective – or as Bratman calls it 'joint' – but still hold on to the idea that all the intentionality in play must be in the heads of the individual agents. So, while the helpful-soloists were all cooperating towards performing individual instances of the same act type, i.e. each personally trying to get to the top of the mountain, what they were not cooperating towards was the performance of a single joint activity. By pairing up the idea that collective action must be cooperative, with the idea that this must be cooperation aimed at achieving a single joint activity, we arrive at the *cooperative joint activity account* of collective action. Strictly speaking, this account is not a summative one. Rather, it is mixed summativist and non-summativist; it says that there is an element of the action (activity) that must be understood by seeing the set of individuals as a singular unit. However, this is a very tame and unproblematic form of non-summativism. Activity *in-itself* is not something that we

79 See Davidson (1971, p.1) for discussion of this example.

80 Frankfurt, 1997.

normally reserve only for agents; we allow that rivers can flow and grass can grow without thinking that this somehow makes water or grass into an agent.⁸¹ Thus in allowing *activity* to exist at the collective level we can still hold that the only *agents* are individuals. This is a break from the framework of individual agency strictly conceived because, in allowing that activity is collective, we allow that collective action cannot be wholly understood as decomposable to separate individual actions. However, it *does* preserve what we might think is the core aspect of the FIA, that is, it preserves the notion that intentionality exists only in the heads of individuals and thus that only individuals can make something into an action. If action is activity plus intentionality then we can say, without violating the relaxed version of FIA, that collective action is irreducibly collective at the level of activity, but not at the level of intentionality.⁸²

The cooperative joint activity account allows us to identify the kind of cooperation that might be thought to give rise to the collectivity that characterises collective action and distinguishes it from mere mutually beneficial interaction. The distinctive feature of this account is that the intentional attitudes of the individuals involved are aimed at the achievement of a *joint activity*. It might be thought that the argument above does not require the existence of cooperation and only motivates us to make do with the idea of shared activity. However, this would not do, as without the inclusion of cooperation we could not rule out examples of multiple actions that fail to be collective by failing to be shared. This can be seen in situations where there are common intentions – which point towards the same activity – but that do not count as situations of collective action. The sum of secret intentions for the collective to act in such a way, does not seem to be the right kind of thing for us to say that the collective acted. For example, if it turned out that all the

81 Hence Gilbert, for reasons along these lines, refers to accounts of this sort as summative accounts. In the interests of clarity, we might re-describe summative accounts to fit this, defining them as: any account that makes *at least* the intentional aspect of collective action nothing more than the mere sum of the individual intentional attitudes of the participating members.

82 Such considerations fit with Michael Keely's suggestion that we must distinguish the notion of aims *of* a collective from that of aims *for* a collective (1981). He accepts that "... organisations [which I think we can fairly read as collectives in the sense used in this thesis], as systems of human interaction, produce events or consequences that are attributable to the organisation" and that these "... are 'more than' the aggregate effects of individual behaviour". However, he claims that "... from this fact that an organisation can so act in the sense of producing an effect, it is a large leap to the claim that it can act in the sense of intending an effect." (Keely, 1981, p.152).

climbers on the mountain each secretly intended for it to be the case that together they counted as swarming, this would still not be collective action. Gilbert makes a similar point with an example of two people on a train who secretly fancy each other. The fact that, each is sitting where they are sitting (e.g. next to each other), because each *wants* it to be the case that they are both travelling in close proximity, does not *make* it the case that, they can yet say that they are travelling together (in a strong sense).⁸³

What appears to be the case is that, the participants in a collective action need to *share* their intentions. That is, they need to *make them public*. We might think that this is implicit in the idea of cooperation that I introduced above, as it would be an odd concept of cooperation that had it that it was possible for us to secretly cooperate with each other. We can secretly act in ways that we know will assist another person, and it can even be the case that unbeknownst to us they are doing likewise for us; however, this would not be cooperation in the fullest sense. Indeed, Bratman introduces such a condition, claiming that we can understand the idea of “being in the public” in terms of the account of common knowledge given by philosophers such as David Lewis.⁸⁴ Roughly put, the notion of common knowledge is that it is transparent to each participant that the knowledge in question is known to all *and* known to be known by all.⁸⁵ Bratman does not claim to have an analysis of exactly what common knowledge amounts to. Rather, he claims to be using it as “an unanalysed idea”⁸⁶ and in this we can follow him, i.e. we can be neutral between

83 Gilbert, 1992, p.159.

84 Lewis, 1969.

85 More precisely, the idea of common knowledge is that it is transparent to each participant that the knowledge in question is openly known to all. So if *p* is common knowledge to agents *X* and *Y* then:

X will know *p*

Y will know *p*

X will know that (*Y* knows that *p*)

Y will know that (*X* knows that *p*)

X will know that (*Y* knows that (*X* knows that *p*))

Y will know that (*X* knows that (*Y* knows that *p*))

X will know that (*Y* will know that (*X* knows that (*Y* knows that *p*)))

Y will know that (*X* will know that (*Y* knows that (*X* knows that *p*)))

X will know that (*Y* will know that (*X* will know that (*Y* knows that (*X* knows that *p*))))

Y will know that (*X* will know that (*Y* will know that (*X* knows that (*Y* knows that *p*))))

X will know that (*Y* will know that (*X* will know that (*Y* will know that (*X* knows that (*Y* knows that *p*))))))

Y will know that (*X* will know that (*Y* will know that (*X* will know that (*Y* knows that (*X* knows that *p*)))))) ...

etc. Ad infinitum.

86 Bratman, 1999, p.102.

accounts of the exact mechanics, because it seems that it is possible to grasp the general idea of something being *out in the open*.

2.3.a – The problem of control

Collective action thus may be seen as being any cooperative joint activity. The notion of cooperation embedded in this account is meant to imply that the intentions of each individual (for it to be the case that they perform said activity together) are out in the open. This account takes us even further away from the simple summative account and it is easy to see why such cooperative activity would feel to involve being bound together (i.e. unity) and having of a commonality of joint purpose (i.e. collective intentionality). Though this account is reductive and in keeping with the FIA (at least with the spirit of the FIA), it is not claiming that collective action is an illusion in the same way that the *simplest* summative account claims that it is akin to the dancing mass of leaves. The key difference between this account and the simplest account is that it allows us to admit that there is a real and distinct social phenomenon that constitutes collective action and that this phenomenon can be distinguished from sets of individual actions that are not related. However, whilst not claiming that it is an illusion that there is a distinct phenomenon, this account is still claiming that it is an illusion that this phenomenon is one where there is *real* collectivity in the strong sense. That is, it is an illusion that there is a collective agent. There is then, even if not a fully collective sense, a real sense in which (on the cooperative joint activity account) intentionality is shared; it is directed towards same commonly known joint goal.

Given the above, such an account is a more credible contender for explaining collective action. However, this credibility comes at a cost: the way in which this intentionality is thought to be shared raises a problem I will call the *problem of control*. Roughly put, the problem of control is that it seems to be illegitimate for each member of a group to intend that they together perform an action *a*. This is because, in order to legitimately have such an intention, they must each think of themselves as able to settle the issue of the joint activity that will constitute *a*'s occurrence. But, it seems impossible that

they can *all* rightly think this to be the case. Intending for the joint activity to take place, appears to require that, the individual intender is in control of its occurrence. However, each individual's being in control appears to require that the other individuals are not in control. Reflecting on this problem will move us from the broad characterisation of cooperative activity accounts, given above, to Bratman's full Shared Cooperative Activity Account.

To explicate the problem of control more fully, I must again direct attention towards the nature of action in general. As I have said above, the performance of a collective action requires not just the existence of some joint activity; it requires that that joint activity is performed *with the intention of being the action in question*. Additionally, it does not seem enough that the agent in question has in their head the kinds of intentional attitudes that would give the activity purpose; rather, the activity has to take place *because* of these particular intentional attitudes. So much can be seen in the following example: let us return to the illustration, given above, of intentionally slipping on a rug. Suppose that I did have the intention to demonstrate safe rug slipping techniques but that I was not paying attention and slipped on the rug by accident. In this case, and others like it, it would seem that although the intentional attitudes are there, they do not make my action intentional. Something more than the mere existence of intention for that activity appears to be needed to turn activity into action – they must be connected in some way.

So what is the necessary link between the intentional attitudes and the activity? We might think that it must just be some causal link. However, this comes up against the problem that we can think of cases where an appropriate intentional attitude causes some activity, but that it fails to be an activity. Roderick Chisholm gives the example of someone who has the desire to kill their uncle and has the belief that they can do this by running him over. As this man is driving along about to commit his devilish deed, the seriousness of his unsavoury intention causes him to become nervous which, in turn, causes him to swerve the wheel of the car and accidentally run over his uncle.⁸⁷ Here, even though he intends to kill his uncle by running him over, *and* his having this intention causes it to be the case that

⁸⁷ Chisholm, 1964.

his uncle is killed in this way, it seems to be untrue to say that he performed the action of doing so. Something about the causal link between his intention and activity in question is deviant – and hence these cases are known as deviant causal chain cases.⁸⁸

What goes wrong in examples of deviant causal chains? Well, defining exact conditions that rule out such cases has proved problematic.⁸⁹ There may be different things wrong with the different cases. However, for our purposes here we need not be concerned with the exact necessary and sufficient conditions. Rather, we just need a general characterisation of what it is that those deviant cases lack. The obvious thing to say seems to be that the intentional attitudes must cause the activity in a way that makes it true to say that *the agent is in control of the activity*. When the agent slips on the rug, even though they held the relevant intentions and these caused the activity, they did not do so in a way that makes it true to say that they *controlled* the activity. There is no reason to think that this point should not hold for collective actions just as it does for individual actions. The positive claim, that it does so hold, is supported by imagining the following collective deviant causal chain example: imagine that together we have planned to descend the mountain on its south side, and knowing that this is our plan causes us to set out over the ridge. However, suppose that we miss the fact that a thick cornice of snow has built up and our trying to enact our plan causes us to walk onto this cornice, which collapses, and we plummet down the south side. This would not be a case of our performing the collective action of descending the south side, even though it would be a case of our collective intentions causing the intended activity. It follows from this that collective actions need more than the appropriate intentional attitudes to exist and more than for these to cause the appropriate activity; rather, these attitudes must further *control* this activity.

88 This is a case of what has been called *primary deviance* in that the causal chain goes astray between the intentional attitudes and the appropriate bodily movements. There are also cases that we can call *secondary deviance*. This is where the causal chain goes astray between bodily movements and results of those movements. We can see this in the following example given by Myles Brand (Brand, 1984 – as reported in Mele & Moser, 1997, p.7). A man may be trying to kill someone by shooting them, their aim is generally good but this time they miss. By chance the noise of the shoot causes a stampede of wild pigs which kill the intended victim. Here, even though the correct intentional attitude of the gunman causes the event and even though the event is describable in a way that fits these attitudes, it does not seem that the death of the man is really his action.

89 See Mele & Moser, 1997.

Returning to the cooperative joint activity account, it thus seems that if the joint activity is to count as an action, the appropriate intentions must not just cause it, they must control it. This gives rise to the following question, “How can it be that the intentions of each person all *separately control* the behaviour?” Take the example of walking to the top of the hill. If one of us is in charge (perhaps they are the elected leader of our hill walking team), then it seems legitimate to suppose that they can see their intention that we climb this hill as being in control of our activity. However, the cooperative joint activity account requires not just that there is one boss whose plan gives purpose to our activity, rather, it requires that each of our individual plans play this role, and thus that each of our individual plans are in control of our activity. If one person controls the situation it looks as if none of the others can. Of course, the thoughts of individuals can lead them to play their parts in the group activity. However, playing your part is not enough to settle the issue of whether the group activity takes place; being in control of the activity, that is ‘playing one’s part’, is not the same as being in control of the activity of the whole group. It looks like in order for each group member to intend that the group performs some collective action, they must be in a position to see themselves as making the decision about whether it does. The problem is, as Velleman puts it, that “... the logical space of decision making is open to those who are in a position to resolve the issue, and it admits only one resolution per issue.”⁹⁰ One way to solve this problem is just to say that there is just one controller – one agent that settles what happens – and that this agent is *everyone* taken together as a collective. However, to take this line would require us to abandon the framework of individual agency.⁹¹ Michel Bratman's alternative is to try to solve the problem of control within the framework of individual agency by using the cooperative joint activity account.

90 Velleman, 1997, p.35. Velleman has us think of an example of two people lifting a sofa and says of it, “Suppose that we jointly decide to lift a heavy sofa together. We thereby exercise a kind of joint discretion over the issue of whether the sofa will leave the ground. The interesting question is precisely how two people can jointly exercise discretion over a single issue. The answer cannot be that each of us exercises full discretion over the issue individually, as we would have to do if each of us were to intend that “we” will lift the sofa. Discretion cannot be shared by being multiplied in this way, since no issue can be settled by each of two people at once.” (1997, p.35)

91 Note that this is Margaret Gilbert's approach, and I shall be exploring it in the following chapter.

2.3.b – Bratman's solution to the problem of control

If we could make control collective then we could avoid the problem of control. There is no problem with saying that *we* are in control; rather it is just with saying that each of us has control exclusively. However, in parcelling out the appropriate intentions to each individual, the collective joint activity account appears to preclude any reference to *we* being in control that does not imply that each of us are individually in control. Bratman's solution to this is to cede something further to the non-summativist by introducing a collective element to the intentional stances of the individuals. But note that for Bratman this collective element need not comprise the intentions *themselves*, but rather can be merely an *enabling background* to them. Bratman believes that this collective enabling background is created by requiring that the intentional attitudes in question be interdependent. That is, the enabling background is created by requiring that the intentions in question are intentions for the collective activity to have a certain purpose *if and only if* all other members of the collective transparently also have such intentions. This attempted solution to the problem of control remains within the spirit of the FIA, as the relevant intentional states remain individual mental states.

Bratman approaches the problem of control first by noting that it is not an issue about the impossibility of the necessary mediation of other agents in each agent controlling the collective action. It is clear that our individual actions need not be the immediate result of our individual bodily movements; rather, they can rely on facilitating mechanisms. Bratman illustrates this point by having us imagine an "... example of the person – call him Abe – who moves the pump handle, thereby pumping water into the house".⁹² Here, use of the pumping system does not make Abe's action problematic, as the predictable reliability of the pump means that there is no issue about Abe's intentions controlling the action. So it is clear that mediation by a facilitating mechanism – between the appropriate intentional attitudes and the resultant activity – is not generally a problem. Now let us suppose that instead of a purely mechanical system (which aids Abe's pumping of water into the house) that there is another person involved. Bratman names her Barbara and has us imagine that

92 An example he reports as originally given by Elizabeth Anscombe (1963).

her job is to keep pushing a certain pump handle so that the water pressure will stay sufficiently high in the system, thus enabling Abe to pump the water.⁹³ Given the background condition that Barbara has the intention to maintain the water pressure, her role does not undermine Abe's intentional stance towards his action itself, for, given her reliability, he still has control over the situation. So mediation between the appropriate intentional attitude and the resultant activity, by another agent, is not a problem.⁹⁴

We can say then that control can unproblematically be mediated by other agents. However, there are several important differences in the collective action situation and the case of one agent merely facilitating the individual actions of another. In the pumping case, Abe relies on Barbara's intentions in order to have his intention to pump the water, but the reverse is not true, Barbara does not rely on Abe's intentions in order to have her own intention to perform her facilitating action. In contrast, in a collective action situation, the intentions and actions of each individual, if they are to be conceived as intending the group to act and performing the group action, are all interdependent. While Barbara's action can be the background to Abe's action, it is hard to see how all the performers of a group action could have their intentions both as the enabling background to the other's intentions *and* as backgrounded by those other's intentions. This appears impossible, in the same way as supposing that one can support the base of a stone pillar on the head of the same stone pillar.

The problem, then, is how the individual intentions can exist in the dual state of both being background to the intentions of the other group members and being backgrounded by those other intentions. Bratman's solution is that this *is* possible if intentions of the individuals have a particular *interdependent* character. He accepts that the background conditions are not right for each individual to start off with a straightforward intention to act; so he supposes that each group member starts with a certain mental attitude that is not quite an intention for the group to act but becomes one when everyone

93 Bratman, 1999d, p.150.

94 Lest, we might not find this convincing, Bratman has us compare it with a situation in which a mechanical device monitors the system and adjusts the water pressure, a situation that does not seem essentially different to one where an agent plays the appropriate role (1999d, p.151).

else shares the same mental attitude.⁹⁵ It is an ‘I will if you will’ situation. In this way each member can come to have their intentions *at the same time*, relying on the others to have their intentions whilst at the same time be relied upon to have their own intentions.⁹⁶ Thus they can all legitimately see their intentions as being in control of the collective action via facilitation of the intentions of the others. In this way, they can all be in control of the collective activity. Imagine two people, let us call them Seren and Freda, standing on the edge of a swimming pool. Seren intends to jump only if Freda intends to jump. Freda intends to jump only if Seren intends to jump. These intentions are common knowledge; they are out in the open between them. Neither can rely on the other's intention as a settled background condition of their own intention, for each knows that the other's intention is conditional on their own. It seems that they are at an impasse but they can solve it. Freda and Seren can *at the same time* come to intend to jump into the pool together through knowing that their intentions mesh in such a way that they are both satisfied.⁹⁷

We can think of the problem of control as asking how, as individual participants in a collective action, each person can frame their intentional attitudes such that the collective action is controlled by each one. Bratman's answer is: “... first, that I can 'frame' the intention that we *J* in part on the assumption that you, as a result, come also so to intend ... Second, even after I have formed the intention that we *J*, in part because I predict that you will concur, I can recognise that you still need to concur: It is just that I am confident that you will. Third, and finally, once we arrive at a structure of intentions that satisfies [the conditions of shared cooperative activity] we can see the matter as partly up to each of us.”⁹⁸

95 Bratman, 1999d, p.154.

96 To clarify, Bratman is not claiming that each individual intends conditionally (i.e. intends for the group to do *a* if everybody wants it to) rather it is the very intention that is conditional (i.e. if everybody conditionally ‘intends’ for the group to do *a* then I intend for the group to do *a*).

97 As Schmid notes, the requirement for there to be an interlocking mesh between individuals' intentions means that Bratman's conception of shared intentionality goes beyond individualism, inasmuch as it rejects the idea that a single brain-in-a-vat could be the bearer of a collective intention (2007, p.206). In this, Bratman's view of collective intentionality contrasts with that of John Searle who argues that the rejection of a group mind means that the collective intentionality must exist wholly in individual minds and thus should be possible for a brain-in-a-vat. (see Meijers, 2003 for a convincing refutation of Searle's Brain-in-a-vat condition and a discussion of the content externalism that is implied by this).

98 Bratman, 1999d, p.157.

Bratman's solution shows that the thoughts of different individuals can be coordinated such that each individual can be said to have personal discretion over the situation. However, the control issued by each is still itself discretionary; that is, it is at the control of each other. This means that the control that rests with each individual is only partial in that the real control lies in the *combination of all the individual thoughts considered as a whole*. His solution, then, introduces a further element of irreducible collectivity with its focus on what he calls an "interlocking web of intentions".⁹⁹ Importantly for staying within FIA, he does not see this web as itself constituting a new agent; rather, its role is as a background condition for the agency of each individual. This does, however, call for a further relaxing of our understanding of the FIA. We must say that while only individuals can have intentions, these intentions can be inter-reliant on the intentions of other agents in an important sense. This is somewhat more relaxed than the original statement of FIA and thus Elisabeth Pacherie is right to say that "... conceiving of shared intentions as an interlocking web of intentions of individuals, Bratman moves away from the classical reductive analyses of collective action".¹⁰⁰ Whilst it is true that Bratman has moved away from the simple reductive analysis, his account is still within the spirit of the FIA, as it respects the notion that intentionality really only exists in the heads of individuals.

2.3.c – Bratman's Shared Cooperative Activity Account

So Bratman believes that there is a real sense in which collective action involves shared intention. However, for Bratman, "... a shared intention is not an attitude in the mind of some superagent consisting literally of some fusion of your mind and mine."¹⁰¹ Rather, it is a sum of *interdependent* individual intentions. As I stated above, interdependence of intentions means that each intention is framed in terms of the existence and character of the intentions of others. This implies that the structure of these interdependent intentions must fit together. Bratman cashes this out in terms of the requirement for what he calls *messing subplans*. This is the requirement that our intentions

99 Bratman 1999d, p.143.

100 Pacherie, 2012, p.352.

101 Bratman, 1999c, p.111.

for the joint activity in question are co-realizable not only at surface level, but also at the level of the subsidiary plans. Bratman has us imagine that we intend to paint a house together, but that, both of us being stubbornly individualistic in our aesthetic tastes, I intend to paint it all entirely blue and you intend to paint it red all over.¹⁰² Suppose neither of us is willing to compromise and we go ahead and paint the house a resulting mess of red and blue. This action, Bratman claims, would not be a shared cooperative action and thus, in the terms I am using, it would not be a collective action.¹⁰³

Holding an intention necessarily requires the holding of sub-plans to achieve that intention. For example, if I intend to walk to the park then I am going to need to intend to put one foot in front of the other. Bratman calls this the principle of *Means-End Coherence* and renders it formally as, “[t]he following is always pro-tanto irrational: intending E while believing that a necessary means to E is M and that M requires that one now intend M, and yet not intending M”.¹⁰⁴ What holds for individual intentions also holds for collective intentions – intentionally *a*-ing implies the necessity of having sub-plans that will bring *a* about. If the different agents do not have meshing sub-plans then they cannot be thought to have coherent intentions about what they both will do. This does not imply that Bratman thinks that sub-plans must always be matching. He allows that, for example, one of our house painters might intend that we paint the house with inexpensive paint and the other that we buy the paint from Cambridge Hardware if “I don't care where we buy the paint and you don't care about the expense ... we could proceed to paint the house with inexpensive paint from Cambridge Hardware. Our activity could be cooperative despite differences in our sub-plans”.¹⁰⁵

Further, this interdependence must not be accidental. Rather, it must be intentional on the part of each agent; they must be committed to having their intentions mesh with the

102 Bratman, 1999b, p98.

103 It is not just that it would not be a collective act of painting the house a single colour, as obviously, it would fail to be this merely because it would fail to be the case that house was painted only one colour. It would, on Bratman's account, fail to be even a collective act of painting the house multiple colours – rather it would merely be an accidental joint activity of doing so.

104 Bratman, 1999f, p.413.

105 Bratman, 1999b, p.98.

intentions of the others with regards to the joint activity. This means that the intentions of each individual cannot be coerced. Bratman gives a case concerning the unacceptability of kidnap as a case of collective action and indeed one person being kidnapped by another does not ordinarily strike us as a collective action.¹⁰⁶ Bratman's ruling out of cohesion fits with the notion that cooperation is key to collective action. However, it is tricky to draw a line between allowable cohesion and non-allowable cohesion. Bratman must allow some level of persuasion and incentives otherwise his account would fail to apply to our real social interactions. In the case of Hillary and Tenzing climbing Everest, Hillary is climbing solely because it is what he wants to achieve. In contrast, part of the reason that Tenzing is climbing is that he must do so in order to get paid. We might imagine that it could have been the case that if he did not receive the money from this job then he would not have been able to buy food and he would starve (and that no other jobs were available) – is this different from the Mafia case, where the Mafioso tells the victim to come with him or he will shoot him?¹⁰⁷

So then, for Bratman, the essence of collective action is that it must involve joint activity, this joint activity must be intended by each member, each member's intention must be dependent upon and mesh with the intentions of each other member, each member's intention must be freely made and include the propensity to work with the others in achieving the joint activity. In contrast to the simple summative account, we have here a complex summative account which we could give a name that was some variant on the 'collective joint activity account' tag used above, such as the fitting but unwieldy '*interdependent, but individually intended, joint activity account*', but for simplicity let us stick with Bratman's own phrase, and refer to it as the *Shared Cooperative Activity (SCA)* account.

Importantly Bratman thinks that his conception of collective action answers the question of what role such action has in our lives. He believes that it helps coordinate

¹⁰⁶ Bratman, 1999b, p.100, also discussed in Bratman, 1999c, p.118.

¹⁰⁷ The actual situation was different from this in an important way in that Tenzing *did* independently want to climb Everest and saw himself as having a dual identity between Sherpa and Mountaineer. I will return to this issue in a further chapter.

activities, helps coordinate planning, and structures relevant social bargaining.¹⁰⁸ It does this because by its very nature it provides a background framework of baseline commitments that agents can take as a given. The fact that the agents are taking part in action that has the structure outlined above means that they must all commit to the same joint activity. Further, this commitment implies that they must also be committed to mutual support¹⁰⁹ as they are not merely committed to the joint activity, but committed to its being achieved through the meshing sub-plans of all of the agents. Thus, if, for example, we think of Tenzing and Hillary climbing Everest together, then on this account, because their seeing their action as joint requires that they intend to achieve it via the intentions of the other; they cannot see the other as a mere non-agentive tool. Rather, they must be flexible and willing to work together to both consciously achieve their goal.

In modifying the simple summative account we now have an account that is, we might say, *robustly social* – it seems to capture unity and something of the joint-ness of collective intentionality. However, one worry might be whether collective actions necessarily must involve such high degrees of cooperation. Further, we might question whether the complex openness requirement is needed. Both these points can be seen most clearly when we examine the phenomenon of detachment, as I shall explore below.

2.4 – Mere Sums -vs- detachment

Now we come to a tricky part of our experience of collective action for a theorist who is trying to stay true to the framework of individual action: detachment. To understand the challenge of explaining this phenomenon, let us return to the examination of the example with which I first set it out: the Greenbank residents' association agreeing to fight the development of their local park. In this example we can see the phenomenon of detachment manifest in the obvious feeling of separation between the individual's perspective (and its relation to the group act) and the collective perspective (and its relation to the group act). This separation generates the possibility of non-correspondence between

108 Bratman, 1999c, p.112.

109 Bratman, 1999b, p.94-95.

the individual perspective and the collective perspective. Because summative accounts, such as Bratman's, see the purpose of the collective act as given by the sum of the purposes the individuals have for it, this generates what I shall call the '*problem of non-correspondence*'.

The Greenbank residents' association, as we will recall, chose to collectively call for 'stopping the development' as the main goal of their group. Their decision to have this as their collective purpose is reached through compromise over various opinions, a rough consensus battered out through long discussion and then endorsed by majority vote as *the stance of GRA*. While this stance is generated by the interaction between the individuals – which may well involve expressions of their individual attitudes – it presents itself to those individuals as something separate. Imagine then that the Greenbank residents' association calls a meeting and twenty of its sixty members turn up. The majority, say fifteen members, vote to write a letter to the leader of the council, they then draft such a letter and send it off. Now it seems perfectly reasonable to say the activity of posting the letter constitutes a collective action by the GRA, a collective action that realises their collective intention to write such a letter. However, it seems that though the institutional framework of the group allows that the group's intentions can be set by a majority of attending members at a properly convened meeting, most of the the members have not directly participated and that their intentional attitudes have not been directly counted. The majority of the members may in fact have forgotten all about the meeting, and thus hold *no* intentions that correspond to the collective intention to send the letter. Gilbert raises this problem. She calls it the possibility of "... compatibility with lack of corresponding intentions"¹¹⁰ or the 'disjunction criterion'.¹¹¹

At first it seems that Bratman's SCA cannot deal with such cases. Indeed, the focus that Bratman puts on small-scale straightforward cooperative acts might lead us to believe that he does not see his theory as extending to them. However, he does admit the possibility of "...complex institutional frameworks"¹¹² and so it seems he must see a way

110 Gilbert, 2000b.

111 Gilbert, 2009, p.493 See also Gilbert's example of how a poetry group might collectively hold a position that does not correspond to the individual position of any of its members (Gilbert, p.190, 1970).

112 Bratman, 1999b, p.94.

that his account can fit with non-correspondence, but it is still unclear what this will be and unclear that it is possible while still keeping in the spirit of the framework of individual agency.

Perhaps we can solve this by using something like David Copp's idea of *operant members*. Copp believes that this phenomena is similar to that which takes place when one individual acts on behalf of another, so just as "... if Jones gives someone power of attorney and orders him to purchase a certain building for him then, if the attorney buys it for Jones, Jones has bought it."¹¹³ He also points out that not just anyone can perform actions on another's behalf (whether for another individual or a group) rather to be able to do so will depend on certain facts.¹¹⁴ These, he says, will include, "... all facts about the constitutional rules or laws, laws and by-laws" for organised collectives and "...facts about the composition of and dynamics of, or patterns of interpersonal relations within un-organised collectives."¹¹⁵ What these facts include in any given social group is perhaps a matter for empirical study, but it is fairly easy to see how the rules of Green Bank residents' association allow the chair of the group to set the collective perspective, but do not allow a stranger from another area to set the collective perspective. Things may be more fluid in non-hierarchical groups. The key seems to be that the performer of the action was in some way authorised by the group to perform an action on its behalf. As Tuomela puts it, the operant members of the group can only be said to act for the group if the non-operant members in some sense passively participate "... in virtue of having some relevant awareness of what's going on in the collective."¹¹⁶

113 Copp, 1979, p.177.

114 He refers to these as "constitution relation" or C-R facts (Copp, 1979, p.179). Also see Tuomela, 1989, p.482 for a similar discussion.

115 Copp, 1979, p.180.

116 As Gilbert says about the invasion of Czechoslovakia by Russia, "... In order for us to feel comfortable with the idea that a certain group [i.e. Russia] invaded Czechoslovakia there must surely be a sense in which whoever organised the invasion and whoever took part in it, was the authorised representative of the group as a whole" (Gilbert, 1992, p. 206) She also makes the interesting suggestion that Plato was getting at much the same sort of idea when he claimed that a city can be wise in virtue of its leaders being wise but only if those leaders are accepted by the citizens as being representatives of the city (Gilbert, 1992, p.470, footnote: 45).

Integrating the notion of operant members of a social group into the mixed account, however, comes at a price. The driving force of our acceptance of Bratman's SCA is the fact that it seems to stay within the spirit of the FIA – it claims that the relevant collective intentionality is spread throughout the heads of all the individual agents. If we allow that an operant member can have intentional attitudes *for* the other individuals, then are we not going against this? Not necessarily. The fact that Jones' attorney can buy a house on his behalf does not take away from Jones' agency. This is because we might think that it is dependent on Jones' acquiescing to having his own intentional states *set by others*. He has agreed that his stance towards the buying of the house shall be whatever stance his attorney takes for him. This would make explanation of the Greenbank Residents' Association case that involves the residents who are absent from the meeting, would be that via, say, their acceptance of the institutional structures that govern their group, each is agreeing to have their individual intentional attitudes set by the operant members of the group – who are constituted by the majority faction of those in attendance.

On a summative account that allows operant members, the important thing is still supposed to be the sum of the stances of each of the individuals; it is just that each individual has their personal stance fixed by means of the operant member's personal stance. It is, in Kirk Ludwig's terms, a situation of proxy agency.¹¹⁷ The problem is that the members of Greenbank residents' association in this second example, *do not* seem to be authorising some individual to set their individual intentions. If one personally thought that the main goal should be to improve the drainage, the result of the vote would not have meant that one's personal view has thus automatically changed. It seems perfectly reasonable for them to say: “*We think* that the main goal should be to stop the development, but personally *I think* that it should be to improve the drainage.” The operant members cannot then be seen as proxy agents for *each other individual* rather they must be proxy agents for *the collective as a whole*.

Bratman's SCA account can be seen as being robustly social, as it locks the individual into the collective perspective in such a way that the possibility of detachment is

117 Ludwig, 2013.

banished. However, detachment is a strong feature of our experience of collective action. As Christopher Kutz points out, even in situations where an individual member's stance does not conflict with the collective stance, it seems fine for them to simply lack the relevant intentions for the collective activity. He gives the example of a cellist in an orchestra. “[T]he cellist need not, and likely will not, engage in planning directed at ensuring that others will accomplish their parts, or worry about whether the bassoonists have properly realized the mood of the conductor's interpretation”.¹¹⁸ The feeling of detachment is part of the feeling of collective intentionality, inasmuch as we feel detached from the purpose of the group as a whole. We feel ourselves not to need to personally intend for the group. Rather, we only feel the need to participate in the group in such a way that there is a collective joint intention.¹¹⁹

Above, I looked at how collective intentionality must involve the correct kind of control over the collective action. Here we can say that the feeling of detachment from collective intentionality can also be seen in how the individual feels towards this control. To say that we feel detached from the intentional control of collective action is to say that it feels to be in the hands *not of any individual* to decide the collective action but rather in the hands of the collective. In Kutz's orchestra case the cellist will feel that it is their collective purpose as an orchestra that controls their activity, and thus makes it into their collective act of so playing. However, she need not feel that her personal intentions are in control of the orchestra's playing. In fact, she appears to lack the appropriate intentions that could be in such control. The control exercised by the conductor is as a proxy for the group as a whole rather than a proxy for each individual. This conflicts with Bratman's account because, as he sees it, the control is in the hands of each individual; it is just that it is mediated by an interdependent web. Each relies on the intentions of others, but nonetheless each still should feel to be in control. Bratman's account, then, does not sit well with the fact that we feel ourselves to be detached from such control.

118 Kutz, 2000, p.23.

119 Kutz (2000) believes that this point is obscured by Bratman where he considers non-hierarchical cases, such as the two painters, because in these – as a matter of practicality – each individual must hold individual analogues of the collective intention in order to achieve coordination. In contrast, in examples such as the orchestra the coordination can be achieved *through* the conductor.

2.5 – Mere Sums -vs- constraint

Bratman's SCA account has been shown to fail to make sufficient sense of the experience of detachment. Similarly, I will now present the case that it struggles to make sense of the experience of constraint. The canonical example of constraint I set up in the last chapter was the obligation of a member of the Leicester Tigers to play his part in the collective act of pushing his team's broken-down bus up the hill. Similarly, Gilbert has us imagine two people going for a walk – and she says of them that each “... gains a special standing with respect to the actions of that other person.”¹²⁰ That is, in participating in a collective action, each participant gains an obligation to play their part and to expect the other to do so also. The existence of this obligation is, Gilbert believes, best seen if we imagine either party transgressing its demands. Imagine that at some stage in the walk, Jack starts to draw ahead creating an increasing distance between himself and Sue. At this point Gilbert says, “Sue might call out 'Jack!’” or she might “catch up with him and then say, somewhat critically, 'You are going to have to slow down! I can't keep up with you!’”¹²¹ Even if Sue is timid and does not issue such rebukes, we can say, at least, that she would be entitled to.

Gilbert believes that the fact that such rebukes are possible, and appropriate, in such cases illustrates the fact that accounts of collective action should meet the following two criteria. Firstly, the *obligation criteria*: each participant has an obligation to promote fulfilment of the intentional goals of the collective action. Jack, by drawing too far away, is failing to promote the goal that they walk together. Secondly, the *permission criteria*: participants understand that they are not (ordinarily) in a position to unilaterally “by a simple change of mind” remove the constraints of the intentional goals of the collective action.¹²² Jack cannot remove his obligation towards them walking together by simply making a personal decision to turn back; he must get Sue's permission for the collective goal to be abandoned.

120 Gilbert, 2000a, p.7.

121 Gilbert, 1996b, p.180.

122 Gilbert, 2000b, p.17.

How can we explain these obligation and permission criteria within a summative framework? Our first attempt might be to suggest that, in forming a collective that is performing an action individuals come to have special moral obligations. We might then suggest that the wrong that Jack is doing to Sue when he walks too far away from her on their walk is a moral wrong. However, while the case of Jack's walking away from Sue might have some moral element, this need not be the case for all collective actions. For example, there seems to be some sense in which a gangster could be reproached by fellow team members for not playing their part. Most common cases need not be morally bad or good but are more likely to be morally neutral. I may have a moral obligation to keep walking with you. However, a similar rebuke seems possible if we had agreed to perform an immoral act, such as to kill someone together. Since we jointly decided to commit the murder, it seems that if we do not jointly decide to call it off then you are entitled to rebuke me for trying to pull out. Gilbert notes that it would seem odd to suppose that "... those who lack the concept of a moral duty altogether be incapable of going for a walk together"¹²³.

Perhaps the alternative is that we have an overarching intention to continue to hold on to our collective intentions. Roth calls this a bridging intention; he has us suppose that "... your intention has the requisite status and corresponding impact on my reasoning in virtue of a special meta-intention I form: the intention to coordinate my intentions with yours."¹²⁴ For Bratman, the existence of such a bridging intention follows from his requirement that the intentions of the individuals involved must be minimally cooperatively stable, that is it is baked-in-to the very nature of Shared Cooperative Activity. Thus he says that if I am a participant in a shared intention then "... the rational pressures characteristic of shared intention are built *right into* my own plans, given their special content and given demands of consistency and coherence for my own plans."¹²⁵

123 Gilbert, 1990, p.4.

124 Roth, 2003, p.67.

125 Bratman, 2009b, p.55 [*emphasis mine*]

However, the problem with the bridging intention proposal is that it seems that each participating member is free to merely rescind their bridging intention. Hence Roth says:

“Suppose I revise my bridge intention to coordinate with your intentions. This is presumably something I can do, given that the bridge intention is, by hypothesis, an individual intention. It is therefore mine to reconsider should circumstances arise that I judge to warrant reconsideration. If something does come up that gives me good reason to revise my bridge intention, I may do so – even though you don’t think it’s a reason to revise this intention. Once I revise the bridge intention, the intentions you have (supposing they are not geared to the revised bridge intention, assuming I even have one) will not have the status of rational constraint for me. I am free to ignore (circumvent or undermine) your intention that would have me *A*. I would be able to do some *B* instead. But this seems to give me a way of shielding myself from any sort of rational objection to my *B*-ing rather than *A*-ing. In revising my bridge intention, I sweep away any sort of authority you may have had as a party to the intersubjectivity between us.”¹²⁶

To the extent that he recognises it, Bratman tries to answer this problem in two ways. Firstly, he downplays the extent to which the obligation and permission criteria hold true. Secondly, he claims that when they do hold true this is just because of the normative necessity of a certain level of stability in individual intentions. This he calls the norm of *intention stability*. Bratman's concept of intentions as plans involves the claim that we must have reasonable stability in holding onto intentions.¹²⁷ Such a norm would account for the problem with Jack changing his mind and starting to walk away if he does so without giving due consideration to his former intention to play his part. In this vein, Bratman says that “... intentions are subject to a demand for stability ... the reconsideration of an intention already formed can itself have significant costs”.¹²⁸ This seems true. However, we may imagine that before he starts to wander off; Jack gives proper consideration to changing his mind. He thinks about how it fits with the rest of his plans and intentions and concludes that it is the best thing to do. In this case it does not seem that we could say that Jack was breaking the rule of reasonable stability in his intentions. Nevertheless, he would *still* be open to criticism from Sue – she would *still* be able to rebuke him. It is hard to work

126 Roth, 2003, p.80.

127 Bratman, 1999.

128 Bratman, 1999c, p.126.

out to what extent Bratman would accept the appropriates of rebuke in these circumstances, he does, however, introduce an element of the norm of intention stability that is specially in play in situations of collective action and might be thought to account for the inappropriateness of Jack's considered rescinding of his own intention. Bratman notes that "...an agent who too easily reconsiders her prior intentions will be a less reliable partner in social coordination." This means that there is a specifically social pressure towards stability. Bratman speaks of this as *cooperative stability* and he says of it that: "...an intention is minimally cooperatively stable if there are cooperatively relevant circumstances in which the agent would retain that intention."¹²⁹ However, while Bratman does give some room for the normative pressure towards cooperation in situations of collective intention, he insists that this is a question of levels of reasonableness rather than violation of obligation. So ultimately he says of the walking case that, "[w]hen I abandon my intention that we take a walk together I am, then, being unreasonable. But it does not follow that in abandoning my intention I am violating a nonconditional obligation grounded in our shared intention."¹³⁰

Given the phenomena picture that has been painted of collective action in this thesis so far, it seems fair to complain that Bratman underplays both the felt normative force of the collective will and the importance of the fact that it feels as if it is that very will (rather than any bridging intention) that is the direct source of the obligations. Given these problems, any attraction of Bratman's view must thus come back then to the attractiveness of the framework of individual agency. However, as we shall see in the next chapter, we may well be able to give a richer understanding of the feeling of constraint by abandoning it. As Gilbert says, "If a shared intention is such that one cannot unilaterally release oneself from participation in it by simple change of mind then there must be something other than a structure of personal intentions at the core of any shared intention."¹³¹

129 Bratman, 1999b, p.105.

130 Bratman, 1999c, p.126.

131 Gilbert, 2009, p.495.

2.6 – Chapter summary

In this chapter, I have set out the set out what might motivate us to accept the framework of individual agency and the challenges to such an account that lead to Bratman's robustly social Shared Cooperative Activity Account. However, while Bratman's account is robustly social, it appears that this might not be equivalent to its being *robustly collective*. The SCA account does seem able to explain the experience of unity, but makes it stronger and more demanding than we might expect. It makes some sense of the feeling of collective intentionality, but in doing so has to make use of complex structures of intentional inter-reliance that may be too demanding for real examples of collective action. Further, the structures of personal intentions as Bratman sets them up are not detached from the individual's own sense of personal agency, and thus he cannot make sense of the feeling of detachment. In particular, he must deny the existence of the feeling of detachment from the sense of control of the collective act. The biggest struggle is its failure to explain the direct nature of the feeling of constraint. One option, in Bratman's defence, is to claim that robust sociality is just the best we can get and that robust collectivity is just an illusion. However, such a counsel of despair only seems credible if all other options have been ruled out, i.e. if there are no alternatives. In the next chapter I will explore Gilbert's Plural Subject Theory as a relevant alternative.

Chapter 3 – Plural Subjects

Let us return to our two hill walkers and their collective act of climbing to the top of Scafell Pike. Recall that the departure point of the last chapter was that when we describe this act, we must see it as the sum of the actions of the two walkers considered *as individuals*. This departure point is the *framework of individual agency* (FIA). It is motivated by the idea that collective action must either be the act of a mysterious collective organism/spirit, or it must be no more than the sum of individual actions. Its most sophisticated realisation was given as Bratman's Shared Cooperative Activity (SCA) account. The SCA account allows that activity is characterised at the level of the collective (i.e. non-summatively), but remains true to the spirit of FIA by claiming that we can understand the relevant intentional attitudes as characterised at the level of separate individuals (i.e. summatively), by making them interdependent in complex ways. As we saw, this theory goes some way towards making sense of our phenomenology. However, it still failed to *fully* make sense of our experiences of constraint and detachment.

The departure point for this chapter will be to question the founding assumption of the last. Rather than accept that rejecting the organic thesis leaves us with only the sums of individual attributes to play with, here I posit the possibility of a different kind of collective agent. This collective agent is not a separate entity, but rather it is comprised of *a plurality of socially united people*. It will be referred to as a *plural subject*. In the first section of this chapter, I will argue that the existence of such pluralities is part of our everyday understanding of social life and that they are understood as the intentional products of instances of uniting together. Having shown such a notion of intentional social unity to be commonplace, the task will be to explicate it, and I will put forward the case that Margaret Gilbert's *Plural Subject Theory* (PST)¹³² fills this role. The chapter's structure will parallel the dialectical approach of the last chapter – i.e. starting out with some theoretical considerations (in this case, considerations about what it is that distinguishes the social from the individual) and then, in the following sections, seek to refine and clarify our

¹³² This account is developed throughout Gilbert's work but in particular see 1992, 1996b, 2000a and 2006.

understanding by considering the four elements of our experience of collective action in turn.

Gilbert's PST account shares the organic thesis's rejection of the FIA. However, in contrast to the organic thesis, Gilbert's non-individual agent (the plural subject) is not some entity that exists separately from the relevant individuals; rather, it is something that they together comprise through becoming socially united. For Gilbert, this social unity requires the sharing, or as we might say the pooling, of the individuals' agency. For this pooling of agency to occur, Gilbert believes that the relevant individuals must together, in conditions of common knowledge, make joint commitments to hold certain attitudes or do certain things as a *we*. I will unpack these notions in detail as the chapter progresses, but for now let us just consider our two hill walkers. On Gilbert's account, for us to jointly walk to the top of the hill: we must be engaging in the activity of walking to the top, and we, each in conditions of common knowledge, must have *jointly committed* to be a *plural subject* that intends to walk thus. For Gilbert then, the hill walkers are not merely cooperating in having mutually dependent other-directed intentions (as they are for Bratman); rather, they are cooperating in sharing one single *collective intention*. The notions of joint intention, plural subjects and collective intention will be explained as the chapter progresses.

In the last chapter, I made the case that Bratman's summative theory fails to fully get to grips with the four aspects of the phenomenology of collective action. In this chapter, I argue that Gilbert's non-summative theory is able to give a better account of the general character of all four aspects. However, as we shall see, there is a theoretical cost to accepting Gilbert's theory over Bratman's; it involves her explanation of the power of the plural will over the individual who is part of it. Justifying this cost involves grappling with the puzzle of how to square our understanding of ourselves as free autonomous individuals with the notion of normative constraint and detachment. The existence of this puzzle for Gilbert's account might be thought to show that Bratman's theory (which does not face such a puzzle) is to be preferred – even with its apparent explanatory deficiencies. Following the establishment of the superiority of Plural Subject Theory in capturing our social experience in this current chapter, I will approach this puzzle by comparing our constraint by our own

wills and our constraint by the plural wills we are part of. In tackling this puzzle, as I shall in the chapters that follow, I will be able to give a fuller picture of exactly what Gilbert's Plural Subject is.

3.1 – The possibility of social unity

The goal of this thesis is to build a theory of collective action. As we saw in the last chapter, the framework of individual agency is motivated by the idea that if we reject the possibility of an emergent social entity (whether it be a mystical disembodied spirit or a gigantic biological organism) the *only* thing that we have left to do so with is individual agency. Here, I present a case for rejecting this premise is built around what I shall call the possibility of *social unity*. Social unity, as I shall understand it, is the intentional binding together of individuals. Unlike the organic thesis, it posits the existence of no separate new organism. However, it does claim that something distinct from the *mere sum* of individuals can be brought into being. This distinct thing is composed of those individuals united together as a social unit. The possibility of this social unity undermines the motivation for the FIA because the socially-united-plurality-of-individuals is a plausible alternative candidate for the role of collective agent. In this section I establish that the possibility of social unity is part of our everyday understanding of social life.¹³³ In the following sections, I explore Gilbert's theoretical understanding of this general idea of social unity, in terms of her notions of the *plural subject* and of *joint commitment*.

To get to grips with the notion of social unity we must first grasp the central place its possibility plays in our everyday understanding of social life. A fruitful way to approach this is to start with a situation in which the phenomenon is clearly absent: let us imagine a lonely Robinson Crusoe sitting on the beach. Perhaps he is marvelling at his own self-sufficiency as he cooks the wild goat he single-handedly hunted. Such a life is surely the paradigmatic non-social existence. The contrast between Crusoe's existence and the

¹³³ This dialectical method mirrors Gilbert's in her foundational work *On Social Facts* (1992), where she comes to the importance of collective action through a concern with what it is that defines the social. In particular in that work, she is interested in contrasting her view with Max Webber's sociological account of the social as a collection of any actions that take into account, and are thereby orientated in response to, the behaviour of others (Webber, 1907).

experiences of the individuals involved in my paradigm examples of collective action – the lovers sharing a cake, the mountaineers working as a team to reach the summit of Everest, the rugby players jointly pushing the bus and the local residents banding together to fight the development of their local park – is stark. If – as I shall here – we define the phrase 'social life' to mean lives that involve social union,¹³⁴ then, the lives of the agents in these examples are clearly social lives, whereas Crusoe's island life clearly is not. What we want to know is what is the relevant difference between the non-social life of Crusoe and the social activities of our collective actors? As Gilbert asks, "... why do we pick out some phenomena as 'social' and deny to others this description?"¹³⁵

The question might seem odd because its answer appears so obvious; isn't Crusoe's life non-social simply because it lacks other people? However this does not capture the whole story. Of course, it would be strange to deny that an isolated individual is living a non-social life.¹³⁶ This though is not to say that the converse – i.e. the proximity of other people – is *in and of itself* enough to imply the existence of social unity. If this were the case, then merely adding others to Crusoe's surroundings should make his life social, but it does not. So much can be seen by way of an element of the original Crusoe story that is often missed out in the established myth of the archetypal isolated man.¹³⁷ This oft overlooked element is that, our primary character is *not* truly alone; the island he is marooned on is also frequented by so called 'savages'.¹³⁸ The point is that despite the mere existence of a multitude of persons on the island, his life still looks to be a non-social one. Or, to add additional complexity, at least it does until later in the story when one of the

134 It might be argued that this use of the term 'social' is narrower than that employed in common usage.

This may be, but I follow Gilbert in thinking that it is *at least* a paradigmatic use of this term. Gilbert acknowledges that the term "... is sometimes used to express a less rich notion" (1998, p.94). However, even if we were to accept that a looser usage was common, it would not damage the forthcoming argument and we could merely stipulate the usage here to be a term of art.

135 Gilbert, 1996c, p.265.

136 Though consider the case of a club containing a single person who keeps a once popular club going, giving speeches, writing down the minutes etc. even though she is the only remaining member, in the hope that others will soon again join. This may be an example of a solitary social activity, but if so, it is a special case that seems parasitic on the social unity of the club when it does have members. If, say, all the other people in the world had died, and thus that there was no possibility of the club ever again having more than one member it would be odd to continue to regard such activity as social.

137 As noted in Defoe, 1994, the idea of Crusoe as the archetype of the isolated individual has become somewhat separated from the detail of the narrative in the original (*proto*-)novel.

138 Defoe, 1994.

indigenous population, given the name 'Friday', escaping from his own group (and their plan to eat him), comes to join Crusoe. Though this pair are far from becoming social equals, they do appear to form a social unit.¹³⁹ What, then, is the difference between the set composed of Crusoe and Friday, which *does* constitute a social group and the set composed of Crusoe and the other islanders, which *does not*?

Such considerations move us on to a second possible condition for social unity: interaction. Clearly, a Crusoe who interacts with no one is living a non-social life. However, again we must be careful in moving from affirming this fact to asserting its converse. It cannot *only* be the fact that Crusoe interacts with Friday that makes their lives social, for Crusoe may well be interacting with all of the 'savages' – though he fails to form a social group with them. For example, he may be changing direction when he sees them coming, trying to beat them to discoveries of food, wrestling with and killing them when he can, and so on.¹⁴⁰ Gilbert supports the claim that mere interaction is not enough for social life with an example of a population of misanthropic humans living scattered in a large forest. She invites us to imagine that when any individual catches sight of another individual, they attempt to beat and kill them. These vicious people can certainly be said to interact with each other, but they clearly do not live as a society.¹⁴¹

It is not by accident that Gilbert's imagined population of misanthropic humans sounds very much like the fantasy 'state of nature' that is invoked by social contract theories

139 Putting aside scepticism about whether Friday is treated as another rational agent rather than a mere savage animal-like other. (see footnote 150 below for more on this scepticism).

140 These considerations are akin to those that arose when I first tried to illuminate the phenomenology of collective action. In Chapter One I used the example of driving a car down a busy high-street to illuminate the insufficiency of interpersonal interaction to characterise collective action. Driving a car down a road involves interacting with other people on that road, but this does not mean that one is performing a collective action with them. Likewise, on its own, it would not be enough to mean that one formed a society with them. Of course, the difference between Crusoe's relationship with the other islanders and the car driver's relationship with the other drivers and pedestrians is that we are not likely to think that the car driver is a social isolate; she does her driving whilst living a social life with those other drivers and pedestrians. However, this is not necessarily the case – imagine that one has been dropped on an alien planet and is driving around – the fact that one is driving about, or at least this fact on its own, is not enough to make it the case that one can say that one is part of a society with them.

141 Gilbert, 1992, p.36. To save us from imagining that it is only the viciousness of the interaction that make such lives non-social, she also sets out a similar example involving shy mushroom pickers. (1992, P.36). I will return to this example in the concluding chapter of this thesis.

as the conceptually pure non-social life.¹⁴² Identifying the non-social is a key general move for theories that attempt to move us towards an understanding of the bond of society. These discussions often invoke the notion of common interest as key to understanding social life. Is common interest, then, that which defines social union? Not quite, for there must be something more than the objective sharing of interests in a distributive sense. This can be seen in the fact that it may be true of those in the state of nature that they *would* be better off being in a society (i.e. they share a common interest) but this does not mean that they are thus already in a society. Rather, if the notion of common interest plays a part in defining social life, it must be via some *actual* process of coming to agree to be bound by a common interest. This fact may be obscured by the way in which sociologists and political theorists sometimes group together certain sets of people and talk of them as united by their shared characteristics. For example, Marxist social theorists see the set of people with common economic interests, due to their need to work for a wage and their lack of ownership of the means of production, as being united as a social class (namely, the working class). Nevertheless, it is not clear that such theorists mean to suggest that the working class is socially united in the sense I am trying to capture. Indeed Karl Marx's repeated calls for the working class to unite appear to imply that merely sharing the characteristics that make them the working class, does not mean that they are already united.¹⁴³ All of which is not to claim that mere sets do not have sociologically important causal impacts. Take the case of the set made up of all women in China. It is smaller than the set made up of all the men in China, and according to some social scientists this fact has real consequences for the functioning of Chinese social life, increasing the risk of social unrest in the country.¹⁴⁴ However, the objective social importance of their shared characteristics *does not* in itself make them automatically into something that is socially

142 Recall Rousseau's talk (noted in Chapter One, Section 3) of such things as men becoming social by "...uniting their separate powers" such that they "... are directed by a single motive and act in concert" (Rousseau, 2004, p.14).

143 Further, Marx's talk of the working class going through stages seems ripe for interpretation as the process of change from being a mere class to become a social unit. This can be seen in the way he describes social progress. For example, he says that at an early historical stage "... the Labourers still form an *incoherent mass* scattered over the whole country, broken up by their mutual competition". As industry progresses, however, "... the workers begin to form *combinations* (Trades unions) against the bourgeois; they *club together* in order to keep up the rate of wages; they found *permanent associations* in order to make provision beforehand for these occasional revolts." [*emphasis mine*] (2000, p.252).

144 As those low-status young males, who are finding it very hard to find partners, have an incentive to revolt. (Hudson & Boer, 2005).

united. As Gilbert puts it, having “... a shared way of life and common interests ... alone cannot be said to make a population into a collectivity”.¹⁴⁵

So, social unity is not an objective given; rather, it is the result of some actual process of becoming bound together. A fruitful way that Gilbert expresses this point is by saying that participation in social life involves “... a type of alliance, a partnership of sorts.”¹⁴⁶ Exploring this way of expressing our social experience can give us further insight into our common understanding of social unity. What, then, are the allusions that are conjured by talk of forming an alliance? Alliances are generally thought to be founded on the perception of a connection based on kinship or common interest. However, this is never *all* there is to such unions, for it is implicit that the parties must not just *conceive* that they are kin or that they have common interests, rather they must *do something* to make their alliance concrete – i.e. to actually bind themselves together. They are not automatically united by whatever it is they share; rather, they must intentionally form what we might loosely call a partnership. This talk of the necessity of *binding*, of *becoming allies*, of *partnership* is part of the fabric of our understanding of what it is to live a social life and implicit in our understandings of these terms is the notion that such a life involves more than mere objectively shared characteristics. What is common between these terms is that they show that being socially united goes beyond merely being *objectively* groupable together, it is rather *subjectively* making ourselves into groups. I use the terms 'subjective' and 'objective' here not in the metaphysical sense of what is *really* true, but rather in a narrower ontological sense concerning what is it that makes these things true. That is, making ourselves into groups can be understood as a subjective process, insofar as it is a function of the choices of subjects, rather than a function of the brute descriptions of those

145 Gilbert, 1992, p.39. That said, there are times when social theorists talk of sets of people with shared characteristics in a way that implies that they are already united, but as Gilbert points out “... one may sometimes be tempted to call something, T, an X if it is likely to become an X or if you think it ought to be an X. However, if one realises that this is why one is saying that T is an X, one will realise at the same time that one does not believe that T, is, as it stands, literally an X.” (1992, p.228).

146 Gilbert, 1996c, p. 271. A fact that she sees as gaining “... some support from etymology, in that the Latin *socius* means 'ally.’” (p. 271).

subjects.¹⁴⁷ All this can be expressed in more simple terms by saying that *social life is what we make it.*¹⁴⁸

Returning to the example of the end of Robinson Crusoe's isolation, we might ask: what *actually* happens when Friday forms a social union with Crusoe? How do they intentionally make themselves into a social group? Charles Taylor describes the properly functioning construction of society as a process whereby individuals become "... men deliberating together upon what will be binding on all of them."¹⁴⁹ Given that Crusoe, as described in the novel, conceives of himself as elevated by his superior white race over the savage Friday, we might think that it is unrealistic to suppose that they debate together in any politically ideal sense. Nevertheless, if we are to suppose that their relationship is of a different kind to that between Crusoe and the goats he herds (i.e. that they really do form a society) then it seems that we must claim that they do something that is at least in the same vein as Taylor's joint deliberation.¹⁵⁰ One thing that we can say of the pair is that they come to see their fates as intertwined, and that they share in some common understandings and rituals which they see as *their* understandings and rituals (in contrast to the other islanders who do not share this commonality). Such a process can be seen in the plot of the original novel where, though Crusoe does teach Friday to call him 'Master',¹⁵¹ he also

147 I mean to invoke the same allusions as John Searle points towards with his distinction between *epistemic* and *ontological* objectivity, though I do not mean to commit to his exact theoretical position (see Searle, 1996 & 2010). I discuss Searle's use of these terms in Chapter One, Section 1.3.

148 It is because Gilbert's Plural Subject Theory accepts the view of social union that is subjective rather than objective (in the sense described above) that it is an intentionalist project. Thinking that social union involves the active intentional activity of individuals fits with Schmitt's criteria of being "...characterised in part in terms of intentional attitudes and their contents." (Schmitt, 2003, p.22).

149 Taylor, 1985a, p.208.

150 Scepticism about the likelihood that agents with such unequal power would, in the actual world, come to see their relationship as a true social union is nicely captured by the poignancy of Gilbert's observation of the Native Indian Tonto's reply (in the popular 1950's television series) to the Lone Ranger's inclusion of him within the scope of the use of the term we; "We, white man?" (Gilbert, 1996b, p.198).

151 Defoe, 1994, p.149.

teaches him to share his language,¹⁵² his religion and his ways of living¹⁵³ and there certainly seems to be some sense in which they come to see their fate together as a common fate. To put it more abstractly, we might say that they come to see themselves as a *we*. This is the point that Gilbert wants to press home: forming a social group is an intentional act that appears to result from seeing oneself together with others as constituting a *we*.

3.1.a – Gilbert's Plural Subject Theory

I began this chapter with a rejection of the idea that once we have exorcised the separate social entity as an explanation for collective action, we are left with only the properties of individuals to fulfil that role. The task, then, was to locate this alternative candidate. Gilbert finds her answer in the central sense in which we use the pronoun *we*, that is in the possibility of forming a social unit. She calls this social unit a plural agent.¹⁵⁴ To help overcome any fears that such a plural subject is just as mysterious as a separate social entity, I have located the notion of social unity as a commonplace part of our understanding of social life and teased out, from the way in which this unity is seen as an alliance, the fact that we see it as an intentional creation rather than a natural fact. This understanding of social unity allows us to propose an account of collective action such that the performer of a collective act is a plural subject formed through the intentional unity of individual agents.

On one level, this social unity account might be thought to be explanatorily worse than the simple summative account in terms of requiring the existence of an extra kind of agent – i.e. the plural agent. However, in another sense it is clearly a victor in the Occam's

152 Sharing a language is important here because it allows more than merely being able to coordinate their interactions. Rather, the fact that Friday becomes capable of communicating with Crusoe allows more than just an exchange of information about the external world. It allows for the pair to reach what we might call a *publicly open consensus* about their shared lives. By 'consensus', here, I do not mean that they must each individually share the same views (given the divergence in their characters, at least at first, we might think this impossible). Rather I mean to use the term 'consensus' in the same kind of way that Taylor employs it. That is, in his terms, the *sharing of inter-subjective meaning* (Taylor, 1985b, p.36). In this sense the pair create what Taylor calls the 'general will' (Taylor, 1985b), or – as we might alternatively put it – they come to have public purposes. I will return to the importance of things being out-in-the-open/common knowledge later in this chapter.

153 Defoe, 1994, p.156.

154 Gilbert, 1992, p.152.

razor test by virtue of being simpler – that is, it does not require any complicated reduction or interdependent individual intentions. According to this picture the *we* is not some mere unifying description of an interplay of *Is*, rather it is an emergent creation, by those *Is* of something new, of a plural subject. It is a collective achievement – a construction. Thus, Gilbert can be defended from the accusation of invoking the 'scary monster' of a free-floating social entity that was evoked to justify a summativist approach at the start of the last chapter. Rather than suggesting that the plural subject is something mysterious that descends from the heavens to reign over people, her model is that it is something that *the people themselves* create through intentionally uniting together as one unit. There is a united performer, the plural subject, that acts, and thus things can be taken to be as they appear to be.

The question that naturally arises is what makes the notion of a plural subject any less mysterious than that of a separate collective entity? On this question Gilbert gives the impression that she sometimes regrets her choice of the term 'plural subject'¹⁵⁵ – this is because some people “read too much” into it. By this I take her to mean that, they see it as invoking the very mysterious collective entity that we have dismissed.¹⁵⁶ As I have hopefully made clear, such a reading misinterprets Gilbert; she does not think that the plural subject is a separate entity from individuals, rather she takes a plural subject to be a subject that individuals combine together to form. As Hans Bernard Schmid puts it, Gilbert's plural subject is a “...softened and modernized version of the 'collective subject'”.¹⁵⁷ Rather than looking for the author of collective actions in a mysterious collective social realm beyond that in which individuals exist, she believes that the individuals combined *are* the very thing that is the plural agent. Collective action, then, is neither the mere sum of individual acts nor is it the act of a separate singular social entity; rather, it is the action of individuals combined together as a plural subject. That is not to say that the question of what this

155 See for example, in 'Two Approaches to Shared Intention: An Essay in the Philosophy of Social Phenomena' (2008) she stresses that she continues to use the term (after first introducing it in Gilbert, 1989) “... just for the sake of a label” (Gilbert 2008, p.502). I find it hard to understand Gilbert's remarks here, other than to suppose that she is being overly defensive about her theoretical set-up for fear of being mistakenly taken to be pushing a variant of the organic thesis.

156 Gilbert, 2008, p.505.

157 Schmid, 2007, p.204.

softened and modernised collective subject actually is has been resolved. In what follows, by considering the four base phenomenal characteristics of collective action, I will begin to put the meat on the bones of the social unity account, progressing towards Gilbert's Plural Subject Theory (PST). At the heart of PST is the notion of the plural subject, and I will set out how Gilbert believes we construct that subject. The question of exactly what kind of thing it is that we thus create will be returned to in the following chapters; any residual mystery concerning the notion of agents 'combining together' should thus be cumulatively reduced.

3.2 – The *we*-vs- the feeling of unity

To understand how Gilbert's account gets to grips with the phenomenology of unity, we must understand her conception of the formation of a plural subject. As I said above, Gilbert sees in the central way in which we use the English pronoun '*we*' an invocation of the *plural subject*. Take the example described in the former chapters of the complexities of sharing a cake. In the example, Celia balks at being asked “Shall *we* share a cake?” by Bernard. For Gilbert, Celia knows that, to perform what she understands as a collective act (cake sharing)¹⁵⁸ with Barnard, she would have to become a socially-united-plurality (i.e. a plural subject) with him. Hence, Celia's baulking at Bernard's question is to be understood as her resisting the invitation to become a plural subject with him. According to this model, she rejects this invitation because she does not want to share in the feeling of unity that intentionally socially uniting with another necessarily generates.

In a *very* broad sense, Gilbert's model can be understood as a contractualist model of collective action.¹⁵⁹ In political theory, contractualists hold that agreements bind us together. They give us obligations and duties. In this sense, we might think that agreements are the kind of things that are perfect for creating social units. This account thus would say

158 Note that it is acknowledged that there are forms of cake sharing that may not be collective acts; i.e. each individually agreeing to have half. But, this example is stipulated to be one where the act is seen as: eating the cake *together*.

159 A fact she acknowledges in Gilbert, 2006.

that what happens when Friday is lost to his native group and joins together with Crusoe, is that his former agreement with his old allies is broken and he and Crusoe make an agreement that they shall now live together and that they will do things as allies.¹⁶⁰ There is, however, a problem with this as a model of social groups, a problem highlighted by those who criticise contractualism as a model of political obligation. While it may sound nice in theory, if we look at actual societies, it does not seem to be the case that individuals have all engaged in an actual practice of making agreements.¹⁶¹ Even when we scale down from the lofty height of nation states and look at our example of a miniature society of Crusoe and Friday it does not appear that even they must have made an agreement, at least not in the full blown sense, to have the kind of relationship they have. Indeed, surely this must be the case when they first met, as at this stage they have no common language in which to set out a formal agreement. Given that Gilbert believes the question of what it is to be in a society is tied to the question of what makes us capable of being united together such that we can perform collective actions, it should come as no surprise that she rejects the idea that social unity must be generated by actual agreement. Gilbert gives an example of two colleagues who happen to meet up, at first completely by luck, outside of work every day. She has us imagine that as they share part of the journey home; they both walk together. Now, after they have been doing this for some time, they may come to see each other as jointly walking (in the sense I have set out in this thesis). This can happen gradually over time without it ever having to be the case that they explicitly agreed that they would walk as a pair. We can see that they are indeed collectively acting, even without *explicit* agreement to so act, by noting that they will experience the phenomenology that I set out in the former chapters. They feel a sense of unity inasmuch as they feel themselves to be walking together as a *we*. They feel their act to have collective intentionality inasmuch as they feel it to be something that they are doing together. They feel detachment and constraint in that they feel that if another fails to show up for the walk they have violated an obligation that belongs not to any one of them individually, but to both of them collectively.¹⁶²

160 Albeit allies with very unequal power relations.

161 In political theory it is known as the 'no actual contract problem' and it is discussed by Gilbert in her *A theory of Political Obligation* (2006, p.70 – 87).

162 One extreme example she gives is argument as a joint action, such as when a couple says "we argued all day." (2006, p. 117) It seems like entering into an argument is the antithesis of entering into an agreement, for not only are arguments acrimonious, they can also be spontaneous and unconsidered. She

So if there need not be an explicit agreement, then what do we need? It is tempting to talk about tacit agreement – but it is unclear what this amounts to – Gilbert concedes that it is not completely wrong-headed to suggest that in all situations where people are allied in a way that makes collective action possible, there is what she calls an 'agreement-like structure' in play. That is, this allying appears to share features with situations where there are explicit agreements. However, Gilbert claims that we need to recognise that having an “agreement-like structure” is not the same as having an actual agreement.¹⁶³ The semantic question of the scope of the term 'agreement' is not the key point here. If we are being linguistically liberal, we might say that Gilbert's model involves 'agreement' in a very broad sense; however, it is clear that this is not a strict sense implied by people explicitly declaring themselves to be bound to a certain course of action.

Rather than an explicit declaration to form an agreement, Gilbert believes that collective action requires merely the *commitment* of the parties involved and an awareness, by all parties, that such commitment exists. Commitment is something that exists in formal agreement. However, Gilbert believes that it does not require formal agreement. We can understand what Gilbert means by commitment by looking again at the Crusoe example: when Friday is rescued by Crusoe, though he cannot make a formal agreement with him – sharing no common language – he can make Crusoe aware that he wants to be allied with him (and this he does by lying prostrate on the ground)¹⁶⁴ and Crusoe can make him aware that he sees this readiness on Friday's part, and he too is ready to share in such an alliance. Gilbert glosses the notion of awareness that the commitment exists in terms of openness of expression of readiness/willingness to *commit*,¹⁶⁵ saying: “One can say at least that each

suggests, reasonably plausibly, that arguing can be seen as a collective action. It has the phenomenology that I described in the first chapter – when we argue we often feel that there is an 'us' that is arguing, so we can say that though we are at odds we feel ourselves to be united, we feel that we together have an aim to our activity, the aim of arguing. While I think that Gilbert is essentially right to see arguments as an example of the possibility of collective action without actual agreement, this is true only of certain types of argument. For example, two people arguing in the street need not see themselves as a *we*. This does not nullify Gilbert's point, however, which can be made as long as there is at least one kind of arguing that is a collective action.

163 Gilbert, 2006, p. 117.

164 Defoe, 1994.

165 Note that the notion of commitment here is neither moral nor necessarily interpersonal; it is rather what Gilbert refers to as a 'commitment of the will' (e.g. Gilbert, 2006, p.127). I explore the exact nature of

party to the joint action does something expressive of readiness to participate in that action. Further, each party makes this readiness manifest to the others.”¹⁶⁶ The notions of awareness, openness and making things publicly manifest recalls the discussion of common knowledge given by philosophers such as David Lewis¹⁶⁷ (discussed in the last chapter).¹⁶⁸ As I suggested there, though it is an interesting puzzle to work out exactly how we can make sense of such public awareness (especially given the apparent problem of a necessary infinite regress of knowledge), such a notion is understandable, and fixing the puzzle is peripheral to our current investigation.

The fact that the kind of commitment that occurs between Crusoe and Friday is necessarily public in its nature, is precisely the attribute which Gilbert believes means that it is special; it is what allows it to remove us from being mere social isolates. For Gilbert; “Once the concordant expressions [of personal readiness] have all occurred and are common knowledge between the parties then the joint commitment is in place.”¹⁶⁹ This can happen in ways that seem to have structures very much like making deliberated upon agreements – the residents' association voting after a long meeting – or it can be something we can, as Gilbert puts it, “*fall into*”, as in the example the two workers that meet every day after work, and even more so in the example she gives of the collective action of kissing, of which she says that there is rarely any pre-standing discussion, rather “... someone expresses readiness for a kiss, say, by beginning to play one's own part of that process, and the other does likewise.”¹⁷⁰

The picture we have of how collective action takes place now looks as follows: collective action is the action of a plural subject which is created by members of the collective openly expressing their readiness to commit to being united. Much has already been said about Gilbert's model of collective action being founded on individuals

this commitment further in the following chapter (Chapter Four).

166 Gilbert, 2006, p.120.

167 Lewis, 1969.

168 Chapter Two, Section three.

169 Gilbert, 2008, p.502.

170 Gilbert, 2006, p.120 Indeed one might think that pre-standing discussion tends to rather destroy the likelihood of genuine kissing.

intentionally uniting together. Thus, it may seem redundant to ask if her model meets the requirement of explaining the experience of unity. Simply put, if collective action comes about through individuals committing to perform a joint action, and this can be understood as the creation of a plural subject, then they will feel the need to be united because they will feel that they have united themselves.

Returning to Gilbert's example of two people sharing a cake. Previously, I set the example up as follows: imagine a post-conference dinner where most people have just met, but two, Tony and Celia, are engaged to be married. After the main course, Tony asks Celia "Shall we share a pastry?" Celia nods to show agreement. Then, one of the other men, Barnard, turns to Celia, who is sitting on his right, and whom he hardly knows, and asks 'Shall *we* share a pastry?'¹⁷¹ Celia feels put out by this question because it invites her to share something beyond the cake; it invites her to share in a kind of unity with Barnard. Before, I said that this example showed that a bond of *some kind* needs to exist for collective action to take place. Now that we have an outline of Gilbert's model, we can understand what this bond in question is. It is not some objective fact about them that bonds them together.¹⁷² Rather, it is that they have committed to *do* something together. In this sense they have committed to become a *we*. Or to use Gilbert's terminology, they have committed to consuming the cake as a plural subject. They are committed to together forming a plural subject with the joint goal of having the pastry together, and this is what it is to share something in a strong sense.¹⁷³

So, the feeling of unity is explained in Gilbert's Plural Subject Theory in virtue of the fact that collective action requires that individual agents see themselves as creating unity between each other; they must openly express their willingness to commit to forming a plural subject, and – in doing so – become so committed. In this sense, as Gilbert says, "... a joint commitment unifies people in a very real way."¹⁷⁴ For Gilbert, this means that in

171 Gilbert, 1989, p. 175.

172 Though note, as in the discussion of economic classes above, sharing a relevant objective feature may well give them reason to join together.

173 Or at least this is what it is to share in contexts such as this, we can also have plain distributive sharing where each just gets their part.

174 Gilbert, 2000b, p.31.

saying that our lives are social we are saying that we are “...beings both independent and interdependent, units that are unified into larger wholes.”¹⁷⁵

I claimed above that a plural subject is not a separate social entity, but rather is just the individual agents united together in some sense; I have now claimed that this togetherness is created by those agents committing to be united. However, this raises the following question: if they can not be committing to be a separate social entity, then are they just committing to be what they already are i.e. a set of individuals? The answer is that they are committing to become the agent of the acts that they intend to collectively perform. In this sense the thing that they are committing to be (distinct from the mere set they already are) is an agent.¹⁷⁶ A problem with this that the term 'agent' is broad and carries with it much baggage that we may not necessarily want (for example moral/worth etc.) To remove this worry, let us just stipulate that, whatever the wider uses of the term 'agent', it is used here in a weaker sense. In this weaker sense to be an agent is just to be something that acts upon the world. In this way, we can read plural agents as truly being agents. This point, rather, hangs on a question of the metaphysics of agency that has not been addressed as yet; however, in a following chapter (Chapter Five) I will address what it means to be an agent.

3.3 – The *we* -vs- the feeling of collective intentionality

Now let us turn to the experience of collective intentionality; in the first chapter I had us recall the endeavours of Edmund Hillary and Tenzing Norgay, the mountain climbers who together achieved the first ascent of Everest. Clearly, this activity had a purpose; getting to the top of the highest mountain in the world was no accidental activity. Indeed, an activity's being experienced as having purpose seems key to it being an action – a point that appears to hold for collective actions just as it does for individual actions.

¹⁷⁵ Gilbert, 2006a, p.1.

¹⁷⁶ At one point Gilbert says something slightly different, which is that “ ... [t]he parties are jointly committed as far as possible to *emulate*, by virtue of the actions of each, with respect to its intending, a single body that intends to do the thing in question.” (Gilbert, 2008, p.503, [*emphasis mine*]). It is hard, however, to work out what she might mean in this passage by the term 'emulate'. It appears to me that she is here just being overly cautious in emphasising that her account does not imply that the plural subject is an actual distinct organism.

Further, the fact that Tenzing described their purposeful activity by saying that “*We stepped up. We were there*”¹⁷⁷ points to the importance of the intentionality being directed not towards individual activity, but towards collective activity. This is the sense in which it must be *collective intentionality*. Further, recall that discussion of deviant causal chain cases¹⁷⁸ showed us that it is not enough that there is a collective intention; rather, this must be *the* thing that settles the occurrence of the event. In this sense, we can say that the occurrence of the collective action must be settled by the collective intentionality. This is what was concluded in the first chapter.

We also saw in the previous chapter that if we try to make sense of this in terms of the intentions of all the individual participants, we come up against Velleman's problem of control. This problem centres upon the apparent illegitimacy of any individual member of a collective intending for the collective as a whole. The issue is that it appears impossible that each could rightly think of themselves as settling the issue of the action's occurrence. The intentionality in question needs not just to cause the action, it needs to control it – i.e. settle the case about whether it comes about. Bratman's solution to this problem was to require that collective action realises a complex structure of interlocking personal intentions. Gilbert's is far easier to express and I will suggest that her solution to the problem of control is not merely more economic, moreover, it is better fitting with the phenomenology. Gilbert's solution to the problem of control is that the collective action need only be fixed by a singular intention: the intention of the plural subject. This fits with how we experience collective acts –it is not that we are all each in control individually of the collective act, rather *we* are in charge of it collectively. When *we* are painting the house, it is *us* that must get it done etc. It also makes sense in terms of possession of the action – collective acts belong to the collective, in a non-distributive sense they are *their* action. In comparison, it is unclear who the collective action belongs to on Bratman's model. Summative accounts seem to be open to the criticism that perhaps according to them actions seem to belong to everyone (over possession) and at the same time to no-one (under possession).

177 Ullman, 1956, p.265.

178 See Chapter Two, Section Three.

This might be a simpler solution to the problem of control – but the question then remains – what can it mean to say that there is a singular (rather than distributed) collective intention. As we saw in the last chapter, one of the things that motivates the framework of individual agency is the idea that, as Searle puts it, all intentionality takes place in individual brains.¹⁷⁹ But is this assumption really as intuitively clear as it first appears? Certainly, having rejected the notions of a literal world spirit and the notions of organic entity (as we have done) means that there can be no collective mind *as in* a collective brain. Gilbert accepts this; however, as we have seen, she thinks that having some gigantic brain is not the only way to generate collective intentions. As we saw in the discussion of social unity at the start of this chapter, rather than something separate from us as individuals, for Gilbert, the plural agent is something that we comprise together as individuals. Collective intentions thus need not exist in collective brains, but can be generated together by us as a united whole.

On Gilbert's Plural Subject Theory, the relevant individuals need not individually share the collective intention; rather they must only jointly commit to holding it as a body – and, as such, form a plural subject. This answer will of course, as it stands, be unsatisfactory to a card carrying believer in the framework of individual agency (i.e. to a summativist). What they will want to know is what does it *really* mean to jointly commit? Summativists can allow that individuals can commit to joint projects, in the sense outlined in Bratman's SCA account – that is, allow that individuals can each be committed to performing a joint activity on the basis of their commonly known and comparable individual intentions. However, for Gilbert “... joint commitment is not a concatenation of personal commitments.”¹⁸⁰ Rather, she sees it as something that is robustly collective, and thus something that – while not creating a physical collective brain – does create a genuinely collective intention. Gilbert's idea is that joint commitments are created by what she calls the '*pooling of wills*'. We can understand this as a process by which people allow their agency to become part of a larger plural agent. This process is like sharing in money

179 Searle, 1996, p.24.

180 Gilbert, 2008, p.502.

by putting it into a joint bank account. If Bob and Jim pool their money then this does not mean that half the money is Bob's and half is Jim's; rather, *all of it belongs to both of them jointly*. Likewise, Gilbert thinks that, with joint commitment, if Bob and Jim jointly commit to become a plural subject with the collective intention x , then it is not that half the commitment is Bob's and half Jim's; rather, the *whole commitment is theirs together*.

One problem we might envisage with Gilbert's plural subject account is that of getting started. If Bob cannot jointly commit without Jim's doing so, and Jim cannot do so without Bob's doing so, then how do they kick the whole thing off? Gilbert's solution is that each expresses conditional willingness to jointly do something and that these combine together to form joint commitment. The problem of getting started seems to be the same as Bratman's problem of each being an enabling background to the other. However, the outcome is collective rather than individual. When discussing Bratman's SCA account, I introduced the thought of two agents standing at the side of a pool, neither wanting to jump in alone. According Bratman's framework, we can imagine that they get around this by each conditionally holding personal intentions to jump which have the other's holding a meshing intention built into them. Each says, "I will if you will", and they both can then jump at the same time. In contrast, according Gilbert's framework, we can imagine that they chose to pool their wills together, by way of each making their willingness to do so public, such that they no longer see the choice over whether to jump as belonging to one or the other of them but rather they jointly commit to it. Each thus says "We will jump!"¹⁸¹ So, for Gilbert, the pooling of wills in this way requires that it is out in the open that each individual is willing for there to be such a pooling. In this way, Gilbert, like Bratman, believes in the importance of common knowledge to create the public nature of collective action. The difference, however, is in what it is that needs to be commonly known. For Bratman, it is the personal intentions of each individual that must be in the public sphere. For Gilbert, it is rather just the willingness to be jointly committed that must be in the public sphere. Looking at Bratman's account through Gilbert's eyes we can see why Bratman's SCA account provides sufficient, though not necessary, conditions for the

181 To fully understand the notion of joint commitment we must, of course, say more than this. I will return to examination of what it means to jointly commit – particularly, what it means in relation to one's individual autonomy – in the next chapter.

creation of a collective action. This is because any act that has the form that Bratman sets out is going to, by virtue of having this form, also end up satisfying Gilbert's conditions for joint commitment to the plural subject.¹⁸²

3.4 – The *we* -vs- the feeling of detachment

In the last chapter, I argued that it is difficult for summative accounts, such as Bratman's SCA account, to explain the feeling of detachment. This is, because they require that the purpose of collective action must be reflected in the sum of the individual views about the action. That is the intentionality of the collective act implies that each individual holds a corresponding individual intention that is directed towards the collective. Thus, there is, for them, no distance between the collective intentions and the individual intentions. A non-summative account, however, does not suffer this problem for it allows that there *is* distance between the collective agency and the individual agency.

Recall the example of the Greenbank residents' association coming to hold their collective view on the basis of a vote where not all the members are present. Each member does not feel the collective view to be a sum of their view together with the views of the other group members, but rather each feels it to be something separate from their own view. On Gilbert's account this possibility is easily explained. For, on it we can say that, the individuals involved in the group are free to jointly commit to becoming a plural subject with a certain view, without this thereby having to be the view that they individually hold. So, for example, the minority, who lose the vote over whether the GRA should write a letter say, can nonetheless accept that *together* they are committed to sending such a letter. Importantly (for explaining the phenomena at hand), they can do so *without* themselves changing their minds about the wisdom of doing so. The possibility of compatibility with lack of corresponding intentions is thus not a problem on the PST account. So, according to Gilbert, the single thing that constrains the practical reasoning of both parties is

182 There is some ambiguity about whether Bratman is trying to provide a fully sufficient explanation for collective action – while the examples that I have set out in the previous chapter appear to imply that he does, at certain points he appears to be more equivocal. Gilbert complains about this ambiguity (2008, p.489).

something over which neither party has solitary control. Thus, at the same time, each feels both to be part of the collective action, and also to be detached from it.¹⁸³

Given the above, Gilbert's account does not need to make use of operant members in the same way that a summative account does. However, this does not mean that such operant members cannot at times play a part in collective actions. Indeed, we might imagine that Gilbert's theory must be able to explain this in order to fully capture the example of the GRA voting example, for she must be able to explain how it is that the members who are not present at the meeting come to be jointly committed to writing the letter. The difference between the role played by operant members on Gilbert's account and that played by them on accounts like Bratman's is that while on summative accounts the operant members must be seen as setting the views *of each member* – on Gilbert's account the individuals can continue to hold their individual views, but merely commit to the operant members setting what they are committed to holding *as a collective*. This fits the phenomenology of detachment better, as we feel detached from the actions of operant members, that is, we feel their actions not to be directly for us *as individuals*, even though we do feel them to be acting for us *as a collective*. Further, this phenomenon of detachment can be experienced as complete ignorance of the collective intention of which one is part because one has ceded power to an operant other. Gilbert gives the example of a couple: “Asked about his vacation plans, George might turn, in ignorance, to his wife Rosa, and ask ‘What are our plans, love?’”¹⁸⁴

For large groups, because individuals will not necessarily be acquainted with each other, there is a need for a weaker version of the claim that it must be out in the open to the participants that each has made themselves part of the required commitment. The common knowledge condition must be tempered somewhat in large groups, for it does not seem to be the case that each agent must know all there is to know of the intentions of the other members of her group. Does this mean that we must abandon the requirement for common knowledge? Gilbert claims not; she claims that we merely need to understand that

183 Which is precisely what makes it such that Velleman's 'Problem of Control' (as discussed above) is not a problem for this account.

184 Gilbert, 2000b, p.23.

in large groups individuals have what she calls *population common knowledge*. As she says, in “ ... many populations, particularly large ones, the parties do not know one another *as particular individuals*. For instance, they may know that many people live around them on a particular island but they do not know of them as particular individuals ... [here] the parties of the commitment would understand themselves to be jointly committed insofar as they are *living on the island qua island dwellers*.”¹⁸⁵ That is, while a participant in a large collective action may not be able to point out every other member, they can, however, make reference to them by virtue of some distinguishing feature that makes them a member of the collective in question, and it is out in the open that agents who have this feature are part of the specified commitment.

3.5 – The *we*-vs- the feeling of constraint

Given that PST is able to found the experience of detachment on the real phenomenon of the collective view being something other than the sum of individual views, it is easy to see how this detached-other-view can thus be experienced as coming into conflict with the individual’s view. In the case of the GRA, the individuals who vote not to send the letter will clearly feel that the stance that they come to hold collectively – as a plural subject with their fellow members – is in conflict with the view that they hold as individuals. Recall, however, that the feeling of constraint does not just consist in the possibility of opposition between the individual view and the collective view; rather it also comprises of the fact that the individual feels some sense of obligation towards that collective view and towards playing their part in realising it. Hence, in the example of the Leicester Tigers engaging in the collective action of pushing their broken down coach up the hill, Bob, a lazy team member who personally does not care if the bus reaches the garage at the top, feels constrained to play his part in the pushing. If Bob manifests his personal preference and fails to push, he is open to rebuke, and this rebuke is founded on his allowing himself to be guided by his personal goals, even though they conflict with those of the collective.

185 Gilbert, 2001, p.52 [*emphasis mine*].

The feeling of constraint then appears to reflect a real obligation that is a direct result of the team membership; recall the helpful bystander and how she is not obliged to continue. Further, the rebuke that the other team members are justified in issuing involves them having the standing to issue such rebukes. They have this standing because they are all intimately connected through their pooled will to the joint commitment to push the bus as a team. For Gilbert, then, the constraint that issues from being part of a collective action is not, as it is on Bratman's account, an indirect result of morality, politeness or feelings of loyalty, rather “...one who calls another on his inappropriate action may well [rightly] justify his intervention *by reference simply to the shared intention*.”¹⁸⁶ and this is because “[t]ogether they constitute the creator of the commitment; the 'one' who imposed the relevant normative constraint on each of the parties”.¹⁸⁷ Of course, as Gilbert also notes, that your fellow team members have the standing to rebuke does not mean that they are always justified in rebuking.¹⁸⁸ In the rugby club example, though Bob is part of the plural subject that is pushing the bus, and his fellow players thus have the standing to rebuke him if he fails to do this, they may not be justified in doing so *all things considered* if he has, say, a bad back. This is because, it would be cruel to make Bob honour his commitment if doing so will cause him lots of pain and thus there may be a moral obligation to excuse him from his commitment.¹⁸⁹ So the difference between the way that Bratman's SCA attempts to explain the feeling of constraint and the way that Gilbert's PST attempts to explain it, is that: SCA allows for some obligations, but requires that these are just “downstream factors”,¹⁹⁰ (i.e. that they require additional factors to justify the nature of the shared intention itself); whereas, according to PST, the “... obligations of the parties are, one might say, purely *internal* to the commitment”¹⁹¹ (i.e. that they “.. inhere in the shared intention itself”).¹⁹² It

186 Gilbert, 2008, p.499 [*emphasis in original*].

187 Gilbert, 2008, p.504 We might worry that this implies that it is only together, i.e. as the *we* that they can rebuke. Gilbert's answer is that a “... given party is in a position to demand conformity or rebuke for non-conformity as co-owner of the action in question ... he might say, 'Give that to me that's mine – qua one of us!’”(Gilbert, 2008, p.507).

188 Gilbert, 2008, p.498.

189 A further complication is that, you might suppose that this 'if-injured–get out' clause may be implicitly written into the plural commitment to push the bus. I discuss this possibility in section 4 below.

190 Gilbert, 2008, p.500.

191 Gilbert, 2008, p.507 [*emphasis mine*].

192 Gilbert, 2008, p.500.

is this difference that means that, while Bratman must essentially explain away the feeling of constraint, in contrast, Gilbert can show how it is an essential feature of our experience of collective action.

3.5.a – The complication of hierarchy and standing conventions

So we have arrived at a picture of collective action that sees it as an essentially voluntarily entered-into social union. The complication of this picture is that collective actions do not always take place in the smooth ways that paradigm examples that I set up might suggest that they do. Thomas Smith makes the following point; “... each of us engages in such activity against a background of circumstances, including the actions and states of others, which we do not choose, and cannot control. For these reasons, one often freely and rationally partakes of intentional joint activity, of which one does not really approve, and which one does not intend.”¹⁹³ In the next chapter (Chapter Four) I am going to be arguing that Gilbert does not fully grasp the nettle of this point and that there are further issues that are brought up by the notion of constraint; chief among them is the relation to autonomy. More immediately I will return to this in the next chapter, where normative constraint will be discussed in more depth, and where I will suggest that it raises a challenge for Gilbert, the challenge of understanding how this constraint can be squared with our conception of individuals as necessarily being at least minimally autonomous, which ultimately requires moving past her conception of voluntary pooling of wills. Because of this, I will come ultimately to suggest that her Plural Subject Theory – pooling of the wills - is overly voluntary and must be replaced by a notion of '*entangling of the wills*'. However, that discussion must be delayed for us to finish our current task of gaining focus on the general shape of Gilbert's account. Here, let me set out how Gilbert does attempt to engage with the possibility of what she calls *hierarchy*.¹⁹⁴

Gilbert understands the possibility of hierarchy as the existence of 'followers and leaders' within collective actions. Whilst army generals or heads of states might most readily

193 Smith, 2011, p.230.

194 Gilbert, 2008, p.99.

spring to mind as embodying this phenomenon, it is clear that even the smallest groups can be hierarchical – as we can see in the unequal power relationships in the Crusoe example. The term 'hierarchical' has a rather oppressive ring to it, but examples of it need not be oppressive. For example, the voting and holding of executive positions in the Greenbank residents' association may be thoroughly democratic, nonetheless, it still involves hierarchy. These are for Gilbert, “[n]on-basic cases [that] involve authorities whose status derives from a basic joint commitment”.¹⁹⁵ They involve 'tacit' commitments and conventions (in the sense explored below) to explain how we can accept the idea that these make possible complex groups with multiple layers of hierarchy and power but still staying within Gilbert's conceptual framework.

Gilbert allows the idea that when we enter into joint commitments we may have special background understandings.¹⁹⁶ This might be explicitly stated, or they might just be implicit because they involve standing conventions. Standing conventions can be separated into two categories: *private conventions* – such as two walkers have generally acknowledged rules that either one can call off the walk and more widespread *societal conventions* – for example, that it is taken as a given that someone with a broken leg is excused from playing their part in pushing buses up hills.¹⁹⁷ Take the lovers from our example above; Tony and Celia form a plural subject, thus they will have a number of ongoing joint actions and goals and they will face rebuke if they break them. However, imagine that one partner is unfaithful; Gilbert says "...suppose that two people are living together as man and wife ... If one partner is discovered to have engaged in sexual activity with a third party, the offended partner may aver, "We're through!" and the other may not question that point ... Such language suggests the existence of an established condition of the kind in question [i.e. background condition]".¹⁹⁸

195 Gilbert, 2008, p.502, footnote: 48.

196 Gilbert, 2009a, 494.

197 Gilbert, 2006, p.110.

198 Gilbert, 2006, p. 112. Gilbert's description of the phenomenology seems right here, however, her reading of it seems to me rather odd. Is there such a convention? I'm not sure. Examples such as lovers tend to push peoples' buttons when talking about constraint. Andrea Westlund (2008), for example, takes it to be incompatible with the reciprocity that is at the heart of loving relationships that the collective will could constrain either partner (I will return to this particular question later in the thesis).

A major complication in accepting this possible role of standing conventions is that it complicates our experience of the phenomenology set out in the first chapter. Namely, it allows for situations in which our experience of constraint is mitigated. Gilbert discusses Bratman's example of duet singers as an example of this. In it, Bratman says that it is fine to think that there is a collective action going on, even if each party sees the other as free to pull out at any time. Gilbert, however, thinks that the "... possibility of such a side-agreement does not force us to reject the obligation criteria."¹⁹⁹ It is rather just that there is a pre-standing joint commitment that the collective position defaults to ending the joint action if any one member individually pulls out. Related, Gilbert also admits that there are circumstances where one had better take the reins.²⁰⁰ Imagine, for example, that after the endeavour has gone on for half an hour the team captain, Jim, calls a stop to the activity. "This is far too risky to our health", he might say, "We should stop!" The team manager might take a different view and a row might break out. In such a case, would we say that the players remained obliged to continue pushing the bus? If not then, does the team captain's ability to opt to stop to the collective activity, and thus to remove the collective obligations, undermine the permission criteria?

To see why not, we have to ask what makes Jim's calling a halt to the activity different from Bob's attempt to extricate himself from it. The key is the fact that Jim is not trying to unilaterally remove himself from the collective obligation, and is thus not challenging the notion that constraint is a universal part of our experience of collective action, rather he is trying to change the collective goal. Further, because we have given Jim a position of authority, as team captain, he seems to have the right to at least try to do this. So, while a team member might not *ordinarily* be able to unilaterally set the collective goal, things will be more complex when there are positions of social power involved. Whatever power Jim has to set the collective goal comes from the fact that he has a social position within the group; inasmuch as he can make decisions for the group, he can do this only because the group is structured in such a way as to give him authority over the others. Making room for such cases, Gilbert says that participants can "... authorise someone (or

199 Gilbert, 2008, p.499.

200 Gilbert, 2006, p.107.

somebody) to set up shared intentions for the people in question.”²⁰¹ This is the point, addressed above, about the role of operant members according to Plural Subject Theory. Note, though, that the example included the idea that the team manager might disagree with Jim. In this case, the structures of authority may be unclear. The club may have a written constitution that governs team matters; however, whilst this is likely to cover arguments over who gets the final say in picking the first team or new team colours – it is very unlikely to cover who has the final say when the team are trying to fix their bus. Whether the team members feel that the captain's calling an end to their collective action ends their obligation to push it will depend then on each player's conception of the structures of authority within the group.

3.6 – Chapter summary

In this chapter, I have argued that Gilbert's PST can successfully reject the FIA without committing the sin of metaphysical extravagance. It does this by using the notion of social union that is embedded within our general understanding of social life. Further, because PST allows that there is something robustly distinct from mere sums of individuals, it can do a better job of explaining the strong sense of collectivity of our experiences of unity, collective intentionality, detachment and constraint. Gilbert's account thus does a better job explaining our actual social lives than summative accounts such as Bratman's SCA account. Let us suppose that the critic concedes that the notion of social union is part of our general understanding of social life; nevertheless they may still insist that the argument – as I have presented it so far – still leaves the notion of joint commitment in a somewhat mysterious state. Bratman's reading of Gilbert seems to be that she commits just this error – i.e. that she makes joint commitment an un-analysable primitive. One move would be to accept this, but plead that we have to hit the ground of a primitive phenomena at some point. However it seems fair to think that we can do better than this, and in the next chapter I will suggest that there is much more Gilbert has to say regarding the question: “What is joint commitment?”

201 Gilbert, 2000b, p.23-24.

Chapter 4 – Harmony, Discord and Autonomy

Moving from Michael Bratman's Shared Cooperative Activity Account to Margaret Gilbert's Plural Subject Theory requires a shift in the way we understand our pair of walkers. We must no longer view them as merely two independent agents cooperating to reach the summit; instead we should see them as a pair whose agency has been pooled through a joint commitment to reach the summit *together*. In understanding things in this way, we do not conjure up any mysterious social spirit, nor do we posit an emergent organism with them both as parts. Rather, we see the two individuals as united by means of intentionally binding themselves together. As I have explored, the notion of forming such self-bound units is not some rarefied artifice of Gilbert's theorising, but rather it is a part of our general understanding of social life.

Plural subject theory trumps summative accounts in getting to grips with the essential *collectivity* of our experience of collective action. However, in doing so, it generates the following question: what does it mean to understand joint commitment as truly *joint*? Answering this question must start with grasping that Gilbert understands joint commitment as a type of *commitment of the will*. As we shall see, she distinguishes this type of commitment from moral or contractual commitment. It is to be understood as an expression *of the will*. With this centrality of the notion of the will in mind, my exposition of joint commitment will have as its foundation a consideration of how individual and plural wills relate in situations of collective action. I will set out this relationship as characterised by *harmony* between wills and *discord* between wills. These phenomena appear to be conflicting but, as I shall explain, they are in fact complementary. This is because our wills can be harmonised inasmuch as they together form a singularly directed collective will, but at the same time, they can conflict with this collective will and with each other. So much can be seen as our two walkers pass Mickledore on route to the summit of Sca Fell: they feel their wills to be harmonised, in that they both share in the collective goal of reaching the summit. However, there may be discord between their wills, in that the

individual goals of each (say to sit down and have a rest or to divert to an easier path), may be in conflict with those of the other and with their collective goal.²⁰²

I explore the exact relationship between harmony and discord in the first section below. By clarifying this relationship, I shall then, in the second section, be in a position to explain how it is that Gilbert thinks that the public sharing of mutual readiness to jointly commit can create a plural subject. Further, I will be able to express the exact sense in which Gilbert thinks that collective commitment is normative; which is that she thinks that it is a matter of *rational requirement*. In this light we can see that if, say, one of our walkers was to turn around and retreat (without seeking a change to the collective goal they share with their partner), the distinctive wrong they commit would be a violation of that which the collective will rationally requires.

While this additional clarity will help to further demystify the notion of a plural subject, it will also bring to the fore the fact that PST entails the possibility that one individual can be constrained by the rational requirements of a will that is not their own. I follow Abraham Roth in calling this phenomenon *practical intersubjectivity*.²⁰³ In the final section of this chapter, I discuss the fact that the possibility of practical intersubjectivity raises a puzzle. This puzzle centres around the relationship between the assumption that autonomy is a part of our conception of ourselves as agents, and the fact that we feel ourselves to be constrained in situations of discord. To illustrate again with our walkers: we might worry that, in allowing that their individual wills might be constrained by the collective will, we make them both into mere puppets rather than true agents. This is conceptually unappealing because it appears to involve denying that they have the kind of autonomy we generally take individuals to have. I will refer to this as *the puzzle of intersubjectivity*.

202 Gilbert similarly uses an example of two walkers to illustrate the same point (2008, p.491), though using different terminology.

203 Roth, 2003.

4.1 – Harmony and Discord of wills

In the last chapter I asked what it is that makes Robinson Crusoe's relationship with Friday different to that which he has with the natives he battles. The answer was found to be that they together form a social bond; that is, they are caught up in the collective action of living a shared life. In a straightforward sense, we can see, then, that there is a contrast between, on the one hand, the antagonism between Crusoe and the natives, and on the other, the harmony that exists between Crusoe and Friday. Crusoe and the natives try to outwit and destroy each other. In contrast, Crusoe and Friday try to assist and benefit each other. However, not all collective actions are harmonious in this sense. Think for example of the collective act of playing chess; the players do not seek to assist and benefit each other, rather they seek to battle and beat their opponent. Nevertheless, even in such cases, there *is* a sense in which there is a harmony: the harmony that can be seen in the experiences of unity and collective intention. The competitive chess players do not see their general interests as united, however they do see themselves as being united as a *we* with at least one harmonised aim, the aim of playing chess together. Harmony in this qualified sense is not harmony of interests, rather it is *harmony of wills*.²⁰⁴ That is, it is the harmony of the creation of a joint goal which all parties join in *willing*.

But recall that this phenomenon – harmony of the wills – is not the full story. Even for Crusoe and Friday, who (unlike the chess players) are not competing against each other, there is also the possibility of a tension that threatens to oppose each individual will. This is experienced as constraint and detachment, as described in the previous chapters. Such tension is possible because in creating a harmony of wills from their individual wills, they create something which can be in opposition to their individual will. Because this plural will can constrain those individuals and because it can seem to come apart (in its content) from the individual wills, it can thus be in discord with them. In this sense, *the possibility of*

204 If we are being precise, it is the harmony of wills *within a given scope* defined by the collective act in question. For example, the chess players will see themselves as united during their game of chess but if they, say, had to abandon their game because of a fire alarm, they may not feel themselves to be socially united (with their opponent) while standing in the car park.

harmony of wills creates the apparently converse phenomenon: the possibility of discord between the wills.

With these notions of harmony and discord we can gain greater insight into the different explanatory attractions of the theories that this thesis has worked through. To recap the argument advanced in this thesis so far: the experiences, which underlie the special sociality of collective action, are unity (our feeling bound together), collective intentionality (our feeling that we have a joint purpose), constraint (our feeling obliged by this collective purpose), and detachment (our feeling that this joint purpose can be separated from our individual purposes). In attempting to make sense of this special sociality, I first swung from the ontologically implausible separate collective entity to the descriptively inadequate mere sum reduction. Dismissing both, I presented Michael Bratman's Shared Cooperative Activity Account. However, hampered by the framework of individual agency, this account was found not to fully get to grips with the phenomenology, and thus I argued for its replacement by Margaret Gilbert's Plural Subject Theory. At its core, Gilbert's theory is based on the idea that individuals come to act collectively as plural subjects, by forming joint commitments to do so. I argued that Gilbert's theory captures the special sociality of our experience without positing any mysterious separate entities. In order to understand more fully what Gilbert says about the nature of joint commitment as a type of commitment *of the will*, in what follows I will now explore how we can think about the comparison between these competing theories in terms of the notions of harmony and discord.

Firstly, the total reduction to individual acts: while such a reduction cannot explain the totality of the unity of collective action, it captures something of the harmony between wills by requiring that all wills aim towards goals of the same type. Though, as we have seen, such harmony does not in actuality fit with the reality of social life, it apes true sociality in that it involves wills harmonised in facing the same way, even if not properly harmonised in becoming one. Total reduction thus partially mirrors harmony, however, it does not fit at all with the possibility of discord; the individual wills all pointing towards the same goals cannot be at odds.

Secondly, the ontologically distinct collectivity (i.e. the mysterious social spirit or the emergent organic entity): this can be seen as a partial possible explanation of the discord involved in collectivity, for in setting up the notion of some separate ontological realm of social reality we allow for something that can stand apart from the individual and their will; something that can be a general will in opposition to it. Distinct collectivity thus partially mirrors discord, however it does not fit at all with harmony; the individual wills are superseded rather than harmonised.²⁰⁵

Thirdly, Bratman's summative theory: this is able to deal in a much fuller way with the idea of harmony; it is able to explain the way in which we think of collective action as harmonising our differing interests and aims because it makes those aims interdependent. Further, (by stipulation) it requires that our individual wills do not merely have equivalent contents but also that they have contents that are united in aiming towards achieving something together. On his account, harmony is a matter of the interdependence of the intentions of the participating individuals, the fact that each plans to perform their own activity in accordance with the meshing plans of the others, as well as their being directed towards a singular goal. It does not, however, allow for the possibility of discord between the wills of individuals and the interests and aims they come to share in, for, as in the total reduction, the individual wills must be all pointing the same way and thus cannot be at odds.

Lastly, Gilbert's Plural subject account: what makes Gilbert's account special is that it does not lean, á la Bratman towards harmony of the wills, nor need it lean, á la the notion of an ontologically distinct collectivity towards discord between the wills. Rather, it can *encompass them both* – it is, in this sense, a *synthesis* of them. On this account wills are

²⁰⁵ This is perhaps the sense in which political theories which have a similar form might be considered to be totalitarian; they are seen as permitting the imposition of a discordant general will onto the people, while not allowing for the liberal possibility of harmony between wills. For example, Benjamin Constant's criticism of Jean-Jacques Rousseau's notion of the legitimacy of the general will can be seen in this light. Constant complains that “[Rousseau] forgets that all the life-preserving properties which he confers on the abstract being he calls sovereignty, are born in the fact that this being is made up of all the *separate individuals* without exception.” (2003, p.28) [*emphasis mine*].

harmonised in a very real sense as individuals pool their wills; individuals commit together and they are together bound by this commitment. At the same time the possibility of discord is also alive. This is because the plural will need not reflect a mere sum of individual wills, rather, it need only reflect the aims and interests which the individuals are willing to commit to sharing in the willing of.

Gilbert's account achieves something the others do not; it encompasses both the existence of harmony of wills *and* that of discord between wills. This way of setting out Gilbert's achievement emphasises its strengths; however, it also makes stark the challenge it faces. In order to really be a synthesis of harmony and discord, Gilbert's account requires a literal understanding of the notion of pooling of wills. That is, it must not just be a convenient shorthand for talking about distributed wills of the same kind, rather it must really be the creation of something over and above this. Some might well argue that failing to encompass harmony and discord, as more individualistic accounts seem doomed to, is less troubling than accepting the anti-individualistic idea that individuals can *actually* pool their wills. To meet this challenge, we must first return to clarifying the notion of commitment that is at stake and its place within Gilbert's theory.

4.2 – Harmony of commitments of the will

The notion of commitment is key to Plural Subject Theory. Becoming committed to the collective act is the way in which we are able to intentionally bind ourselves together and become plural subjects. Commitment, we might say, plays the role of harmonising the wills of the participants. This way of expressing the centrality of commitment brings to the fore the following important question: what is the relationship between the notion of 'commitment' and the notion of 'the will'? This section will explore this question. As we shall see, for Gilbert, the relevant notion of commitment encapsulates the activities of *deciding, intending, planning, aiming to* and such like.

The word 'commitment' has wider and narrower usages. Dictionary definitions tend to distinguish between meanings such as: being steadfast with fixity of purpose;

binding oneself intellectually or emotionally to a course of action; making an engagement by contract; and making an official pledge.²⁰⁶ However, Gilbert's notion of commitment is different from all of these; it can be thought of as providing the core element that unifies these phenomena as all being kinds of commitment. That is, it provides the common thread that runs through them all. As we shall see below, Gilbert's notion of *commitment of the will* need not have the features that we might commonly attribute to commitments of the more enlarged conceptual type (encapsulated in the list above). The type of commitments that Gilbert is interested in, need not be well considered; they need not be inter-personal; and further, they need not be morally obligatorily. Carving away these features from the wider notions leaves us with Gilbert's notion, of which she gives the following definition: "... one has a commitment of the will if, simply by virtue of an act or state of one's will, one is *bound in the way that is common to decisions, intentions, and efforts*".²⁰⁷

We can interrogate our understanding of this notion of commitment by examining an example Gilbert gives of an agent making such a commitment. Let us imagine that, the day after she has finished walking to the top of Scafell Pike, one of our intrepid hill walkers, Jane, is thinking of a new activity to pursue. Let us suppose that she decides to go fishing. Having so decided, Gilbert says that, Jane is "... now *personally committed* to going fishing, as long as she does not change her mind."²⁰⁸ Here Jane, by virtue of her own will, commits herself to a certain course of action; this is the sense in which we can say that Jane makes a personal commitment. Of course, this is not to say that Jane has definitively closed the question of what she should do on that day. Perhaps new information might come to light about the relative fun to be had playing ping-pong in contrast to fishing. She remains rationally free to rescind her commitment and, again by virtue of her own will, become no longer committed.

206 Gilbert also notes that the term is used in an altogether different sense by some economists to indicate *any* incentives to behave in a particular way. (2007, p.261) It is unclear to me that this peculiar usage is anything other than confusing.

207 Gilbert, 2006, p.131 [*emphasis mine*].

208 Gilbert, 2000a, p.3 [*emphasis mine*].

Examining the different ways that we could fill in the details of the story allows us to see that the commitment in question need not be well considered, interpersonal nor need it be moral. In regard to the first of these, consider the following: we might imagine that Jane has been thinking about taking up a new hobby, in addition to walking, for a long time – she has considered the merits of a myriad of different sports, their relative costs, the level of skill they require, the social status (stigma or prestige) attached to them, and so on. Having thought about all this, she comes to a decision to go fishing. In such a scenario, the activity she has personally committed herself to matches that which she has – all things considered – most reason to do. However, Jane need not necessarily be as reflective; we might imagine that upon unexpectedly passing a sports store with a fishing kit in the window, she simply makes a snap decision to go fishing. In such a case, going fishing is not the action that Jane believes she has most reason to perform, it is merely one that she commits to through an immediate moment of minimally reflective choice.²⁰⁹ When Jane has made the decision on a whim, she may have more liberty to change her mind compared to if she has spent a lot of energy coming to the decision; however in both cases if she *does not* change her mind then she is committed to go fishing nonetheless.

It is also clear from the example as it stands that the commitment need not be interpersonal: no others need be involved – Jane can commit to go fishing alone. As a comparison, we might imagine a situation in which Jane commits to go fishing with her friend Bob. This scenario would also involve a commitment of the will. Additionally, it would perhaps also involve a commitment in one of the wider senses, discussed above. For example, if we suppose that Bob is paying Jane to come fishing with him, then Jane might be seen as making a contractual commitment with him to be his fishing partner. Such considerations would add ways in which we could assess Jane as an agent. They would not,

209 Of course, there *are* differences between the two cases that are relevant to the issue of our assessment of Jane as an agent. The seriously contemplative Jane appears to have made a better decision, or at least employed better decision making processes, than the impetuous Jane. Though not necessarily, as pressures of time and limited resources may mean that Jane is better off making snap decisions than spending her precious time deliberating. Similarly, we may well form different expectations about each Jane, e.g. we might expect that the seriously contemplative Jane is more likely to carry out her decision than the impetuous Jane. However, despite these differences, both of these cases are examples of commitments of the will, though it may be that one case counts more typically as a commitment in the wider sense outlined above.

however, show that being interpersonal is essential for all kinds of commitment. Another possibility is that Jane's going fishing with Bob is not a contractual commitment – as such – but rather that it is a collective action. If this is the case, then, as I shall explore below, on Gilbert's account the commitment in question *would* be a commitment in the sense we are interested in – i.e. a commitment of the will.

Lastly, that the kind of commitment in question is not a matter of morality can be seen in the following: suppose that we take seriously the idea that fish can feel pain and that their pain might be morally important. If this were true then plausibly Jane ought not to make a decision to go fishing. However, this does not show that Jane *cannot* make a commitment of the will to go fishing, rather it merely shows that, morally speaking, she *ought not to*. It does not show that Jane's intention to do something that morally questionable is any less a real commitment of the will than a decision she might make to do something we consider to be morally more worthy. Indeed, it is because we think that in holding the intention to fish that she is (in the sense we are interested in) committed to do so, that we think that morally speaking she is under pressure to rescind her decision. One *ought not* to commit to do what is bad, but that does not mean that one *cannot* so commit. Generally speaking, we can say then that, questions about the moral value of intended activities speak to the issue of what intentions one ought to form, rather than to the issue of the force of intentions once they are made.²¹⁰ As Gilbert notes, commitments of the will can be “... forces for both good and evil: we can intend to save the world, and we can intend to destroy it”.²¹¹

Commitments in the wider sense (such as duties, contracts, promises and such like) exert what we might call *external pressure* on the will. If it is a moral good to feed the poor then this fact implies that I should change my will such that I form the commitment to feed the poor. In contrast, commitments in the sense Gilbert uses the term are not external pressures to change the will; they are facts about the way the will is already orientated and

210 Gilbert notes that this discussion may be confused because talk of “... what one ought to do” may be thought to imply obligation in the sense that is derived from some moral principle, i.e. to be moral requirement. She agrees that “... it is impossible to have an obligation of *this* type to do evil.” (2008, p.507).

211 Gilbert, 2008, p.487.

the rational requirements that it being so orientated creates.²¹² They are *internal pressures on the will*. This is the sense in which, as Gilbert puts it, they are commitments *of* the will.²¹³ As Gilbert makes clear, when we say that Jane, having decided to go fishing, ought to go fishing; “[t]he 'ought' here is a matter of what might be referred to as a *rational requirement*.”²¹⁴

Commitments of the will can be said to have what we might call *normative powers*. The notion of normativity is a slippery one, and subject to much debate that is beyond the scope of this thesis. For our current purposes, it will do to understand normative power as anything that gives rise to an *ought* or an *ought not*. Understood in this way, we can see that the normative realm is wider than the moral realm; for example, if fishing is cruel then for moral reasons Jane ought not to fish, but we can also say that if Jane is attempting to perform the ascent of Everest in pure alpine style then, even though she would break no moral law in not doing so, she ought not to make use of additional assistance for reasons of proper sporting style.²¹⁵ In terms of commitments of the will, we can say that, if Jane has decided to go fishing, and continues to hold to this decision, then she ought to go fishing. It is important to note that, these differing *oughts* need not be taken to be all-out universal imperatives. They give reasons for acting, but this does not mean that they supply conclusive reasons. Thus, when Gilbert talks about commitments generating obligations, she means only that, to have an obligation to act is to have sufficient reason to act.²¹⁶ We might thus say that they provide *pro tanto* reasons. This normative power means that commitments-of-the-will trump mere inclinations. For example, Jane's decision to go fishing means that she ought to ignore her inclination to stay in bed.²¹⁷ Failing to do what

212 A complication of this picture is that certain views of moral obligation, such as Immanuel Kant's, *do* try to make our moral obligations into facts about the nature of our will. If Kant were right that the categorical imperative is a function of our rational will, then, the distinction between moral commitments and commitments of the will collapses. Nevertheless, even if this were the case, there would still appear to be important distinctions between the universal demands that the categorical imperative imposes – on *all* rational wills – and the particular demands generated by particular commitments of the will – on only those *specific* wills. My *internal/external distinction* could thus be replaced with the *universal/particular distinction* whilst retaining the import of the distinction.

213 Gilbert, 2006, p.127 [*emphasis mine*].

214 Gilbert, 2008, p.501.

215 Confusingly, and no doubt reflective of the general muddiness that exists around conceptions of the normative realm, climbers refer to such rules of sporting style as 'ethics'.

216 Gilbert, 2006, p.29.

217 A fact I discuss further in the following chapter.

we intend to do, is wrong in a way that, failing to do what we feel inclined to do, is not wrong. The former is always a weakness of will. The latter at times may be an expression of strength of will. This is the sense in which Buno Verbeek has us imagine that Ulysses makes a decision to bypass a pub, and is rational in doing so, even though he desires to go in and drink with his friends.²¹⁸

Above, I noted that we might divide commitments of the will into various different categories, such as decisions, intentions, and efforts. Unfortunately, these categories, informed as they are by common usage, are not all that well defined. One possibly robust distinction is that between intentions and decisions. Gilbert sees the dividing line between the two as being that, “[w]hile a personal decision may be characterised as an *act* of will, a personal *intention* may be characterised rather as a *state* of will”.²¹⁹ Further, Gilbert thinks that “... decisions but not intentions have trans-temporal reach”.²²⁰ That is, decisions continue to stand unless explicitly rescinded; however, it is possible for intentions merely to fade away by failing to be continued to be held.²²¹ This seems a reasonable potential distinction. That said, as I have noted, meaning is a slippery thing. In competition with Gilbert's suggested way of carving intentions as a particular form of commitments of the will, there is an alternative sense in which we might say that Gilbert's commitments of the will are synonymous with intentions. This can be seen by the fact that it is acceptable to reply to questions like “what do you intend to do?” by saying “I have decided to do ... [such and such]”. This appears to imply that both terms are interchangeable. In contrast with Gilbert, Bratman is happier to use the term intentions as a 'catch all' blanket term to cover all of these.²²² In the next chapter – in which I discuss Bratman – I shift from the notion of 'commitments of the will' to that of 'intentions'. But note, I shall take the sense

218 Verbeek, 2007, p.150. The use of the character of Ulysses is a nod towards the story in which he has himself tied to the mast of his ship to avoid the temptations of the sirens.

219 Gilbert, 2006, p.130 [*emphasis in original*].

220 Gilbert, 2008, p.501.

221 For terminological simplicity, I will however view both of these processes as types of rescission.

222 While generally Bratman's use of the term 'intentions' covers all commitments of the will, he does, at least at one point, specify that 'plans' denote “intentions writ large” and 'policies' denote intentions that are vaguer and hold over a longer time. It is also worth noting that there is a meaning of 'plans' which does not fit Bratman's or Gilbert's usage, such as the plans *to* a building or a plan *for* making jam. As Bratman says, “I might know a procedure to achieve a certain end. In this sense I can have a plan to roast lamb whether or not I actually intend to roast lamb. On the other hand, for me to have a plan to roast lamb requires that I plan to roast it.” (Bratman, 1999f, p.29)

of 'intention', used there, to be equivalent to 'commitment of the will', i.e. as a general term that covers the linked phenomena of intentions (in a narrower sense), decisions, plans etc. The discussion in this current chapter is framed in terms of 'commitment of the will', rather than in terms of 'intentions', for the exegetical reason that Gilbert places the term 'joint commitment' at the heart of her setting out of Plural Subject Theory. It is thus important that the discussion makes it clear that Gilbert's notion of 'joint commitment' *does not* involve moral nor contractual arrangements, but is rather a question of direction of will.

So above I introduced the notion of harmony of wills as an element of collective action. Now that we can see commitment, as Gilbert uses the term, to be *commitment of the will*, we can understand why she thinks that joint commitment can be understood as pooling wills. Joint commitment is *joint commitment of the wills* – or to put it another way, it is *joint willing*. One might wonder how it is possible that all the members of a group can come to participate in making a singular commitment-of-the-will together? I gave the foundations of Gilbert's answer to this question in the last chapter. Gilbert believes the acts of each member expressing willingness to constitute – with each other – a plural subject combine together to create a joint commitment. The idea is that the expressions of willingness are not in themselves commitments, rather they become a singular commitment, to act as a group, when they are combined together. As Gilbert puts it, “[e]ach person expresses a special form of *conditional commitment* such that (as is understood) only when *everyone* has done similarly is *anyone* committed.”²²³ Now that we understand that the commitment at stake is commitment of the will, we can see more clearly that she must mean to claim that the combination of expressions of willingness *unifies the agency of the participants*. That is, it creates a will which is *the will of all of them together*, but of none of them apart. Thus we can say that Gilbert sees joint commitment as also being a type of commitment, but a commitment that, rather than being *personally held* by single individuals, is *jointly held* by multiple individuals and thus “... unifies people in a very real way”.²²⁴

223 Gilbert, 1992, p.7 [*emphasis in original*].

224 Gilbert, 2000b, p.30.

As we have seen in the previous chapter, when power relations are in play (as they arguably are, to a greater or lesser extent, in all social relationships) the holding of a collective commitment can be a complex affair. In simple cases, a collective commitment may be created by the direct involvement of each member of the collective. In complex cases, a collective commitment may be created through the activities of a subset of the membership.²²⁵ The uniting factor between all these cases, if they are genuine cases of collective action that is, is that they are characterised by all the group members together comprising the subject of the joint commitment *as a whole*.

4.3 – Discord between commitments of the will

The picture we have, then, is of our two individual walkers surmounting the limits of their individual agency. They do so by harmonising their separate wills to create a single, internally harmonious, plural will – specifically, a plural will with the goal of reaching the summit of Scafell Pike. Now, let us suppose that when being planned the day before, the hard slog to the summit seemed like a great idea to both of them, but that half-way up the first hard step climb one of the pair changes his mind – or as we might put it, he changes what he individually wills – he personally decides to turn back. Here we have a situation of discord; one individual's will no longer meshes with the plural will of which he remains a part of. A key difference between collective commitments and individual ones, as is apparent in such a case, is that while we understand individuals as normally able to absent themselves from the duties imposed by their own personal commitments merely through rescinding them, collective commitments can only be escaped by being rescinded collectively.²²⁶ On this model then, the errant walker is reprimandable *directly* because what their declared individual commitment demands of them (i.e. to give up on the walk) conflicts with what the commitment they hold collectively with the other walker demands of them (i.e. to continue to the top of the hill). Moreover, they cannot absent themselves

225 I discuss such complications in Chapter 3, Section 3.5.a.

226 I use the term 'rescinded' here in a wide sense that includes the possibility that some kinds of commitments can be rescinded by fading out while others require active rejection (see Gilbert's distinction between intentions and decisions discussed above). (p.101 inc' fn. 221)

from this criticism by abandoning the collective commitment, for it can only be *rescinded jointly* with the other walker.

We can see, then, that the very process of harmonising one's will with that of others gives rise to the possibility of discord, that is, to the possibility that the detached collective will can constrain the individual will. On this description there is a direct interplay between the wills; the collective will is directly providing the framework within which the individual members can decide how to act. A helpful way in which to understand the mechanics of this is in terms of Abraham Roth's notion of *practical intersubjectivity*. This is the idea that our practical rationality is not just sensitive to the demands of our own subjective circumstances, but rather that it can be restricted by the willed states we share with others. Further, it is the idea that this can be the case without the need for the individual reissuing of that other's willed states.

Practical intersubjectivity may well involve the kind of enlargement of the concept of the individual practical self that will make individualist-minded philosophers recoil in horror. Nevertheless, Roth believes that it is evident in the way we normally treat certain social situations. Roth sets up a simple scenario where the will of one individual appears to unproblematically directly impact on the practical deliberation of another. He has us suppose that two people, let us here call them Nina and Jack, have decided to drive to Vegas together. This journey will require each of them drive in turn. We can imagine that Nina takes it upon herself to decide that she will drive the second leg and, thus, that Jack will drive the first leg. Let us suppose that on hearing this Jack does not object and goes ahead and drives the first leg.²²⁷ Roth thinks that a natural way to understand this situation is that Jack simply does not concern himself with resolving the question of the order in which he and Nina should drive. Rather, he sees this as already having been settled. He sees Nina as having resolved the issue *for both of them*. Because Jack and Nina are to perform a collective action, namely driving to Vegas together, and we might add, though Roth omits this, because Jack accepts Nina's authority to set their collective intentions.²²⁸ To use the

227 Roth, 2003, p.66.

228 Or more precisely, Jack accepts Nina's authority to set the collective intentions with regard to achieving this trip's collective goal of driving to Vegas. It may be that Jack is always happy for Nina to act in this

terminology developed in the last chapter, Nina is an operant member of the collective that is composed of herself and Jack. The key point is that Nina's intention to drive the second leg figures in Jack's reasoning "... as a rational constraint rather than as a mere consideration".²²⁹ What this also means is that in situations of discord we have a problem that is not merely a coordination issue; rather, we have problem that generates the possibility of irrationality. The irrationality of being obliged towards conflicting actions by practically conflicting commitments. That is, we have a problem of being obliged to act in one way by the collective will and another by one's own will. We might call this *intersubjective irrationality*. This phenomenon can be seen by imagining that Jack and Nina each hold, and know each other to hold, conflicting intentions about who will drive. If neither is willing to change their intention in the light of the other's intention, then there is a tension. This is because, at a group level, it is irrational for these conflicting intentions to both be held.

In contrast to the situation described above, socially interacting people can have conflicting commitments in a way that does pose a coordination problem, but does not equate to a situation of intersubjective irrationality. For example, Roth points out that there is nothing irrational about me having an intention to buy a car and you having an intention to buy the same car, even though our intentions are incompatible and we both know them to be so. Such a scenario does generate a problem, but this problem is practical rather than rational; only one person can win, but this makes neither irrational in trying to beat the other. In contrast, just as Jane (discussed above) violates a practical norm if she intends to go fishing but then does something other than fish, Nina and Jack violate (at least at the collective level) the same practical norm if together they intend to drive to Vegas, but they fail to mesh their conflicting intentions as to who shall drive. Jack and Nina's conflicting intentions are unlike those of the two agents engaged in a battle to buy the single car because their conflict is *within a united will* (that jointly intends to go to Vegas) rather than just between separate wills (that each intend to buy the single car).

unilateral fashion, say if she holds a position of general authority over him. Alternately, it may be just that in this instance, Jack cedes authority to Nina just because at this time he has no desire to have a say in this particular decision.

229 Roth, 2003, p.67.

An alternative way to explain the irrationality in the situation above is to suppose that each has a supplementary-intention that their intentions will coordinate with those of the other party. So, in the case of the two people intending to drive together to Vegas, it is not directly the fact that their intentions, regarding driving the second half of the route, are in conflict that makes them irrational. Rather what makes them irrational is that each individual's intention to drive the second half is in conflict with *their own* supplementary intention, i.e. in conflict with each individual's intention "... to coordinate intentions pertaining to driving to Vegas together".²³⁰ The problem with this 'bridge intention' solution, as I have already discussed in Chapter Two, Section 2.5, is that it does not seem to make the link between the intentions of the two individuals strong enough. Because of this, it does not fit with the collectivity of the phenomenology of collective action (as has been described in this thesis). Simply put, because the supplementary intention is an individual intention, each individual remains rationally at liberty to rescind it. As Roth notes, "[b]y revising my bridge intention, your intentions no longer make any claim on me"²³¹. In the driving case it seems like, the fact that the two participants are doing something collectively, ought to mean that neither have the power to unilaterally remove themselves from the rational constraint of their collective activity. The important point is that the intended driving action is not that of two separate but cooperating individuals but it is rather a joint action. The important thing is not the inconsistency between the two drivers' intentions as such, but rather the inconsistency between each driver's intentions and a harmonised collective intention. Thus, the rational inconsistency is not (directly) between the intentions of the two individuals, rather it is between the individual's intentions to play their parts in the collective action and what is necessary for the collective action to take place.

By moving from a focus on conflict between individual commitments to conflict between individual commitments and the joint commitment, we can see that there is a possibility of intersubjective irrationality even in situations where there is agreement

230 Roth, 2003, p.77.

231 Roth, 2003, p.80.

between the individual's commitments. So, for example, suppose that both Jack and Nina secretly did not want to go to Vegas, and thus that both intended to refuse to drive. If they both express their misgivings, they would decide, together, to cancel their trip. However, it seems possible that each fails to openly express their individual stances. If they do this, then, on Gilbert's view, the joint commitment will stand, and thus, if we take collective rational constraint seriously, they will both be intersubjectively irrational in intending to refuse to drive. This move also makes it easy to generalise this model from one about two-person collective actions to one that can cover many-person collective actions. Multi-person models, as I have noted at various points in this chapter, involve questions of authority, hierarchy and social power; intersubjectivity in multi-person cases can come about in situations where an individual stands in control of others.

We can now understand more clearly what it means to think of the force of the constraining power of the collective will as direct; it is direct inasmuch as it operates intersubjectively for multiple subjects in the same fashion that the individual will operates internally for the individual subject. The collective will in this sense has *intersubjective normative force*; its force is direct – or as we might say, *internal to the will* – rather than externally constraining on it. A bridging intention, belonging subjectively to an individual agent, would also be directly internally constraining to the will of its owner. However, because it is completely within the subjective control of the individual, it can be ditched by them. In contrast, joint commitments are both internally constraining and *not* controllable by the individual. So, the wrong that we are interested in – the wrong of breaking with a collective commitment – is not that of betraying another person in a moral sense (though of course in some collective action situations this may also additionally exist), rather, it is a *rational wrong*, it is *betrayal of the will*.²³²

232 See Gilbert, 2006, p.134 for discussion of the powerful (but potentially unhelpfully ambiguous) notion of betrayal in this context.

4.3.a – The puzzle of intersubjectivity

Collective action involves committing together and together being bound by this commitment. Further, this is something that each individual makes happen; as Gilbert puts it, “[f]or each party to enter into a joint commitment is for him to *allow his will to be bound*”.²³³ However, though Gilbert's theory requires that each individual jointly commits, because their will is thus collectively pooled they can then be directly normatively constrained in their action by intentions, other than their own personal intentions. This, depending on your outlook, is either Plural Subject Theory's biggest advantage or greatest problem. Those of an individualist bent might argue that this is too high a price to pay. The fuzzy, but nonetheless intuitively forceful, idea which underlies their complaints is that to allow practical motivation to arise from a location external to the individual agent is counter to the fact of *individual autonomy*.

The notion of autonomy is at once straightforward, in that it exists in common usage rather than just obscure philosophy, and hard to pin down, in that, the exact content of what it means is often disputed. At the core of the notion of autonomy is the notion of being in control. Sarah Buss cashes out this general sense of autonomy by saying that “[t]o be autonomous is to be a law to oneself”.²³⁴ Within this, however, we can separate wider and narrower usages. Autonomy, according to the wider usage, requires that we govern our external environments; let us call this *environmental autonomy*. In contrast there is a narrower sense of 'autonomy' that requires only that we count as governing our own actions, or as Buss puts it, to be free to “... make one's own mind up”.²³⁵ I will follow Buss in referring to this as *personal autonomy*.²³⁶ There is a strong link between environmental autonomy and personal autonomy, but they are different phenomena.²³⁷ Environmental

233 Gilbert, 2006, p.154 [*emphasis mine*].

234 Buss, 2008.

235 Buss, 2008.

236 Buss, 2008. The notion of autonomy that is perhaps most discussed is that of 'political autonomy'.

Political autonomy appears to be autonomy in a wide sense akin to what I have called 'environmental autonomy'. However, depending on the account of political autonomy we focus on, it may apply more exclusively to freedom in relation to social laws, customs and other institutional rights.

237 Buss makes a similar point in different terms when she says that “...every agent has an authority over herself that is grounded, not in her political or social role, nor in any law or custom, but in the simple fact that she alone can initiate her actions.” (2008)

autonomy requires personal autonomy because, we might reasonably assume, you can only count as governing the world if you count as governing your behaviour within it. But the reverse is not the case. Imagine a woman locked in a prison cell; she has no control over the time she is allowed to sleep, who she is allowed to see and so on. This person has lost autonomy in the environmental sense, yet those limited actions that she can perform – such as pacing back and forth – still count as being personally autonomous. In contrast, if we were to imagine that our technologically advanced prison guards had a way to manipulate our prisoner's mental states – say to control her mind such that she wants to commit to the rule of the dictatorial regime – then she would not just fail to be environmentally autonomous, she would also fail to be personally autonomous. In what follows, I will be using the term autonomy in the limited sense of personal autonomy.

The complaint against Plural Subject Theory is not just that the constitution of a pooled will seems a little mysterious. Hopefully, I have done some work in dispelling this mystery; at least to the extent that any remaining mystery is no greater than that which concerns the constitution of individual wills.²³⁸ Rather, the problem is that, the notion of individuals being guided by collective commitments seems to be in conflict with what, for individualists at least, is a key part of our understanding of ourselves as autonomous agents. Hans Bernard Schmid invokes the notion of zombies to characterise such a phenomenon; people not in control of their own intentions he thinks of as *intentional zombies*.²³⁹ The image of zombies has a certain resonance for they are seen as creatures who have lost the ability to truly govern their own worlds; they have become 'mindless brain eaters'. However, it seems to me that the imagery of puppetry is even more appropriate for our discussion of autonomy, for it conveys not just the idea of lacking in control, but, importantly, also that of being *controlled by others*. Using the terms explored above, we might say that the fear is that Plural Subject Theory requires that engaging in collective action must rob us of our personal autonomy and turn us into *intersubjective puppets*. Intersubjective puppets appear to lack personal autonomy. In the illustration above, Jane

238 Indeed, in the coming chapters, I will argue that the remaining mystery of the plural will can be lessened even further by applying the same techniques as can be applied to lessening the mystery of the individual will.

239 Schmid, 2009, p.19.

freely chooses to go to fishing. No doubt there might be many complex psychological explanations as to why she does so. However, the following minimal description is intuitively compelling: her actions are at least personally autonomous; that is, they are governed by her. Now, imagine that Jane is part of a collective that is intending to go fishing. Following the plural subject model, Jane is thus jointly committed to playing her part in the collective fishing action. The question is: if Jane acts in line with this pooled will, can we still understand her actions as personally autonomous? Would such an understanding reduce her to a mere intersubjective puppet, when we should understand her as a fully-fledged agent? This is the *puzzle of intersubjectivity*.

The plausibility of the idea, that it is worrying that an account might involve individuals becoming intersubjective puppets, can be seen by imagining a modification of the rugby club example: suppose that in his rebuke the keen bus pusher had made reference to his own personal goal rather than to that of the team. Rather than saying, “You can't stop, *we* said *we* would push the bus”, he says, “You can't stop, *I* said *we* would push the bus!” Now let us assume that there are no issues of hierarchy at play here, e.g. that the presumptuous team-mate is not the team captain. Further, let us also assume that there is no standing commitment to have his personal pronounced goals automatically count as the goals of the collective. If this is so then there may still be *indirect* ways in which this kind of phrase might sometimes be an appropriate, even if rather rudely put, rebuke. For example, if the errant team member had promised to do whatever the keen member wanted. However, suppose that we try to account for it in the same way that the PST account proposes that we account for the rebuke attached to the collective commitment. That is, suppose that we propose that the rebuke is legitimate because one individual can be *directly* normatively constrained by the individual intentions of another. Against such an explanation, it certainly seems fair to complain that the individual's autonomy is not being taken into account. *We appear to have reduced the agent who is constrained by the intention of the other into a mere puppet controlled by that other.* We can take Gilbert as acknowledging that this is how things are, when she says that in situations that are not collective actions, “I can persuade you to change your mind, but I cannot directly change it.”²⁴⁰ This case is

240 Gilbert, 2008, p.491.

clearly not completely analogous with that where the keen team member invokes a shared collective intention. Whereas, the errant team member is a part of the collective that holds the collective intention, in contrast, he plays no equivalent part in the constitution of the keen team member as an individual. It is not immediately clear, however, what exactly it is about this difference that might make the concern about autonomy disappear. One might argue that, just as in the presumptuous walker example, the individual has become a puppet of something external to them, in this case a puppet of the collective.

The puzzle of intersubjectivity is a puzzle precisely because we *do not* think that people have become puppets in any strong sense when they perform the kinds of collective actions that have been documented in this thesis so far. We can see this by noting the heavy contrast between normal cases of collective action and science fiction examples, such as the Borg in the television series *Star Trek: The Next Generation*. The Borg are all connected to a collective hive mind and are clearly all puppets of its singular will. The Borg do perform collective action, but it is not collective action in the sense that we perform it; for them there is no possibility of discord because there is no individuality left – rather they have a total harmony of the wills, an embodiment of the organic theory of collective action. This is why being assimilated²⁴¹ is such a scary concept for liberal individualists.²⁴²

Perhaps we might counter worries about individual autonomy by stipulating that, as a basic fact of being social creatures, we *just can* be directly motivated by the intentions of others. One might think that such a possibility is ignored only because of an attachment to what we can call the *individual motivation thesis*: roughly that, individuals can only be motivated by their own will.²⁴³ Hans Barnard Schmid suggests, contrary to this theoretical thesis, that our actual social experience is of routinely acting directly on the intentions of others, without seeing this as problematic. He gives the simple example of one agent moving aside on a park bench to fulfil the intention of the other to sit down, claiming that it seems perfectly natural to suppose that the shuffling agent saw themselves as acting

241 The term for being forcefully made part of the Borg collective.

242 Indeed, Lawrence M. Krauss claims that it is what makes the Borg “...the most frightening, and intriguing, species of alien creature ever portrayed on a television screen.” (1995, p.111).

243 Schmid rather pointedly calls it a product of the 'Cartesian Brainwash' (Schmid, 2009, pp.29 – 42).

directly on the intention of the sitter.²⁴⁴ He proposes we label acceptance of this possibility as *motivational heterarky*.²⁴⁵ In a similar vein, Abraham Roth suggests that we accept what he, perhaps more informatively, calls *practical intimacy*. He defines this as the idea that; “[i]t is possible for one individual to take up and act on the intention formed by another without re-issuing the latter's intention”.²⁴⁶ In the terms we have been employing here, we might best call this phenomenon: *the possibility of intersubjective motivation*.

How does the possibility of intersubjective motivation sit with our concerns about autonomy? Well, Schmid believes that there is no reason to think that they pose any challenge to autonomy, for an individual can still be fully autonomous even when acting on the intentions of others, because it can still be 'up to them' whether they act or not. For example, in the case I described above of one agent, *A*, moving aside on a park bench in order to fulfil the intention of another agent, *B*, to sit down: “[i]t is not that *B* somehow acts *directly* through *A*'s behaviour, bypassing and displacing *A*'s agency ... Rather, *A*'s behaviour still instantiates *A*'s own action”.²⁴⁷ Likewise, the case I first used to characterise one of the four primary features of the phenomenology of collective action might also be examined here: the Leicester Tigers pushing a bus up a hill. We could say that the autonomy of each member of the Leicester Tigers, is not challenged by their being externally moved to push the bus by the collective intention to get the bus up to the top of the hill, because they each choose to be moved by this non-individual attitude.

Unfortunately, such a reply does not quite address the concern about autonomy in relation to the pooling of wills in the Plural Subject Theory, for (as Roth points out in his discussion of Gilbert's examples) the issue is precisely that PST gives rise to the possibility of cases where it is *not* up to the errant individual whether the collective intention has authority over them. The trouble is that the collectivist model proposes that individuals who have formed collectives are thus normatively constrained by the particular intentions of that collective, *whether or not*, in that instance, they want to be. Roth's own solution to

244 Schmid, 2009, p.19.

245 Schmid, 2009, p.19.

246 Roth, 2003, p.383.

247 Schmid, 2009, p.19.

this problem rests on the idea that in certain circumstances one agent can have authority over the actions of another.²⁴⁸ Such considerations take us in the right direction. However, the issues I raised in a previous chapter, in relation to the possibility of operant members²⁴⁹ complicate this picture. Recall that, we saw that the role of operant members in collective actions does not appear to be that of one individual having authority to set the subjective will of another individual. Rather, operant members appear to have the authority to set the collective will. Given that the collective will is detached from the mere sum of individual wills, the group leader in this case is properly understood, not as having direct authority over any other individual, but rather, as having authority over the collective will. The issue thus remains of what it is that makes it legitimate for this collective will to constrain the individuals, given that this collective will is disconnected from that individual will.

So where does this leave us? Schmid and Roth may be right in their insistence that the mere possibility of intersubjective motivation (i.e. of acting directly on the intentions of others) does not run counter to agentic autonomy; however, we need to go further than this if we are to get to grips with the concern about normative *constraint* by intentions that are not solely one's own. There is reason to think that the fuzzy idea of agentic autonomy does not provide quite as clear-cut an objection to normative constraint by collective intentions as it might first appear. We might think that agentic autonomy is compromised by allowing an agent's free choices to be normatively constrained by anything at all, believing that autonomy is akin to freedom to do anything at any time. However, we need to note that we clearly do not see it as problematic for autonomous agents to be normatively constrained *by their own intentions*. We can thus ask the question: why is an autonomous agent not free to merely act as they please rather than being constrained by their intentions? In the following chapter, I will suggest that answering this question can give us both a model that we can apply to understanding normative constraint by collective intentions, and also bring to light the necessary limitations of this constraint.

248 Roth, 2004, p.391 – 397.

249 Chapter 3, Section 3.4.

As we shall see, Bratman, with his planning theory of agency, provides us with a potential answer to this question. His theory aims to tell us not just why constraint by one's own intentions does not conflict with autonomy, but also why it is fundamental for being an autonomous agent *at all*. Given that collectivist theories propose that we should understand the constraint of collective intentions as being of the same type as the constraint issuing from individual intentions, there seems to be clear motivation for attempting to apply an explanation that mirrors Bratman's account of individual commitments of the will to collective commitments of the will. This is not a use to which Bratman, as far as I am aware, has attempted to put his theory of (individual) intentionality. This should not surprise us, given the way I have characterised Bratman's Shared Cooperative Activity Account.²⁵⁰ Nor should it preclude us from attempting to do so, for his non-collective account of collective intentionality does not directly or necessarily follow from this account of individual intentions.

4.4 – Chapter summary

In this chapter, I have put forward the case that Gilbert's notion of joint commitment is best understood by setting out its relationship to the way in which the wills of agents involved in collective actions are related. I did this first by resetting out the argument of the thesis so far in terms of the ideas of discord and harmony between the wills, arguing that Gilbert's account is unique in its ability to capture them both. I then argued this allows us to understand Gilbert's notion of joint commitment as a *harmony of commitments of the will*. This has the benefit of expressing the real sense in which wills are pooled according to Plural Subject Theory, however it has also brought to the fore the puzzle of intersubjectivity. In the next chapter I will tackle this, and attempt to understand the nature of intersubjectivity using the tools provided by Michael Bratman's planning theory of agency.

²⁵⁰ Recall that, Bratman rejects the collectivist account of collective intentionality. In contrast, he believes, roughly, that collective intentions require the existence of interlocking of conditional personal intentions which have the same orientation towards the collective act, and thus does not allow the possibility of conflict between the collective's intentions and the individual's personal intentions. (See Bratman, 1999, pp.93 – 142). See Chapter Two of this thesis for full discussion of this position.

Chapter 5 – Constraint by the will (individual and collective)

Through taking the phenomenology of our social lives seriously, we have been led to adopt Margaret Gilbert's *Plural Subject Theory*. This account captures the apparently contradictory, but in fact complementary, phenomena at the heart of collective action: *unity between wills* and *discord between wills*. Adopting Plural Subject Theory requires that we move beyond the framework of individual agency. I have argued that it is reasonable to do so because the possibility of individuals uniting together – and so creating a plural subject with a collective will – is part of our basic understanding of social life. The last chapter, however, raised a challenge: how can we marry the fact that the collective will has intersubjective normative force with the apparent autonomy of individual agency? In this chapter, I will be looking to disarm this challenge. I do so, firstly, by setting out how personal autonomy requires constraint by one's own will, and secondly, by arguing that parallel considerations hold in the collective case.

The above will involve attempting to understand the nature of intersubjectivity using the tools provided by Michael Bratman's *planning theory of agency*.²⁵¹ Whilst there are many theoretical points of dispute between Bratman and Gilbert, their debate is often framed as a disagreement about the base phenomenology. Gilbert claims that we experience the collective will as directly constraining us and that we invoke this collective will when rebuking those who transgress. Thus, in Gilbert's world, the keener of our two hill walkers may well admonish the lazy one by saying something along the lines of, "You can't stop. *We* said *we* would go to the top!" Bratman, meanwhile, claims that all that we experience is a pressure towards a certain level of stability regarding the direction of our own will and everyday moral obligations towards each other. Thus, in Bratman's world, the keen walker is more likely to exclaim, "You can't stop. *You* said that *you* would go *with me* to the top!"²⁵² The first chapter of this thesis described how the phenomenology can be seen as

251 The argument set out in this chapter is an extended form of that which I make in *How where I stand constrains where we stand* (Kisolo-Ssonko, 2013).

252 Or at the very least, in the world according to Bratman's theoretical schema, where people do say the former, they can be understood as *really* meaning something like the latter.

firmly on Gilbert's side. However, armed only with such descriptions, we may hit a brick wall when arguing with intransigent individualists who simply insist that the world is otherwise. By giving an account of intersubjectivity that starts with a theoretical understanding of *the nature of our agency*, I hope to contribute towards breaking this apparent stalemate.²⁵³

This chapter is split into two parts. The first part, entitled 'Constraint by the individual will', sets out Bratman's planning theory of agency and his explanation of the power of individual intentions. The second part, entitled 'Constraint by the collective will', sets out the application of this theory to collective intentions. In applying Bratman's planning theory of agency to Gilbert's Plural Subject Theory, I am looking to defend the plausibility of Gilbert's theory in its general form. However, my investigation will result in the need to modify Gilbert's theory. I claim that we must replace Gilbert's idea that pooling of wills takes place as the result of essentially voluntary 'willings', with the idea that the wills become entangled to various degrees over time. I will discuss the full implications of this change to Gilbert's Plural Subject Theory in the next chapter.

Before I continue, a brief terminological note: as I have set out in previous chapters, Gilbert's notion of 'commitments of the will' is *not* akin to the sense of the term 'commitments' in the general (moral or contractual) meaning. Rather, Gilbert's term covers the linked phenomena that are loosely referred to as intentions, decisions, plans, aims and such like. In his discussions of agency, Bratman tends to use the term 'intentions' as a blanket term to cover this same set of linked phenomena.²⁵⁴ Because I will be engaging

²⁵³ As I noted at the end of the last chapter, Bratman appeared to acknowledge the need for a turn towards the nature of agency at the 2012 *7th Conference on Collective Intentionality* where, in the questions section of Gilbert's talk, '*Saving the appearances" with joint commitment*' (CIVII, 2012), he proposed that the way to move beyond these debates over the base phenomenology is to try to build a model of the nature of the individual will and its social analogue. This is the challenge I see myself as taking up. I further clarify the nature of this challenge at the start of section 5b, below.

²⁵⁴ While Bratman does not use the terminology 'commitments of the will', he does speak of intentions as involving "... a characteristic kind of commitment" (1999f, p.15) and can clearly be seen to be pointing towards the same phenomenon as Gilbert. Of course, Bratman does not believe that there is a collective *will* as such, but in this sense the term 'commitment of the will' is not necessarily any more problematic than the term 'collective intention', for strictly speaking Bratman does not think that there can be *a thing* that is the intention of a collective. Nonetheless, just as, on Bratman's account, talk of collective intention is actually to be understood as talk of sharing in having joint intentions, collective commitments-of-the-will could be taken to be seeing people as sharing in collectively willing. Though it is

mainly with Bratman's formulations in this chapter, here I will mostly switch to his convention and use the term 'intentions' as synonymous with 'commitments of the will'.

generally true that Bratman's use of the term 'intentions' covers all commitments of the will, this is not always the case; for example, he acknowledges that 'plans' can be seen as "intentions writ large" and likewise that 'policies' can be seen to denote intentions that are vaguer and hold over a longer time (1999f, p.29). A further complication is that there is a usage of the term 'plans' which does not fit Bratman's or Gilbert's usage, such as the plans *of* the building or a plan *for* making jam. As Bratman says, "I might know a procedure to achieve a certain end. In this sense I can have a plan to roast lamb whether or not I actually intend to roast lamb. On the other hand, for me to have a plan to roast lamb requires that I plan to roast it." (Bratman, 1999f, p.29).

Chapter 5 Part (a) – Constraint by the individual will

In the last chapter, as an illustration of the normative constraint that our *intentions* (or, as Gilbert refers to them, the *commitments of our wills*) place upon our practical deliberation, I gave Gilbert's example of Jane forming an intention to go fishing. Gilbert takes it to be obvious that having formed this intention, and not having rescinded it, Jane can be criticised if her actions and further practical deliberation do not fit with it. This is the sense in which I have set out that agents can be *normatively constrained by their intentions*. Unlike, say, the rope that tied Odysseus to his ship's mast, such constraint does not delimit what is *possible* for an agent. Rather, it sets the boundaries of what is *normatively permissible*. To refer to such constraint as normative is to say that, rather than delimiting what an agent practically might do, it delimits what they ought to do. All this seems intuitively correct, a straightforward description of what we take to be the nature of intentions.²⁵⁵ However, if we are to get to the bottom of the nature of this constraint, we need to progress beyond noting that things *are* so and ask *why* they are so. What is it about intentions that means that the constraint that they issue has real normative force for agents like us?

Another way to put the point is to imagine a wayward sister to Jane; let us call her Anarchist-Jane. Anarchist-Jane is like Jane in that she responds to the practical world by forming intentions. However, unlike Jane, she then fails to see the constraint issuing from her intentions as binding upon her. What exactly is errant about Anarchist-Jane? In what follows, I shall argue that Bratman's *planning theory of agency* holds the key to answering

²⁵⁵ Though note that not everyone accepts this. Bratman acknowledges that it is possible to reject the idea that the normative constraint of commitments of the will is real, calling theories that do so *myth theories*. He attributes such a position to Joseph Raz (2005) and Niko Kolodny (2005) and states that they believe that it is a myth to think of the constraint appending intentions as having any *distinct non-instrumental* normative significance in each particular case (see 2009a, p.419 – I explore the importance of the normative significance being distinct and non-instrumental below). However, while such positions are possible, given that the idea that our intentions do constrain is the intuitively natural starting point, Bratman – rightly, I think – believes that *one should only accept a myth theory if one cannot give a compelling account of the power of these norms* (2009a, pp.418 – 419). Thus, if we can find a workable way to explain what gives the norms their force, then myth theory will be redundant.

this question.²⁵⁶ Broadly put, Bratman's planning theory starts by noting that we have a practical need to structure our lives. That is, to have a rational framework that allows us to avoid having to make practical decisions at every possible juncture. For example, it benefits Jane to see her decision to go fishing as a fixed point as she decides what clothes to wear, what bus to catch and so on. She can determine her actions in light of her standing intention to go fishing. In this way, she does not have to revisit the rationale for this action at each juncture. I will explore the idea that intentions fill this function in the first section below (5.a.1). However, this is not the end of the story, for Jane does not simply see following her standing intentions as generally leading to better practical outcomes; rather, she experiences them as things that (as long as they continue to be held) *ought to be followed in each instance*. In the second section (5.a.2) I will introduce the idea that intentions are experienced as having *distinct non-instrumental force* and present Bratman's notion that this is linked to the way in which intentions allow agents like us (i.e. planning agents) to count as *governing* by giving us a standpoint. That is, by giving us a rational perspective we can identify as being the place “where we stand”.²⁵⁷

Understanding this dialectic will depend on appreciating the interplay between two concerns about the *metaphysics of agency*. Firstly, the concern, raised by Harry Frankfurt, that we need an account of how agents like us can count as authentically governing our own lives.²⁵⁸ Secondly, the concern, originating from John Locke, that we need to identify some state that links the mental life of an agent to account for their identity as a singular agent across time.²⁵⁹ As we shall see, Bratman thinks that by allowing us to have a standpoint, intentions can meet both concerns. Further, he believes that the normative authority of intentions is grounded in the fact that they can only play this role in virtue of being taken to be universally constraining (i.e. constraining in every particular instance).

256 Bratman first began to develop this planning theory in his paper *Intention and Means End Reasoning* (1981), and has subsequently gone on to develop it in his book *Intentions, Plans, and Practical Reason* (1999f) and throughout his work in various papers. Bratman's presentation of this theory is extended and subtle. It is thus open to different interpretations. In what follows, I put forward what I take to be the key elements that make his theory work.

257 Bratman, 2009a, p.431. As I shall explore below Bratman takes the concern with 'where-I-stand' to be a “Frankfurtian concern” (2009a, p.431).

258 See Frankfurt, 1971 & 1997.

259 Locke, 1748.

From this metaphysical point – regarding the nature of agency – I will take it that the epistemic point – that agents can only properly understand themselves as agents if they see their intentions as constraining – follows.²⁶⁰ Thus, the error of Anarchist-Jane will be that by failing to appreciate the constraining force of her intentions, she fails to be able to see herself as the kind of agent she in fact is.

5a.1 – Intentions as rational scaffolding

There can be no doubting that intentions pervade our practical lives. They come in various guises, some are short term – such as my intention to finish typing this paragraph before I pause for a rest; others are longer – such as my plan to finish this chapter by the end of the day; and others are much longer – such as my aim to finish this thesis by the end of the summer. Why, one might wonder, do I bother with any of them? As I began to sketch above, Bratman's answer is that we are the kind of creatures that have both the practical need and the mental capability to organise our lives. According to this picture, intentions are akin to maps; they offer us pre-considered paths across our practical choice-laden landscape. I will call this function that of offering *rational scaffolding*.²⁶¹

The importance of the rational scaffolding created by intentions can be seen by inventing an additional sister for Jane; let us call her Intention-Free-Jane. Unlike their sibling Anarchist-Jane who ignores the force of her own intentions, Intention-Free-Jane simply fails to make any; she lives her life from moment to moment, making choices freely at every possible juncture. What would life be like for Intention-Free-Jane? Well, suppose she lives on a desert island, similar to that of Crusoe discussed in the previous chapter, but

260 Bratman focuses on the metaphysical point, but I take it that the epistemic point flows easily from this metaphysical point. Indeed, Bratman seems to be suggesting as much when he says “When I recognize inconsistency in my own intentions, I see that in this specific case there is no clear answer to the question, ‘Where do I stand?’ This question about myself is, with respect to this domain, simply not settled; there is as yet no fact of the matter.” (2009a, p.431)

261 The project is thus to understand intentions from a first person perspective. An alternative starting point would be to enquire as to the importance of such rules of reasoning in helping us understand *other* agents and their actions, i.e. to understand intentions from a third person perspective. While Bratman recognises that understanding how intentions work is an important component in our interpretive and evaluative assessments of other agents, he sees this as a secondary function of the fact that they structure “first person practical reasoning” (2009a, p.413). Thus, approaching the topic in this former way will allow us to unravel the latter problem.

let us say, so richly endowed with edible plants and free of dangers that it is never much effort for her to fulfil her very basic needs: to eat, drink and sleep. We might imagine that in such a land Intention-Free-Jane could get away with this form of mental life. She sees a piece of fruit and, if she wants to eat it, she eats it, and if not, she does not. She sees a nice patch of ground and if she is inclined to sleep at that moment, she lays down to sleep, if not then she does not. Sadly, the real world is not like this abundant Eden. Moreover, actual people have interests and desires far more complex than merely to eat, drink and sleep. So let us complicate Intention-Free-Jane's life by imagining that she is not happy with her island existence and wants to escape back home. Naturally, she realises she needs to build a ship. It immediately becomes obvious that to achieve this she will need to think beyond her immediate present; she will need to think about her *practical choices as extending across time*. As Harry Frankfurt puts it, she will need to move beyond the "... way of nonhuman animals and of small children."²⁶² The reason for this is that if, when deciding whether to pick up this piece of wood or that, Jane has to every time revisit the question of what she wants to use the wood (rather than see her intention to build a ship as a fixed point), she will fast get bogged down in cognitive overload.

In virtue of the fact that intentions are rational scaffolding, they are, as Bratman notes, a "... more or less all-purpose, universal means" to any end.²⁶³ To say this is just to recognise that you cannot usually directly make it the case that what you intend just comes about; rather, you need to do those things that are needed to bring it about, in order to bring it about. Moreover, if you do not want to be at cross purposes, to bring about things other than that which you wish to bring about, holding conflicting intentions will frustrate you in achieving your aim. These points are obvious when we think about the problems that Intention-Free-Jane would face trying to escape her island. Intentions are defined by being the mental means by which we make it such that we do not have to treat each moment as one where we must decide what to do. Given the above, it is fruitless to ask *why* intentions structure. To ask such a question is like asking why buses accept passengers. For, in the same way that accepting passengers just is part of what it is to be a bus, *providing*

262 Frankfurt, 2006, p.6.

263 Bratman, 2004, p.1.

rational scaffolding just is part of what it is to be an intention. The question that I am asking is rather, why do we feel ourselves to be bound by this rational scaffolding?

According to Bratman, what I have called rational scaffolding operates in terms of the following two principles: *means-ends coherence* and *intention consistency*.

Means-ends coherence is the rule that if we have an intention to *E* then we must commit to doing that which we believe is necessary to do *E*. So, for example, if Jane intends to go fishing, and believes that it is necessary to buy worms to go fishing, then she ought also to intend to buy the worms.²⁶⁴ Bratman gives the following formal definition of this rule: “The following is always pro-tanto irrational: intending *E* while believing that a necessary means to *E* is *M* and that *M* requires that one now intend *M*, and yet not intending *M*”.²⁶⁵ This rule is important because the intentions we form tend only to contain partial plans for their realisation, and because we know that filling in the gaps is necessary as a means for achieving our goals, we must thus commit to filling in these gaps. In this vein, Jane's intention to go fishing may not specify the location at which she will fish, but, because having a plan about where she will fish is a necessary means to going fishing, she will need to form such a plan. Further, if, say, the chosen location is far away, she will need to have a plan about how to get there.

Intention consistency is the rule that if we are committed to one action then we cannot also intend to do some other action that we believe is impossible in conjunction with that intention. So, for example, if Jane intends to go fishing today and knows that this will not leave her with enough time to also get in a round of golf, then she ought not also plan to play golf. Bratman gives the following formal definition of this rule: “The following is always pro-tanto irrational: intending *A* and intending *B* while believing that *A* and *B* are not copossible”.²⁶⁶ Further, in interaction with the rule above, we can say that the agent is under pressure to *fill out their partial plans* in ways that are compatible with their prior

²⁶⁴ Suppose that Jane knows that she must catch the early bus to get to town before the shop selling the worms shuts. However, perhaps because her scoleciphobia (fear of worms) overcomes her, she forms the intention to catch the later bus. In such a case, Jane's fishing partner, Jim, could rightly criticise Jane.

²⁶⁵ Bratman, 2009a, p.413.

²⁶⁶ Bratman, 2009a, p.413.

intentions. Given this, we can see that, in filling out her plan to fish by specifying a location, Jane must take into account her prior intentions – such as, let us imagine, her decision never to fish at the lake where her brother was eaten by a carp.

Of course, the average person might not explicitly acknowledge that they are guided by the principles of means-ends coherence and intention consistency in their practical reasoning, at least not in these terms; nonetheless they are easily recognisable as formalised versions of our everyday practical reasoning (as the example of Jane shows). Thus, when Bratman says that “[g]uidance by our (implicit) acceptance of these norms is central to the proper functioning of our planning agency”²⁶⁷ he seems to be right: implicit acceptance of these principles is thoroughly natural. To say that our acceptance is implicit is just to say that these are not rules which we explicitly have in mind as we engage in practical reasoning; however, they are ways of thinking that can be seen to be embodied in the ways in which we think practical reasoning ought to function. They describe kinds of reasoning that we can recognise and admonish divergence from.

The special nature of intentions in providing rational scaffolding can be further elucidated by comparing them with desires. On one level, intentions and desires are similar; both contain descriptions of certain sets of affairs, and both can be said to motivate us towards realising those sets of affairs. So, a desire to keep fit has the same aim as an intention to keep fit and either would motivate one towards exercise. In this sense, both are, to use Donald Davidson's terminology, *pro-attitudes*, that is, they are attitudes that motivate some set of affairs.²⁶⁸ However, while our desires do motivate us to live in certain ways, unlike our intentions, they do not normatively constrain our actions. Take Jane: she desires to go fishing, but she might also desire to stay in bed, desire to start training for a marathon and desire to become the kind of person who does not like to fish. However, it is not the case that if Jane fails to start training for a marathon, fails to stay in bed, or fails to stop liking fishing that she must have behaved incorrectly. Indeed, as these desires are not

²⁶⁷ Bratman, 2009a, p.413.

²⁶⁸ Davidson lists 'pro-attitudes' as including “... desires, wantings, urges, promptings, and a great variety of moral views, aesthetic principles, economic prejudices, social conventions, and public and private values” (2001a, p.4)

copossible, not only do they fail to provide rational scaffolding but they could not do so. The fact that we hold a persistent plurality of conflicting desires is understandable only because desires fail to structure our options in the way that commitments of the will do.²⁶⁹ On this picture then, intentions allow for the possibility of *evaluative judgement* in a way that desires cannot. For example, Bratman says “[s]uppose that I know that I can stop at one of two bookstores after work, Keplers or Printer’s Inc. but not both. And suppose I find both options equally attractive. I judge all out that stopping at Keplers would be just as desirable as any act of stopping at Printer’s Inc., given my beliefs”. How can I thus motivate myself to go to one store rather than the other? The answer is that I can simply form an intention to do one rather than the other without changing my evaluative beliefs, just as a real Buridan's ass would be able to avoid starving by forming an intention to go towards one bale of hay rather than the other, even though it has no evaluative preference between the two.²⁷⁰

So the rigidity with which the normative force of intentions holds, i.e. the fact that it is constraining, allows intentions to play a special role in our practical reasoning; I will call this the *rigidity of rational scaffolding*. As I have set out above, this rigidity is characterised as imposing two rules: means-ends coherence and intention consistency. This talk of rigidity should not be thought of as masking the fact that we can, of course, change our minds. That is, we can rescind our personal intentions and their rigid hold over us is thus abolished.²⁷¹ Bratman does not think that if I form an intention that this intention becomes an alien force that rules over me deciding what I should do; “[f]uture directed

269 It is for this reason that Bratman rejects the belief-desire model of intentions, such as that held by Davidson. See Davidson, 2001b for his account of the belief-desire model, and see Bratman, 1999e for Bratman's rebuttal of it.

270 According to Bratman, Buridan's ass type cases are common in the lives of rational agents like us; for example, “Just consider choosing one of the many boxes of Cheerios from the supermarket shelves.” (Bratman, 1999f, p.11) As Bratman sees it, in Buridan type cases the beliefs and desires of the agent under-determine the choice that the agent has to make. However, we are still able not only to make a choice but to act in the fully agentic sense (Bratman, 1999f, p.11 & p.20). Once I have an intention for some particular action, this becomes directly relevant to the rationality of my future plans. So in the Buridan’s ass case, if that ass had chosen the left pile of hay then moving towards that pile becomes more rational than moving towards the other – even though the ass’s beliefs and desires still do not give it reason to go one way rather than the other.

271 For Gilbert, this of course is the key difference between the plural will and one's individual will, i.e. that we are in a position to unilaterally change our individual wills. However, each individual is not in a position to unilaterally change the plural will.

plans”, he notes, “... are after all revocable: they don’t control one’s future conduct by way of some mysterious action at a distance; and many times in the face of new and relevant information we recognise that it would be folly to stick rigidly with our prior intentions.”²⁷² So, to be clear, when I speak of rigidity of rational scaffolding, I am referring to the rigidity of the force of intentions *when held*. The question of whether we must stick rigidly with our prior intentions (i.e. not change our minds) is a separate one.²⁷³

What the introduction of Intention-Free-Jane shows is that we can carve up the category of intentional action more finely than I have been doing so far. In the first chapter, I tried to clarify the notion of intentionality with the distinction between the completely non-intentional movement of windswept sand and the clearly fully intentional act of climbing to the top of Everest. I said that, at a most basic level, the difference between the two is that action is purposeful. However, Intention-Free-Jane’s actions are intentional in a minimal sense that distinguishes them from the movement of wind swept sand, and yet they are not agentic in the full sense in which normal human adults’ actions are. Bratman sees such a point and presses the idea that unlike many creatures, “... we – normal adult human agents in a modern world – are not merely purposive agents”,²⁷⁴ rather we make intentions that we see as providing rational scaffolding. In this sense, the agency of us (as in we actual adult humans) is *planning agency*.

5a.2 – Why practical utility is not enough

Imagine that we accepted the story of the role of intentions above; would this be enough to get us to the bottom of the normative authority of intentions? It certainly does explain why it is reasonable for a sensible agent to let themselves be guided by their intentions. However, there is a problem. The conclusion, that an agent *must* stick to these rules in *every particular case*, does not follow from the premise, that being constrained by intentions is useful for achieving our ends. We can see this by returning to an examination

272 Bratman, 1999, p.2.

273 Recall that Bratman does think that there is a pressure towards reasonable stability in not changing our minds. I discussed this back in Chapter Two, as part of his attempt to provide an individualistic explanation of why we face normative pressure to not renege on playing our parts in collective actions.

274 Bratman, 1999a, p.5.

of Anarchist-Jane: suppose that she reads the above argument and accepts it; she says that she can see that it is useful to be constrained by intentions. *However*, she then goes on to say that, all the same, she would rather not be. Maybe because she has an underlying attachment to the power of spontaneity, an attachment that she feels outweighs practical utility. We want to be able to say to Anarchist-Jane that she has missed something. We want to be able to say that *she cannot just choose to ignore the force of her own intentions* in this way. Unfortunately, the truth of the instrumental utility of rational scaffolding does not give us enough ammunition to defeat Anarchist-Jane's intransigence.

That there is something more to the normative authority of intentions – than the general utility of being guided by them – can be seen in cases where the particular demands of the intentions in question do not match that which the agent has most objective reason to do. Take the example of Jane and her intention to go fishing. I noted in the previous chapter that believing fishing to be morally unacceptable would not in any way make Jane's intention any less a real intention (or as I spoke of it there, any less a real commitment of her will). Likewise, we can say that I might intend to push over an old lady (which would be morally reprehensible), or I might intend to cut myself (which would have a negative effect on my health). In all these cases, the fact that, objectively speaking, we ought not to positively value the intended ends does not appear to change the fact that if I intend them and yet fail to be constrained by those intentions, then (in some sense) I am in error.²⁷⁵ Further, we might even question the idea that being guided by the rational scaffolding of one's own intentions is always in one's general interest. Consider the following: suppose that Anarchist-Jane's defiance of the rules of practical reasoning is not based on mere libertarian wilfulness, but rather that she believes, *correctly*, of herself that she is prone to make very bad decisions. Thus, she supposes that her general utility will not be adversely affected by ignoring these rules. It seems not beyond imagination to suppose that not following the rules and thus failing to effectively act might actually be instrumentally

²⁷⁵ While it seems clear to me that the idea that evil intentions constrain us fits with our every experience, those coming to action theory via a concern with the foundations of ethical action, often find it problematic for us to be normatively constrained to commit bad ends. Bratman admits that in the past he tried to avoid the conclusion that we can have such constraints, but he now admits that they do exist as there is an intrinsic reason for self-governance even where “... self governance involves volitionally necessary bad ends” (Bratman, 2009a, p. 443 - particularly footnote. 75).

advantageous for her. The important thing, however, is that even if this is the case it still seems that Anarchist-Jane is doing something erroneous by ignoring the force of her own intentions over her.

The problem is that, it appears that we are rationally free to act in ways that are contrary to acting in the manner that would be of the greatest utility. Imagine that one of our two walkers has been up all night writing up a plan – based on extensive research – about the best route to take to the top of Scafell given their walking experience, the time of year, the weather conditions etc. The fact that *objectively* this plan describes the best route does not appear to bind the other walker to it. In fact, it does not even seem to bind the walker who drew up the plans to follow them. Both walkers are at liberty to exercise their personal autonomy and ignore the plan. They are not bound by it, even though it would be of great practical utility to follow it. In contrast, not feeling oneself to be bound by the plans that one commits to is not like this – it is wrong.²⁷⁶

What is missing from the above is, as Bratman puts it, the fact that intentions appear to have a “... *noninstrumental normative significance* in the particular case, a significance that is distinctive in the sense that it is not merely a matter of the promotion of your particular ends”.²⁷⁷ Using the terminology developed by John Broome, we can say that intentions are *strict normative relations*. That is, they *require* that those who have them must act appropriately. In contrast, Anarchist-Jane appears to wrongly take them to be *slack normative relations*, that is relations that *merely recommend* that those who have them act appropriately.²⁷⁸ In this way, the normative force of the rational scaffolding provided by intentions is universal (for the agent in question); it applies in each particular case.

Given all of this, we will want to be able to say that, by failing to see what her intentions require of her, Anarchist-Jane is not being rational. However, her irrationality

276 Recall (from footnote 254 above) that there are two senses to the term 'plan'. Bratman (1999f, p.29) distinguishes between having a plan as to how to cook lamb and planning to cook lamb. It is in the former sense, i.e. not the committed sense, that I am thinking of the walker having a plan as to the best way up the mountain.

277 Bratman, 2009a, p. 418 [*emphasis mine*].

278 Broome, 1999, p.409.

does not appear to be of the standard theoretical sort; that is, it is not a matter of her holding contradictory beliefs. We cannot, for instance, locate her irrationality in her holding of both the belief that she will do as she intends, coupled with the belief that she will not. This is because, though it would seem odd indeed if she believes that she would definitely not carry out her intentions, it is not a requirement that she believe that she *will* carry them out. For example, it is possible that she may intend to go fishing but yet not believe that she will certainly do so. All that she is required to believe is that fishing is not impossible, that there is some chance that she will do it. As Bratman has rightly shown, “... there need be no irrationality in intending to *A* and yet still not believing that one will.”²⁷⁹ Anarchist-Jane could, for example, be unsure whether she can reach the location of a suitable site for fishing and thus be aware that she may fail to fish, even though she is confident enough of the possibility to give it a go and thus to intend to do it (though, of course, being Anarchist-Jane she does not see her intention as binding upon her). “[T]he demands of theoretical rationality”, as Bratman puts it, “do not strictly speaking engage intentions, they only engage associated beliefs”.²⁸⁰ In this way Anarchist-Jane's irrationality is a matter of her incorrectly grasping what she ought to *do* rather than a matter of incorrectly grasping what she ought to *believe* – and, in this sense, it is *practical irrationality* rather than theoretical irrationality.²⁸¹

279 Bratman, 2009a, p. 38.

280 Bratman, 2009a, p.427, footnote 50. The converse account is dubbed by Bratman the '*Cognitivist Position*'. He notes that holding it would require a commitment to a close connection between intentions and beliefs. Against such a position, Bratman claims that rather than requiring a flat out belief that we will certainly do that which we intend, to have an intention we must rather just see the possibility of our doing as we intend as something we take for granted. That is, it must be, in Bratman's terminology, something that is *accepted as part of the cognitive background for our deliberation* (2009a, p. 38). Partly what we accept will mirror what we believe. However, Bratman says that, “There is an important phenomenon of acceptance that is context relative in a way in which belief is not”. Thus, while we cannot reasonably “... at one and the same time believe that *p* relative to one context but not relative to another” we do, in planning for our future actions, reasonably “... accept that *p* relative to one context but not relative to another”. (1999f, p.27) For example, in making her plan to go fishing tomorrow, Jane's 70% confidence – that the shop, from which she needs to buy equipment will be open – may suffice. However, in making a plan to hold a fishing party for her friend Jim's 80th birthday party, which she knows Jim will consider to be a very important event, her 70% confidence may seem too low. In light of this, Bratman says that, to form an intention to *A* we do not need to flat out believe that we will *A*, however we must ‘bracket out’ any doubts, that is, we must make a (context relative) acceptance that *A* will be achieved (1999, p.32).

281 In distinguishing between theoretical irrationality and practical irrationality I am following Bratman.

5a.3 – Self-governance and there being somewhere where-/stand

In setting up the puzzle of intersubjectivity in the previous chapter, I introduced the notion that we see ourselves as necessarily personally autonomous.²⁸² Further, I discussed how the possibility of intersubjective constraint can be seen as troubling because it appears to violate this autonomy. I ended that chapter, however, by asking what is it about constraint by our own intentions that means that it is not problematic for personal autonomy? We can now ask a related question, which pushes at the same issue: what is it that means that Anarchist-Jane (who lives in the moment, paying no heed to her intentions) is not a legitimate representative of the most authentic autonomy possible? Below I will present an answer, from Michael Bratman's work, founded on the idea that *our being autonomous and our being constrained are intimately linked*, such that, *the reason for the latter derives from our reason for the former*.²⁸³ At first this might seem odd. The rational scaffolding provided by intentions constrains our practical deliberation. How can something that constrains us be intimately linked to our being autonomous? Is autonomy not freedom from constraint? Answering this question will take us to the heart of the ground of the normative authority of intentions. As we shall see, Bratman's perception of the situation mirrors Frankfurt's insight that doing whatever one is motivated to do "... misses entirely... the particular content of the quite different idea of an agent whose will is free".²⁸⁴ For one's actions to be realisations of one's will being free – or as I am calling it, one's personal autonomy – those actions must be governed by one's will. While Bratman's thoughts share something with Frankfurt's – both are concerned with what it is that allows us to count as governing *as the very agents that we are* – he departs from Frankfurt in what he thinks underlies the possibility of governing. Frankfurt thinks that counting as governing comes from having 'second order desires' (see below). In contrast, for Bratman, counting as governing comes from being constrained by the rational scaffolding of

282 I contrasted this sense of autonomy with that of environmental autonomy, which requires control of one's surrounding world. Only in the personal sense of autonomy, i.e. that of making our own minds up, is necessary for the kind of agency we are interested in here.

283 As Bratman puts it, "... our reason for conforming to these norms of practical rationality derives in part from our reason to govern our own lives." (Bratman, 2009a, p.412).

284 Frankfurt, 1971, p.14.

intentions. The normative authority of intentions thus ultimately resides in their place in the metaphysics of the agency of creatures like us.

Understanding Bratman's argument requires an appreciation of the complexity of human psychology. Our beliefs, desires, memories and other psychological states change as time passes. Even in any one particular moment we will hold a diverse array of mental states, some of which may even be contradictory (as with the example, above, of Jane's conflicting desires). However, even though our psychological content is heterogeneous, we consider it possible for our actions, which must be brought about by some of those psychological states, to be *governed by ourselves as unitary agents*. The general existence of the feeling that – ourselves – govern in this way can be seen in the peculiarity of examples where we do not. David Velleman gives the following example: Sigmund Freud is reported as noting that, whilst sitting down at his desk, he moved his hand in a remarkably clumsy way and knocked an ink pot to the floor, smashing it. Freud's explanation for this was that his knocking of the ink pot was unconsciously done so as to get rid of it after his sister had remarked on its ugliness.²⁸⁵ Velleman's claim is that, even though Freud's behaviour is motivated by an element of his psyche (his desire to get rid of the ugly ink pot), it is not something that he truly governed. In this sense, it is not agentive in what we might call *the strong sense* but is rather, in Velleman's terms, merely 'motivated activity'.²⁸⁶

If most of our actions, or at least a great deal, are to be understood as expressing strong rather than weak agency, *and* strong agency cannot be understood merely in terms of our acts being expressions of our psychological states, then it becomes necessary, as Velleman puts it, to find a way to locate the "... agent at work amid the workings of the mind".²⁸⁷ That is, to switch back to Bratman's terminology of governance, for the agent to

285 Velleman, 2000, p.2.

286 Velleman distinguishes the category 'ungoverned actions' thus; "... contains the things that one does rather than merely undergoes, but that one somehow fails to regulate in the manner that separates autonomous human action from merely motivated activity" (2000, p.4).

287 Velleman, 2000, p.131. One might worry that in introducing the idea that we cannot reduce action governed by us to action caused by any proper part of us, we are putting an anti-naturalist bar on any reductive understanding of human agency. However, for Velleman, the idea that 'the agent should be in control' is not an absolute bar on reduction. Rather, it just sets certain criteria for what constitutes a successful reduction. It demands reduction that captures the functioning of the agent. That this type of reduction is possible can be seen by examining some of those operations of a person that are not agentive.

govern is “... for the relevant attitudes that guide and control [their] thinking and action to have *authority to speak for the agent*”.²⁸⁸ In sum, they must have *agential authority*. Frankfurt's solution to the quandary of how the relevant attitudes that guide and control our behaviour get to have agential authority (with which Bratman agrees) is that these attitudes must comprise a standpoint that we can identify with.²⁸⁹

How, then, might the relevant attitudes come to comprise *a* standpoint that an agent can identify as *their* standpoint? Frankfurt looks to find attitudes of the agent that can be said to be those with which the agent most strongly identifies. He thinks that the relevant attitudes are a certain kind of desire, *second-order desires*. Second-order desires are not desires about things external to the agent's mental world, rather, they are *desires about desires*. In a sense, they are desires of the agent that they *be* a certain way. They are attempts, we might say, at governing ourselves. For example, Jane's (first-order) desire to go fishing might be coupled with a desire that she desires to fish more than she desires to sleep-in in bed. These second order desires are thought to represent a standpoint that the agent can identify with because they describe the kind of person they want to be. This proposal has its merits. It is certainly preferable to any account that simply identifies an agent's standpoint as the sum of all their psychic stew. This is because it allows us to understand examples, such as Freud's smashing of the inkpot, as not being fully agentive. On Frankfurt's account, such situations are explained by noting that though Freud does have a desire that the inkpot be smashed, he does not desire that he should so desire.

Bratman rejects Frankfurt's proposal that simply any second order desire can play the role of setting up an agent's standpoint. This is because, he claims that, there is nothing essential to the nature of second order desires that means that we cannot be equally as alienated from them as we can be from first order desires. If Freud can reject his desire to smash the inkpot as part of his authentic standpoint, why is it not also the case that he could reject a desire, to desire to smash the inkpot? Similarly if, say, Jane had a higher order

For example, Velleman notes that it is easy to understand a person as being a 'digerster of food' in virtue of the functioning of certain of their parts (i.e. their stomach and the chemicals within it) (Velleman, 2000, p.138).

288 Bratman, 2000 [*emphasis mine*].

289 Frankfurt, 1997.

desire, to not desire, to desire to be the kind of person who desires to fish more than to sleep, where would this leave her second order desire? Which one would count as her governing and why? The problem is, as Bratman poetically puts it, there appears to be no principled characteristic by which we can identify any desire, of whatever order, as being anything more than just “... one more desire ... in the psychic stew”.²⁹⁰

To be able to understand a psychological state as capable of being part of a standpoint that counts as our standpoint, we need a state that we cannot dissociate from; a state that is part of the core of *our being the agents that we are*. As Bratman notes, “... it is only if there is a place where you stand that *you* are governing in the corresponding domain, for in self-governance where you stand guides relevant thought and action”.²⁹¹ He thinks that, unlike desires of whatever order, intentions can play this role. This is because, they are for an agent “... attitudes whose role it is to support the temporal organisation of her agency by way of constituting and supporting Lockean ties characteristic of her temporal persistence”.²⁹² This is the nub of the importance of intentions and the rigidity they provide. Bratman's reason for thinking that intentions can do a better job than second order desires in setting up a standpoint, from which the agent can govern, is that intentions can provide the kind of links between disparate mental attitudes that create a *whole agent*. That is, they play what we might call a Lockean role in constituting our (particular kind of) personal identity. It is because of the role they play in structuring agency across time globally that they can be seen as having agentive authority in each particular instance.

How, though, do we move *from* our heterogeneous psychological content *to* having the potential to act from a standpoint that is authentically ours? Bratman's answer follows in the tradition of Lockeanes who seek to explain the unity of the self in terms of psychological continuities that bring about “... cross-temporal organization and integration of thought and action”.²⁹³ While Bratman acknowledges that we can be united by mental links of many types (such as memories and the like), he believes that intentions provide an

290 Bratman, 2000, p.37.

291 Bratman, 2009a, p.431.

292 Bratman, 2000, p.46.

293 Bratman, 2009a, p.430.

especially strong kind of bond; a bond that has the power to make a standpoint not only united, *but also governing*.

In a general sense, there are plenty of characteristics that could be thought of as unifying an individual; for example, it is possible to think of me as being linked to Joseph Kisolo-Ssonko from 2011 by virtue of having the same jacket, the same mother, or the same memories etc. No doubt some of these are better candidates for providing a robust account of my identity over time than others. However, in searching for agentive unity we are looking for something particular; we are looking not merely for something that links a person together as a singular object, but rather something that links them *as an agent*. We are clearly looking for links between mental attitudes, but we must have a stronger sense of 'link' than mere causal connection, because one's mental attitudes could be causally linked, without this necessarily meaning that, in terms of their content, they were linked in such a way that we could conceive of them as creating a coherent agent across time. Bratman acknowledges that we might conceivably argue that other mental attitudes, such as memories, could serve as bridges creating a unitary agent across time. However, memories, like desires, do not have the power to structure our future choices. The fact that I can remember wanting to go to the beach yesterday, does not tell me what I must do today. Intentions, on the other hand, carry with them *practical authority*. If I continue to hold an intention to go to the beach then this not only tells me what I must do now (e.g. find my towel), but also structures the choices I can make into the future (e.g. getting out of the house in time to catch the bus). Intentions are special for Bratman because they are "... authoritative policies" that are "... embedded in structures of planning agency."²⁹⁴ Intentions create not just any kind of identity over time, but *agentive identity over time*. Moreover, agentive identity over time in the sense of strong agency.

The special characteristic, which we want the unifying factor to have, is that it should be a mental state that in some special way relates to the agent being the agent that they are. Intentions meet the requirement of having this special characteristic because they not only merely unify an agent, but give them a standpoint from where their actions can be

²⁹⁴ Bratman, 2009a, p.430.

governed because of the way in which they structure future rational deliberation. If we did not see intentions as necessarily structuring our future reasoning in this way, then they would not be the kind of thing that has the power of speaking for us. Here, my earlier comparison, between an agent who merely desires to keep fit and one who intends to, is germane. If an agent has the desire to keep fit, but not the intention, then if she fails to keep fit, she, other things being equal, commits no error. We can have inconsistent desires without this devastating where we stand, because desires, by themselves, are not enough to determine where we stand. However, if one has an intention to keep fit, then this frames what one ought to do – we can say that *it frames one's standpoint*.

It is important to grasp that Bratman's claims about the metaphysics of agency are claims specifically about the kind of agency that creatures *like us* exhibit.²⁹⁵ Thus, what it takes for us to be autonomous, given the kinds of creatures we are, may be different from what it takes for other creatures to be autonomous. Hence, a creature that simply acts instinctively may be understood as acting autonomously merely in virtue of their actions reflecting their current whims. However, we are not creatures like this. For the practical reasons, given above, we are planning agents. The fact that we plan, however, is not just an additional bolt-on to the nature of our agency; rather, *it is part of what it is to be the agents that we are*. This is the context in which we should understand Frankfurt's claim that, an "... essential difference between persons and other creatures is to be found in the structure of a person's will."²⁹⁶ Though there might be some room for dispute as to whether this distinction clearly splits humans from all other animals,²⁹⁷ it is fairly clear that it is a correct description of us; our actions are not just isolated reactive responses to the external world. Rather, they are integrally part of our planned response to that world. They do not just

295 Creatures like us broadly being "... adult humans in a broadly modern world" (Bratman, 2000, p.35).

296 Frankfurt, 1971, p.7. Frankfurt speaks of agents that are merely purposeful as '*wantons*' rather than '*persons*'. Intention-Free-Jane would be a wanton under this description, rather than a person. We might object that the term 'person' might have other implications, for example the application of moral rights, that do not map this distinction. Intention-Free-Jane may still have the legal standing as a person even if she does not exhibit strong agency. We need not resolve this question here, however. Rather, all we need to say is that what Frankfurt takes to be a necessary condition for being a person is at least a necessary condition for being the kind of agents *we take ourselves to be*.

297 As in this thesis I am interested ultimately only in getting to grips with human collective actions, I can leave this question of the nature of the agency of other creatures to one side. Lest one object to this focus as species-ist, we can say that this investigation applies to agents like us, and merely leave the question of whether any other animals are like us in the relevant way, open.

have a purpose, their purpose is *governed by us as agents who form ongoing plans about the shapes of our practical lives*. Bratman says that “... in self-governance the agent directs and governs her practical thought and action”.²⁹⁸ For planning creatures like us, directing and governing requires that we are guided by the framework which is created by the rational scaffolding of our intentions. In comparison, the things that Intention-Free-Jane does *are* intentional; they are not like mere windswept sand. However, they lack something. Her life is lived moment to moment, it is purely reactive, and in this sense it is not governed by her as a planning agent with a unified standpoint over time.

If Freud were like Intention-Free-Jane, then, we might think that it would be wrong for him to dissociate himself from the smashing of the ink pot. As Intention-Free-Jane acts always solely on the basis of her immediate inclinations, the only way we can understand actions as being hers is to see them as arising from her mental stew. However, Freud – and the rest of us – are different from Intention-Free-Jane. We are not agents who merely react; rather, we are planning agents. In this way, having intentions is, for us, a matter of our seeing “...ourselves as agents who persist over time, who begin, develop and then complete temporally extended activities and projects.”²⁹⁹ We see ourselves as having the potential to act today, as the very same planning agent we were yesterday, but this possibility requires something in addition to the mere fact that behaviour comes about as a result of elements of the same psychic stew. It requires that we have a *where-I-stand* built from the rigid scaffolding of our intentions.

As we have seen above, intentions provide us with rational scaffolding within the landscape of our practical choices. The rigidity of the rational scaffolding provided by our intentions is comparable to scaffolding supporting the construction of a bridge. For, just as that scaffolding – once erected – constrains the bridge's developing shape, intentions – once formed – provide stable platforms from which we can construct our practical lives. This metaphor is apt, for stretched just a little further, it brings us to Bratman's key insight: just as scaffolding only facilitates the construction of a bridge if (while erected) it is rigid

298 Bratman, 2007, p.4.

299 Bratman, 2000, p.35.

enough to hold the relevant materials in place, our intentions only structure our practical lives because (while held) they are rigid enough to constrain our practical reasoning.³⁰⁰ The rigidity of the scaffolding allows us to have a standpoint from which we can act as the agents we are. This is the very nature of the strong agency.

In summary then, the explanation for the normative force of the rules that append intentions in every particular case, is that seeing these rules as obligatory is a necessary element in the metaphysics of self-governance, that is, in making oneself into a united agent capable of governing.

300 These considerations, at least partially, mirror Christine Korsgaard's discussion of the normative force of the rules we set ourselves as deriving from ability to unify an entity: for example, she says that "... according to Plato, the normative force of the constitution consists in the fact that it makes it possible for the city to function as a single unified agent" (2009, p.152). Though note that Korsgaard conceives of normative force as being necessarily moral in a way that I have rejected in this – and the last – chapter.

Chapter 5 – Part b – Constraint by the Collective Will

I have set out above how Michael Bratman's answer to the question of the origin of the normative authority of intentions is tied to his *planning theory of agency*. Or to be more precise, how it is tied to his planning theory of the kind of *strong agency* that *agents like us* have. I have been arguing that we can understand collective action as action that is governed by a plural subject (comprised of individuals pooling their wills). The argument I will now advance is that just as *I* need *my* intentions to rigidly constrain *me* in order that the rational scaffolding which this provides can create for *me* a where-*I*-stand from which *I* can govern, *we* need *our* collective intentions to rigidly constrain *us* in order that the rational scaffolding which this provides can create for *us* a where-*we*-stand from which *we* can govern together.

While this is not a claim that Margaret Gilbert makes, it does fit with her project. For, as Thomas Smith notes, she believes that “... joint commitment is a *plural analogue* of the sort of reflexive commitment that an individual agent performs when he unilaterally decides or intends to do something.”³⁰¹ As it is such an analogue, we should be able to take Bratman's account and simply plug collective intentions into the place occupied by individual intentions. In the first section of the second part of this chapter (5.b.1), I argue that such substitution is indeed fruitful. However, while it is possible to successfully ground the normative authority of collective intentions in this way, we shall see that a gap remains between the necessity of a where-*we*-stand *from a collective perspective* and its necessity *from the perspective of the individual members* of the collective. I will first try to bridge this gap with the notion of identity submersion (in 5.b.2) but dismiss this as not correctly satisfying the phenomenology of constraint. In light of this, I go on to develop (in 5.b.3) a transcendental argument that the process of engaging in collective action entangles one's own idea of one's individual standpoint with that of the standpoint of the collective of which one is a member, such that a failure to be able to understand the latter undermines the ability to be able to understand the former.

301 Smith, P.1127, 2007 [*emphasis mine*].

Before I set out on this task, it is worth noting that Bratman would not accept this use of his planning theory for (as I made clear in Chapter 2) he does not accept the notion of a plural agency. However, that is not to say that he does not see his own theory of collective action as connected to his theory of individual planning agency.³⁰² Bratman calls his project for understanding collective action a *constructivist project* because it “...begins with an underlying model of individual planning agency ... then seeks a conceptual and metaphysical bridge from such individual agency to modest forms of sociality.”³⁰³ The difference between Bratman's constructivist project and the constructivist project that I am arguing for here is that, while we both start from his planning theory of agency, he does not believe that the explanations will be *symmetrical* in the way I am proposing. That is, while he accepts that collective commitments of the will have specific structuring and coordinating roles, and that these generate characteristic social norms, he does not believe that these roles create the possibility of *plural agency* in the sense that their individual analogues create the possibility of *individual agency*.

In setting out his vision for his constructivist project, Bratman criticises Gilbert for introducing what he calls the “... non-reducible social concept of a 'joint commitment'”. In this light, he claims that in Gilbert's Plural Subject Theory, the concept of joint commitment is necessarily “... conceptually primitive”.³⁰⁴ This critique appears to contain the implicit claim that Gilbert's Plural Subject Theory negates the possibility of understanding the plural will using the same conceptual tools as used to understand the individual will. Hence, Bratman claims that *he can*, and – by implication – that *Gilbert cannot*, explain the nature of our obligations towards collective commitments of the will “... while staying within the conceptual and metaphysical resources of the planning theory.”³⁰⁵ In this section I take myself to be refuting this critique. My claim is that while Gilbert's

302 In a yet unpublished book to be titled *A Theory of Shared Agency (forthcoming)* Bratman promises to paint a more explicit picture of his vision of the continuity between individual planning agency and modest sociality. Elements of this project are sketched in an essay (mooted as a prospective chapter of said book) entitled “Group Agency” (2009b).

303 Bratman, 2009b, p.57.

304 Bratman, 2009b, p. 58.

305 Bratman, 2009b p.58.

notion of joint commitment *is* explicitly non-reducible to a sum of individual commitments, this *does not* mean that we must see it as unanalysable. Indeed, as I have set out above, I will be arguing that we can understand its construction in a way that is symmetrical to Bratman's understanding of the construction of the individual will. As Pierre Demeulenaere notes – in a comment on Bratman's "Group Agency" article – Bratman's "... thesis is to build up shared social norms on the basis of individual norms".³⁰⁶ Hence, Bratman speaks of his constructivism as conservative.³⁰⁷ Adopting this language, we can say, in contrast, that the constructivism I set out below is a *radical* one. It is radical in that, though using the same initial conceptual apparatus as Bratman, it aims to justify the *move beyond the framework of individual agency* – that is, the move beyond the idea that the only norms of agency are individual norms.

5b.1 – Constructing the Plural Agent

So what happens if we simply take Bratman's planning theory of agency and plug collective intentions into the place he reserves for individual intentions? Straightforwardly, this results in what we might call the *collective self-governance explanation*. According to this explanation, the normative authority of collective intentions is grounded in the necessity of rigid rational scaffolding in order for the plural subject to count as governing its actions.

The plausibility of this explanation requires underpinning by a number of suppositions. Firstly, that we think of collective acts as being performed by plural subjects. The argument that this is the case is one that I made in chapters Three and Four. Secondly, that collective intentions constrain rational deliberation for plural subjects, in the same way that individual intentions constrain rational deliberation for individuals. From which would follow that, the collective engages in a strong form of agency. This too has been covered to some extent previously,³⁰⁸ but I will briefly clarify the argument for it again below, presenting it in light of the discussions of the previous section of this chapter. Lastly, and most crucially, that collective intentions can play the same part in constructing a

306 Demeulenaere, 2009, p.61.

307 Bratman, *forthcoming*.

308 In particular, see Chapter 3, Section 3.3.

collective standpoint as individual intentions play in constructing an individual standpoint. Moreover, that, it is only if there is a place where *we* stand that *we* can govern our collective actions. This I shall explore below.

On the question of the rigidity of collective intentions, reflecting the discussion in Chapter Three, we can say that they give rigid structure to practical choices in situations of collective action, just as individual intentions structure practical choices in situations of individual action. This can be seen in the example of Gilbert's walkers; together they have the joint intention to walk to the top of the hill and it is as natural to suppose that, as a pair, they are constrained by this, as it is to suppose that each individually would be constrained by their own individual intention to walk to the top of the hill. To say this is just to acknowledge facts such as, if they collectively intend to walk to the top of the hill, then they cannot rightly also intend to abandon the walk halfway up. Nor can they rightly intend not to wear down the path to the top of the hill, if they know that walking to the top of the hill will require this. Further, their joint plan to walk to the top of the hill structures their joint practical life; for example, it settles the practical questions of what together they should do as they progress past different landmarks (i.e. keep going until they reach the hill's summit) and allowing them not to have to keep deliberating at each stop. This point is noted by Gilbert, who says that, “[just] as personal intentions are subject to demands for coherence and consistency, and so on, these shared intentions would appear to be subject to similar demands”, for example, “[i]f Zena and her friends share such an intention to bring peace to the world, they have reason, by virtue of that intention, to develop concordant sub-plans.”³⁰⁹

Does this mean that collective action is agentive in the sense that Bratman calls strong agency? If we accept the conceptual apparatus of Gilbert's Plural Subject Theory, then it turns out that it must be. This is because according to Plural Subject Theory, *collective intentionality is never purely reactive*, but rather it is always the result of joint commitment. It is, in this minimal sense, always planned. This can be seen in the fact that while Intention-Free-Jane was seen to be practically impossible (given the pressure to

309 Gilbert, 2008, p.511.

organise our lives), her collective analogue is not even conceptually possible. This is because it is not possible for a plural subject to act merely on its momentary inclinations as *there is nothing more to the practical perspective of the plural subject than the collection of the intentions which its members collectively commit to*.³¹⁰ Therefore, it does not make sense to think of collective action as being weakly agentive; rather, when it exists, it must always be strongly agentive. Collective agency is *collectively planned agency*.

The possibility of the strong agentive nature of collective acts can be seen – just as it could in the case of individual agency – by noting the contrast with situations where there is no planning agency. In this vein, imagine that our walking couple are setting out on their hill climbing adventure on a particularly hot day. Let us suppose that there are two ways to get to the start of the hill, one is short and quick and best facilitates starting the hill climb, the other is much longer and requires wading through a river. It seems that in such a situation, just as Freud's subconscious desire to smash his ink pot (discussed above) could lead him to do so, our walkers might have unexpressed individual desires to get wet in order to cool off, and these could lead them to together take the longer path without individually realising, or collectively expressing, their motives. While the walkers' collective act of ascending the hill is something that they can lay full authoritative claim to as being their action, something that they truly *jointly* govern, in contrast, this is not the case with their diverting along the path that takes them through the river. In terms of collective action, this is not collectively agentive *at all*, as it is rather driven by their *individual* aims to keep cool.³¹¹ So we can say that, there is a need for an authentic collective standpoint for agents like us to achieve collective action, just as there is need for an authentic individual standpoint for agents like us to achieve individual action. Indeed, the point here is even stronger than in the individual case for, given that we have accepted the legitimacy of Gilbert's claims about the necessity of joint commitment for collective action, we can see

310 On the organic thesis (which we rejected), where the collective is seen as an emergent organism, this would be possible. However, for collectives *of the sort we actually do have* – rather than the sci-fi Borg discussed in last chapter – this is not true.

311 It may be that the walking past the stream differs from the smashing of the ink pot in that from an individual perspective it may be strongly agentive, i.e. individually they may intend to go via the stream. However, even if this were the case on the individual level, the important point, however, is that it is not collectively so.

that the act of the walkers is only a collective act of the type specified *at all* in the circumstances of strong agency.

So, it seems that we do see rational deliberation for us as plural subjects as being constrained by collective intentions. Moreover, it seems that in doing this, collective intentions can play the same part in constructing a collective standpoint as individual intentions play in constructing an individual standpoint. Given this, we can say then that all the members of a collective together failing to see their collective intentions as normatively constraining would undermine their ability to act from a standpoint that can count as *their* own standpoint. In the case of our walkers, this means that if the pair do not see their collective intentions in general, and their joint plan to climb to the top of the hill in particular, as normatively constraining them, they will fail to form a collective agent capable of the governed intentional act of collectively walking to the top of the hill.

A wrong-headed objection to this line of reasoning would be to say that just because a collective fails to be an agent, this does not necessarily mean that it will fail to be a collective *per se*. There are other ways in which we can identify the collective as a unitary entity, for example, that its members have common characteristics or certain physical relationships between each other and such like. Alternatively, we might consider it to be a unitary entity because of its legal or conventional status as one. This is correct; however, these facts do not blunt the force of the argument above, just as the following similar point does not blunt the argument regarding individual intentions. One could still identify an individual who failed to be bound by their own intentions as being a unified thing *of some sort*; this could be achieved by reference to their physical properties, to their social position, or their legal standing and so on. However, whilst being a united thing of some sort, this person would fail to be an agent united in a way that made her capable of governing her own actions. As I stated in the last section, what we require is not simply any kind of identity over time, but *agentive identity over time*. This, it seems to me, is akin to Locke's claim that rather than search for any old sameness, we must specifically search for "...the

sameness of a rational being.”³¹² The same holds true for the collective; it is not what it takes to be a united object of any sort that is relevant, but rather what it takes to be united as a possible collective agent of actions.

Assuming that the above argument is sound, i.e. that there really is a correspondence between that which explains the unity of the individual standpoint and that which explains the unity of the collective standpoint, then, just as individual intentions have force over the individual, collective intentions have force over collectives. Unfortunately, this does not yet quite get us where we need to be in order to understand the foundations of the force of normative constraint in situations of discord between individual wills and collective wills. This is because it is not yet completely clear, as I will set out below, *what is at stake for the individual in the collective's being understandable as a planning agent*.

5b.2 – Identification

The collective self-governance explanation presented so far allows us to understand the necessity of collective intentions being seen as constraining for the plural agent to exist as a plural agent. However, there is a gap between this and explaining why each individual member of the plural subject must see these intentions as constraining on their *personal* practical rationalities. This gap exists because each member is not, on their own, identical with the plural subject and thus each member can ask the question: “Why should I care about the existence of the plural agent?” The intrinsic reason that each individual has for their own existence as an agent cannot be straightforwardly transposed onto the collective case. In the case of the individual, the agent cannot abandon their own standpoint as they have no other standpoint to fall back on – *one's existence as a planning agent cannot be separated from one's existence as the very individual that one takes oneself to be*. By contrast, in the case of an agent's attachment to the collective perspective, we might wonder why that

312 Locke, 1748. Though we might quibble over whether being rational requires being a planning agent. For example, we might suppose that Intention-Free-Jane could be rational if her actions are appropriate responses to her momentary desires even though she fails to plan across time. However, it seems likely that at least sometimes when we think of what it takes to be a fully rational being we are thinking in terms of the need for something like what has here been called planning agency.

agent cannot simply abandon the *standpoint of the we* and retreat back to the *standpoint of the I*?

We are justified at this juncture in making the following conditional claim: if an individual agent wants to see their actions as part of a collective action, then they *must* see the intentions of that collective as normatively constraining on the participating agents together. This is because it would not make sense for an agent to see their actions as part of a collective action if they fail to be able to rightly see *that collective as an agent capable of action*. On this account, then, collective intentions have the power to normatively constrain because the individual sees themselves as participating in the action of a plural subject. This requires that they see their collective as the kind of thing that we can understand as a plural subject; the kind of thing that can act with strong agency. This, in turn, requires that collective intentions are seen as having the power to constrain such that they can provide the structure necessary for the plural subject to be able to govern its actions.

This is not the end of the task however, as we must still ask: what is the relation between the plural agent, which the individual sees themselves as part of, being constrained and that individual seeing themselves as constrained. In order to answer this question I will turn to the notion of *identification* with the collective. As well as asking what it means for the agent to identify with the collective, I will examine how it is that, in situations of discord, such identification can become less than voluntary.

Firstly, we must clarify exactly what is meant by 'identifying with the collective'. We sometimes talk of identifying with others in a very loose sense, where what we mean is that we see ourselves as similar to them in some relevant sense. A recent example of such usage, which generated heated debate amongst political commentators, involved members of a British anti-government campaign group claiming to identify their movement with that of the successful Egyptian campaign to oust the Egyptian President Hosni Mubarak.³¹³ The dispute amongst the commentators was in regard to whether or not the two movements shared significant relevant characteristics for such an identification to be

313 See for example Jenkins (2011) & Gopal (2011).

legitimate. However, even if we assume that they did, this does not seem to be the kind of identification that could help with the current issue. We might well imagine that the British protesters' attempt to identify themselves with the Egyptian protesters was aimed at forging a sense of solidarity and a commonality of aim between the two groups. However, it seems obviously far too strong to suggest that this would mean that the British campaign group should see themselves as constrained by the present intentions of the Egyptian campaigners. For example, the Egyptian campaigners might be said to have held the collective intention to occupy Tahrir Square, in Cairo; if the British protesters were constrained by this intention then they would have to hop on a plane and fly to Egypt! Amongst other things, this would seem to problematically conflict with their own collective plan to occupy Trafalgar Square, in London.

If the protesters' seeing themselves as importantly similar to the others is not enough, then how might we strengthen the identification requirement? Well, sometimes we talk of people identifying with others in a much stronger sense than that above. Sometimes we talk about individuals identifying with others in such a strong sense that they see themselves as sharing a *singular identity* with those others.³¹⁴ Cases of very highly organised, strict and centralised religious and political organisations, what we might call 'cults', appear to be illustrations of this type of phenomenon. In such cases, individual members of the collectives in question might feel that they have no individual standpoint of their own, rather, they identify the standpoint of the group as *their* perspective. The submersion of individual identity beneath the identity of the group to this extent might strike us as politically worrying, but it may well be functionally useful. Imagine that one of the individual members of the British anti-government protesters in the example above saw themselves as having no agentive identity beyond that of the group, or at least no agentive identity in the context of the time in which they were engaging in the collective action of occupying Trafalgar Square. If this were the case, then – free from considerations about personal safety and personal desires and the like – they would be focused solely on the

³¹⁴ While these two senses of identification are no doubt linked, it is not entirely clear how identification in the widely used sense of sharing significant relevant characteristics engages with the identity in the philosophical sense of being a singular object. To be clear, it is the possibility of this relational philosophical sense of identity that is important for the discussion that follows.

'higher' goal of maintaining their collective capture of the Square. Thus they might well be more likely to achieve their group goal.

If we understand identification as submersion of singular identity, then this not only tells us why the individual cannot help but see themselves as part of the plural subject, it also tells us why this identification is not experienced as conditional. For if the agent truly has submerged themselves, then they will have no individual standpoint left to retreat back to. If they understand themselves only as a group member, then constraint by collective commitments will be inescapable. On this account then, collective intentions have the power to normatively constrain individuals because the individual fully identifies themselves with the plural subject and thus they have no standpoint of their own. This requires that they see the collective as the kind of thing that can act as an agent. This, in turn, requires that collective intentions have the power to normatively constrain, for their doing so provides the rational scaffolding necessary to create a standpoint from which the collective can govern.

Does such submersion happen? Well it may well represent a utopian ideal for some collectives, but it seems doubtful whether it is ever really achieved, even in the case of cults. The situation which seems to fit with such a model best is that of the fictional Borg discussed in the previous chapter. However, while individuals assimilated into the Borg collective can (in their fictional reality) fully submerge their personal identities, this possibility seems too extreme to represent the formation of collectives in the real world. In most situations, the notion of completely giving over of one's identity is obviously far too strong. This can be clearly seen if we return to Gilbert's canonical example of the two walkers. The extreme nature of requiring full identity submersion is tempered slightly if we bracket the submersion of identity to the activity taking place at that time, i.e. to the context of the walk, rather than extending it to such things as the breakfast they ate before they set off and the drive home they will complete after the walk, both of which they may do very much as individuals. However, even with the caveat of this bracketing, it seems very odd to think that either walker completely subsumes their agentive identities underneath that of the collective which they form. As much as we want to say that there

really is a collective 'we' that is endeavouring to reach the top of the hill, we do not want to say that there no longer exists two individual walkers at the same time.

So, identification is not a matter of losing one's own individual identity; rather, it appears to be a matter of being both part of the plural subject but also still an autonomous individual. However, if this is the case then can collective intentions be strict normative relations in Broome's sense? In one sense, yes. For, if an agent wants to see their action as part of a collective action, then this does not merely *recommend* that they see the collective intentions as normatively constraining; rather *it requires it*. However, it might be complained that the conditionality of this argument nullifies its force. If identification with the collective is a choice, then what stops us from just not identifying? This would violate Gilbert's obligation and permission criteria; namely that an individual faces normative pressure to act in line with the intentions of the plural subject (that they are a member of), even where this conflicts with what their own personal intentions would have them do. Keith Graham appears to mark this point when he claims that “[p]recisely because we *are* individuals whose existence is not exhausted in the social relations we participate in and the groups to which we belong, *questions can arise about whether to identify with or dissociate from collective agencies of which we are members*”.³¹⁵ We can see this in the example I set up as a canonical example of the phenomenology of constraint in collective action: the Leicester Tigers pushing their coach up a hill. The question that I have been seeking to answer is: “What is it that explains the fact that John (a member of the team) is directly normatively constrained by the collective intention of the team?” We have above an explanation that will only work if John conceives of himself only as a team member. But this is not the case; not only does he also see himself as an individual, he is additionally a member of other collectives, for example, his family. So, John has other alternatives to identifying with the team. What, then, would make it wrong for him to say, “Sod the team's existence, I'm off for a beer”? The answer, which I will be setting out in the next section, is that part of John's sense of self will be *entangled* with the collective and thus conceptually require its existence. That is, I will claim that John's ability to understand himself as having a coherent where-*I*-stand and thus to be the kind of agent he in fact is, is

315 Graham, 2007, p.8 [*emphasis mine*].

dependent on his ability to understand the plural subject he is part of as having a coherent where-*we*-stand.

5b.3 – Inauthenticity and the transcendental argument for the inescapability of the plural subject

So, what is to stop John saying “Sod the teams existence”? Well, being able to understand ourselves as parts of plural subjects is something that has many pragmatic advantages. It allows us to navigate a social world which is populated by a vast number of different individuals without having to consider each individual as such; we can engage, rather, with collectives. Moreover, being able to consider practical questions from a collective perspective also seems to solve problems of the rationality of co-operation in 'prisoners' dilemma' type cases, as co-operation is unproblematically the most rational choice from the collective standpoint.³¹⁶ As Philip Pettit says, where “... I see myself as the representative of a group, charged with doing as well as possible by its interests [I] might use this representative identity to get out of game-theoretic predicaments”.³¹⁷

Clearly, not being able to understand the world in terms of being a part of true collective actions would have many disadvantages. This gives us a forceful pragmatic reason to be able to understand the collectives we are part of as potential agents of collective actions. However, if this reason is only general and pragmatic, then why is it not easily trumped by the other pragmatic demands of an agent's own standpoint? As we have already seen, the move from general rules to particular norms is difficult. Whilst we might have a *general* reason to maintain the identity of collective agents in this way, this does not necessarily give us a reason *in each particular case*. So, it is clear that John has a general reason to want the rugby team to exist as an agent that can perform joint actions. For example, he wants them to win the cup as a team. However, in this particular case, the team's existence is causing him to be compelled to push a heavy bus and thus, everything

316 For example, see Hollis (1998) for an elucidation of what he calls the 'team work' solution to problems of the rationality of cooperation (particularly pp.137 – 142). I discuss this further in Chapter One, section 1.3.

317 Pettit, 2007, p.33.

considered, he has no reason, founded upon his own individual standpoint, to see it as existing.

That we cannot locate the strict reason, for an agent to have to understand the collective they are a part of as an agent, in the general utility of doing so, does not mean that pragmatic questions never inform an individual's social interaction with others. There are times where an individual's interaction with others is driven purely by the pragmatic perspective of their own standpoint. Take Bratman's own example of interaction that falls short of collective action: suppose that I am a gangster and that, "I intended that we go to New York together as a result of my kidnapping you and forcing you to join me".³¹⁸ Imagine that I gave you a choice, either you act as if we are travelling together or I have your parents murdered. You would have a strong pragmatic reason to 'go along' with my scheme and take part in what, to onlookers, may appear to be a collective act. The point, however, is that while this may appear to be authentic collective action, in fact it is not. In a similar way, John the rugby player may attempt to be an *inauthentic team member*. That is, he may attempt never to *really* identify with the team at all and rather see himself as merely going along with the notion that they are pushing the bus together for purely individualist reasons (say, that he wants the others to like him).³¹⁹ This would be an interesting kind of social interaction but it would not be authentic collective action. Such interacting individuals could rightly consider themselves to be a unit of sorts; united by their physical proximity, their interdependence or perhaps their mutual goal. However, they would not be bound *as a planning agent* – i.e. as a plural subject – which is precisely why they would not count as performing actual collective acts. Both of us would be free to act contrary to the mutual goal (of travelling on the plane together) without being concerned about this destroying the potential of our collective to act, although of course you would rightly be very concerned about maintaining the fiction of our collective act so that I did not carry out my threat.

318 Bratman, 1999b, p.100.

319 The reason I say John may *attempt to be* inauthentic in this way, rather than saying more straightforwardly that John may *be* inauthentic, will become clear in the following chapter as there I will argue that the reality of social life makes such inauthenticity difficult i.e. that we cannot help, as I shall put it, but become entangled in collective wills.

Most social situations, however, are not like that of being kidnapped or trying to be an inauthentic team member. The former claim is quite obvious; the latter is perhaps more contentious, but I take it that the phenomenology described in Chapter One of this thesis attests to its truth. That is, in more typical social situations we do not conceive of ourselves as merely pragmatically interacting with others. Rather, we see ourselves as involved in *authentically collective acts*. Further, we do not stand outside of each instance of social interaction and at each moment choose whether to engage in collective action – rather, we find ourselves *already bound up in* many ongoing social projects. This can be seen clearly in the example I have been using throughout the thesis of Hillary and Tenzing summiting Everest: as Philip Ebert and Simon Robertson note, “When asked who reached the summit of Everest first, Hillary and Tenzing have always insisted that they climbed it together and that there is therefore little point to that question – after all, *they* did.”³²⁰ In Chapter One, I noted that this showed that the climbers experience the intentionality of their action as essentially collective. Here, we can now go further and say that this shows that because they have understood their endeavour as collective they cannot now, *post-hoc*, reconceptualise it as an individual feat. To do so would not just be to enter into a situation inauthentically, as John the inauthentic rugby player above tries to do. Rather, it would be to be *inauthentic in understanding what one had already (authentically) done*. Graham seems to be noting a similar point in saying that “[s]ome of the things that people do *gain their significance* from being part of some collective action”.³²¹ It is this concept of being bound up, through the social nature our actual lives, in the collective will that I will refer to as *entanglement*.

The argument here, then, is that we can assess the normative force of attachment to the possibility of the plural subject in two parts:

Firstly, prior to an individual’s authentically engaging in collective action. At this point, given the general utility of collective action, each of us has a general reason to engage in it. This gives us a reason to see the collective as an agent, but it is a reason with a slack

320 Ebert & Robertson, 2010, p.102.

321 Graham, 2007, p.60 [*emphasis mine*].

normative force. Slack in that it merely *recommends* that we see our collective as potentially a plural subject.

Secondly, after our authentic engagement in collective action. At this point, given the fact that we need to be able to understand the collective as a plural subject in order to be able to fully make sense of our contributory actions, we have a reason that holds sway in the particular instance. This gives us a reason to see the collective as an agent, and it is a reason with a strict normative force. Strict in that it *requires* that we see our collective as potentially a plural subject.

We set up our actions to be contributions to social agency because it is beneficial to do so. However, once we have done so, we can only continue to understand our contributory action as the kind of thing we set it out to be if we are able to see the collective as an agent capable of governing our collective actions, and because this requires the intentions of the collective to constrain, we *must* see them as doing so. Post-hoc reconceptualising our contributions is logically possible, but it is a kind of inauthenticity. It is a kind of inauthenticity in that it requires self-consciously misleading ourselves about how, at the moment of our actions, we set them out to be. Engaging in such inauthenticity undermines our own agentic identity as singular agents, and, as set out above, we have an intrinsic reason to value the existence of our own agentic identity. It does so because if we fail to see ourselves as bound by our former intentions, in this case our intentions that our actions be authentic contributions to collective acts, we fail to then be able to have a where-I-stand from which we can govern.

This account has what we might call a two-part *transcendental structure*.³²² For it starts with peoples' experience, and says that our social lives are such that we feel ourselves to be part of collective actions. It then presents, (firstly) the existence of a collective capable of governing its own actions as necessary for us to have this experience, and (secondly) the constraint of individuals by collective intentions as conceptually necessary for the existence

³²² I understand the term 'transcendental' here in a general, rather than any specifically Kantian, sense (see Stern, 1999).

of the collective as an agent. From this, it concludes that collective intentions must constrain individuals.

To return to our hill walkers, according to the account given above, they have a general pragmatic reason to start walking together, given the utility of such collective actions. This pragmatic reason does not force them to engage in collective action. They may, for instance, feel that walking up the hill as individuals, perhaps individuals under contract to help each other if needed, is just as individually beneficial as fully-fledged collective action. If they do so, they will not be constrained by a collective intention – although they may face other normative pressures, such as fulfilling their contracts to each other. However, if they do decide to engage in collective action, then, once they are doing so, their understanding of their own contributory action will be bound up with the existence of the collective act. This will require them to continue to see the collective as capable of governing its action. This, in turn, means that they must see themselves as having to abide by the norms that attend the collective intention to walk to the top of the hill, because failure to do so will result in failure for them to be able to understand their contributive behaviour as such, and thus failure to fully understand what they are doing.

We can conclude, then, that an individual's reason for seeing themselves as constrained by the intentions of a collective of which they are a member is both similar to and different from that which they have for being bound by their own intentions. It is similar in that it is a matter of securing a unitary standpoint which can be the authentic agent of actions. However, it is different because the securing of this standpoint is not an a priori necessity for the individual. Rather, it becomes a necessity only after the fact of social interaction. Understanding the nature of collective action in this way means that we must not think of becoming bound by collective intentions as a simple case of voluntarily pooling our wills (as it is on Gilbert's account). Instead, we must understand ourselves as bound by collective intentions because we become entangled in a collective will through the actual process of engaging in social life. I further explore what this means, with regard to how we understand joint commitment, in the next chapter.

5.c.1 – Chapter summary

In the first part of this chapter, I set out Bratman's planning theory of agency and established that the rigid structure of the constraint our individual intentions put upon us is necessary for us to have a where-I-stand from which, as individuals, we can govern our own actions. In the second part of the chapter, I applied this to collective intentions, finding that we can similarly say that the rigid structure of the constraint our collective intentions put upon us is necessary for us to have a where-*we*-stand from which together *we* can govern our collective actions. Modifying Bratman's account allows us to understand collectivist accounts of collective intention, such as Gilbert's, as rightly saying that collective intentions are projects of creating a plural will through constructing binding collective commitments. However, it forces us to acknowledge that these collective commitments will only appear obligatory to the individual to the extent that the individual sees the existence of the plural will itself as a required fact. This does not mean, though, that the individual can merely choose at any moment whether they will or will not see collective commitments as binding. This is because their understanding of their own actions is entangled, through the actual process of living a social life, with the agency of the collective.

Returning to the terminology of intersubjectivity, we can see that our lives are intersubjective because *understanding ourselves as subjects* (in the sense at stake) *is bound up with our sense of being entangled in plural subjects with other people*. Thus, the fact that we can rationally act directly upon intentions that are not wholly our own is tied to the fact that our sense of rational autonomy is far messier and more extended than individualists allow. Given their entanglement in various plural agencies, the socially situated autonomous individual can be recognised as having a choice: accept the normative constraint of the collective intention or abandon the possibility of collective self-governance. Unlike abandoning the possibility of individual self-governance, the latter option is not completely barred to the agent, for the agent can fall back onto their own individual agency. But this option is not without conceptual cost, for, given our contingent but actual social experiences, our own sense of subjectivity is bound up with our sense of

belonging to, and acting as part of, a collective. Rejecting the existence of the plural subject will come at a cost for the individual and in the next chapter I will argue that this is experienced by them as what I shall call *normative pain*, the felt strength of which will be determined by the extent to which they are entangled. Thus, as we shall see clearly in what follows, we do not experience the choice as a plain choice, but rather we experience it as a dilemma, as a situation in which we feel torn.

Chapter 6 – Modified Plural Subjects

Let us return to the reoccurring motif of our hillwalking pair. Previously, I considered their walk in isolation, and their relationship only in terms of their particular act of walking to the top of Scafell Pike. Doing so allowed me to hone down on certain key aspects without the distraction of too much informational noise. However, such a strategy does risk missing something of the complexity of real social life. In this chapter, I will look at how acknowledging such complexity informs the way we should think about collective action. In the example of our walkers, allowing the complexities of their lives to come into view makes an important fact immediately obvious: they are unlikely to have met for the first time as they stood at the bottom of the hill. Rather, they will have a *social history*. They may have climbed hundreds of hills before as part of a hill walking club; they might work together, or have been friends for a long time; they may be members of the same political party; they may even be married. I will argue that these kinds of facts, regarding the histories of such collectives, taken together, constitute their members having lived (what I will call) a *social agentive life*.

I will approach the complexity of social life through the prism of the *Modified Plural Subject Theory* (PST*), as developed in the last chapter by applying Bratman's planning theory of agency to Gilbert's Plural Subject Theory. PST* replaces Gilbert's idea of *pooling of wills* with the notion I expressed as *wills becoming entangled*. My argument will be that, by understanding *entanglement as a messy elongated product of real and complex social life*, PST* can better capture the *actual* phenomenology of collective action.

The chapter will be split into two sections. In the first, I will expand upon the nature of PST*, and the difference between it and Gilbert's PST. I will further tease out the nature of entanglement and ask what this tells us about how we experience normative constraint. In particular, I will relate a notion of the *variability of lived agentive life* to the *variability of the normative pain of going against the collective will*. My claim will be that if, for example, we take our two walkers above, the felt cost of going against the collective will

shall vary in line with the extent to which they are entangled in that collective. Thus, while the strict obligation to their collective decision to walk to the top of the hill will exist independently of their wider history, an aspect of their experience will feel more normatively problematic for them if they rebel when they are more entangled (e.g. if they are married), than if they are less (e.g. if this is their sole joint venture). I shall understand this as *variance in the normative pain* that tracks *variance in entanglement*.

In the second section, I will look at two test cases which, I will argue, show the phenomenology of actual social life to fit better with PST* than with Gilbert's PST. The examples I will explore will be, firstly, that of being in a romantic couple, and secondly, that of political obligation. In setting out these examples, I aim not to make substantive claims about the exact nature of love or political fidelity; rather, I merely aim to set out how our general experience of these phenomena fits with the conception of collective action I am arguing for.³²³ I will be arguing that, with its notion of entanglement, PST* better captures how it feels to become part of the plural will (i.e. that being united need not be fully voluntary, but can instead involve merely living a social agentive life) and, further, the variance in the feeling of normative pain of rebelling against the plural will (i.e. that such variance tracks variation in the level of entanglement.) By setting out these examples, I will also make concrete what I take actual lived social lives to involve, in particular, how I take the experience of normative pain to be expressed. These two points mark the distinctiveness of my variant of Plural Subject Theory from Gilbert's.

6.1 – Pooling -vs- entanglement

According to planning theory, individual intentions, or as Gilbert calls them individual commitments of the will, structure the practical lives of individual subjects. In the previous chapter, I presented the case that our being bound by our intentions is rooted in their allowing us to understand ourselves as the very governing agents that we take

³²³ I aim more directly to give an account of loving union in Kisolo-Ssonko (2012) and a further project will be to construct a theory of political obligation as founding on entanglement as described by PST*, in the same way that Gilbert sets out to give such an account of political obligation as founded on pooling of wills as described by her theory (see Gilbert, 2006).

ourselves to be. Further, I argued that this also holds for collective cases, i.e. that collective commitments of the will enable a collective standpoint and that this grounds their normative authority. Given this outlook, we can understand joint commitment as grounded in our nature as *planning agents*; specifically, as grounded in our nature as *planning agents who operate in a social world*. We can say, then, that to be jointly committed is to be caught up in a structure that allows us to conceive of ourselves as part of a plural subject. In this section, I will expand on this picture and argue that it follows from it that the process of forming a joint commitment need not be thought of as the combining of mutually interdependent individual willings, as Gilbert's PST seems to imply. Rather, I will argue that the process of forming a joint commitment can be merely the very process of living as *socially planning agents*.

To help further unpack what I take the process of living as socially planning agents to be, and how I see it as related to joint commitments, let me marshal another concept: '*agentive life*'. I shall understand this term as capturing those aspects of the existence of an agent (whether individual or plural) that *conceptually necessitate* that they are an agent. By saying that these aspects conceptually necessitate the existence of the agent *as agent*, I mean to say that, they are aspects that are properly understandable only given the assumption of the existence of an agent. Thus, if we live agentive lives, then, in order to authentically understand ourselves, we must understand ourselves as agents, and thus we must understand ourselves as being committed by our intentions. So agentive life is underpinned by the fact that the lived reality of actual people requires the possibility of agency to be *authentically* understood.

So, what are these aspects of one's life? As before, let us start with the individual and then move on to consider the collective. In light of the previous discussion of agency, it is clear that the performing of actions requires that we understand ourselves as agents. For example, if I perform the action of climbing a mountain, then I can only understand this as something that I intentionally do, if I understand myself as an agent. Performing actions is thus clearly an aspect of one's agentive life. But, not everything that happens to us requires *in-itself* that we are able to understand ourselves as planning agents. As I discussed back in

the first chapter of this thesis, we need not see all things that happen to us as being intentional acts.³²⁴ For example, if I fall off a mountain unconscious, I do not need to be able to understand myself as an agent in order to understand this falling. If I had been an inanimate object, say if I had been shot dead before I fell, the falling of my body would still be characterisable as the same sort of event. Though such non-intentional events do not require the existence of a standpoint to be understood in isolation, we *do* need such a standpoint to understand them in terms of their *embeddedness* within the ongoing lives of ourselves as agents. Thus, if I am still alive as I fall, I will want to understand this falling as the falling *of the same agent* that formed and carried out the intention to reach the summit of the mountain. That is, I will want to understand it as the falling of myself the agent. Further, I will want to understand the consequences of the fall in terms of my ongoing plans, for example to get back home and raise my children. What this means is that given that I understand my life in its totality as an agentive life, I am required to make sense of my now-falling in terms of how it fits in with my wider agentive plans, goals and aims. One has no other option but to live in this way because we identify ourselves as singular things that live agentive lives. Unless we have some kind of cognitive issue, such as multiple personality disorder,³²⁵ we can have no rival identities to fall back on – it is my entire understanding of me as the kind of thing that I am. Of course, there are some things that happen to me that are of no consequence to my agentive life. For example, every passing second, trillions of neutrons move through my body, not interacting with it at all. It would be wrong to say that it is required of me that I understand their moving through me in terms of my being an agent. However, what we can say is that as an individual I thus have no choice but to see *all* of those events that are embedded in my life as an agent in terms of my being a singular ongoing agent.

It is clear from the many examples given throughout this thesis (not least the four canonical examples of the lovers, the mountain climbers, the residents' association and the rugby club) that in addition to individual agentive lives, we also live *collective agentive lives*.

³²⁴ See Chapter 1, Section 1.2.

³²⁵ I take it that we have a non-instrumental reason not to want to have such a cognitive disorder. This seems like a plain truth about the nature of being a unified individual. At the very least, we might say that it would be the height of inauthenticity to believe that we could switch between standpoints (i.e. identities) intentionally.

It is plain that the performing of actions, in this case collective actions, requires that we understand ourselves as plural agents who are performing those collective actions. For example, if we perform the action of climbing a mountain, then we can only understand this as something that we intentionally do, if we understand ourselves as a plural agent. Further, just as in the individual case, not everything that happens to an individual requires *in-itself* that they understand themselves as individual planning agents, not everything that happens to collections of people requires those collectives to understand themselves (taken together) as plural planning agents. However, again in parallel with the individual case, if they *have* lived their lives as ongoing plural agents, as joint social planners with continuity of existence, they should thus make sense of even the non-intentional things that happen to them in terms of the embeddedness of such events within their ongoing collective lives. That is, they will make sense of them in terms of their ongoing collective plans, goals and aims.³²⁶ Thus, for example, given that Hillary and Tenzing climbed Everest as a team, if they had fallen from the mountain then this unintentional falling would not *in-itself* require that they understand themselves as a joint agent. However, they would have understood their falling (had such a catastrophe occurred) in terms of their ongoing plural agency; they would understand it in terms of their joint commitment to be the first climbers to reach the summit. There will of course, just as in the individual case, be some minimally embedded/non-embedded events that we need not understand in terms of the ongoing agentive life of the plural subject; for example, if Tenzing were to momentarily gaze at a mountain flower as he passes it. However, what we can say is that *all* of those events that are embedded in their collective life need to be understood in terms of their life as an ongoing plural agent.

There is an important difference between the individual and collective cases. The individual has no choice but to see all events that are embedded in their lives in terms of their ongoing agentive life, because they have no other viable way of understanding who they are. In contrast, because they can always fall back on this individual identity, it is *in some sense* possible for individuals to abandon the understanding of the collective *as agent*.

³²⁶ I take it that the discussions of the previous chapters have shown that we do indeed understand our particular collective acts not in isolation but rather as entangled in ongoing plural agencies (in particular see section 5b.1 of Chapter Five).

In this sense, our being part of any particular collective agentive identity is potentially *escapable* in a way that our individual agentive identity is not. This escapability, however, is not straightforward. Rather, it is reined-in by the fact that abandoning the collective is almost never without the cost of loss of self-understanding; this was established in the last chapter. So, to the extent that the agent *has* in fact lived part of their life in ways that are collectively agentive, they abandon their ability to properly understand this aspect of their lives if they abandon seeing ongoing events in terms of this plural agent. Thus, if Hillary refuses to see those things that happen on the mountain to himself and Tenzing (that are in fact embedded in their plural lives) in terms of their joint commitment to climb the mountain, then he abandons the ability to understand their life on this mountain as a collective agentive life. An important extra point is that, even the parts of their collective activity that do not *in-themselves* necessitate the existence of an agent can be embedded within the wider framework of their ongoing plural agency. Hillary thus abandons the ability to understand the *wholeness* of their activity.

The relevance of the multiplicity of the events which are embedded in our need to understand our – individual or collective – selves as agents, is that it shows the breadth of the agentiveness of our lives. It is thus a broadening out of the transcendental collective self-governance explanation. That is, a broadening out of the argument, made in the last chapter, that having engaged in a particular collective action, we have no choice but to see that collective as an agent.

In the last chapter I said that our need to understand ourselves as part of plural agents differs from our *a priori* need to understand ourselves as individual agents. We have no choice but to engage in individual action. In contrast, though there are many benefits to living a social life, it is merely contingent that we do. Further, I said that, given that we *do* engage in collective acts, we *are* required to understand ourselves as being part of plural agents. The argument in this chapter has the same structure but a different focus. Rather than the consequences of a *particular* collective act, the current argument focuses on how it is that the agent understands themselves *in general*. The need, for this general understanding of oneself as being an agent that is able to engage in collective acts, arises

through our living of socially agentive lives as described above; such that, particular activity is not made sense of in isolation, but rather it is made sense of as embedded in our ongoing plural agencies.

By recasting this argument in terms of this notion of living an agentive life, we can move beyond merely understanding the ground of collective normative constraint (the conclusion of the last chapter) to understanding what this tells us about how it is that we *actually* come to form joint commitments. For Gilbert, the creation of a joint commitment is a conceptually simple thing; it is something that happens through the voluntary willings of multiple individuals together.³²⁷ She thinks of pooling of wills as akin to the readily understandable notion of pooling of money: in both cases we voluntarily will something to come into existence that then cannot be considered as consisting of discreet separate pieces.³²⁸ As Gilbert puts it, “[f]or each party to enter into a joint commitment is for him to *allow* his will to be bound”.³²⁹ More precisely, Gilbert believes that joint commitment comes about when, in conditions of common knowledge, each has expressed their *willingness* to be jointly bound. For the modified plural subject account PST*, the creation of a joint commitment can be as simple as Gilbert suggests – it can be a fully and transparently voluntary act – however, it can also be a *messy elongated product of real and complex social life*. How this works in particular cases will be explored in the examples below of love and political obligation. Here I will expand on how PST* implies that entanglement generally works.

327 Or at least, it is *fundamentally* simple. She does, as I have discussed and shall return to below, allow the messiness of real social life to come into view insofar as she allows for the possibility of implicit commitment. However, as I have claimed before, it is quite opaque exactly what is going on in Gilbert's examples of implicit commitment, such as that of Quiet Harbour (Gilbert, 2008, p. 487). At any rate my claim is that the process of becoming jointly committed is even more messy than Gilbert's account allows for.

328 Interestingly, I have found that undergraduate students find the notion of a pooled bank account not as readily accessible as those of us who are older. My (tentative) sociological claim would be that this demonstrates the motif of this chapter: that it is living a social life (in this case I guess the life in an long-standing and economically co-dependent romantic couple that most undergraduates have yet to experience) that forces us to have to see things in a collective way. See discussion of romantic couples in section 6.2 below.

329 Gilbert, 2006, p.154.

On the picture painted above, joint commitment should not at base be understood as being created by interdependent willings of individuals in conditions of common knowledge. Rather, it should be understood as the – sometimes unintended and ignorant – messy product of individuals simply living as social agents. In this sense, rather than each individual having to pool their will through an individual act of expressing willingness to do so, each individual's will can become *entangled with the plural will by the very act of living in such a way that presumes the existence of this plural will*. This is what I understand as entanglement: the independent threads of people's lives become twisted together to form a mass that one can no longer understand as a set of independent threads, but rather, one can only see as a single cord. Because of the kind of existence we have, entanglement, we might say, *is* the process of living a social agentic life.³³⁰ The upshot of this is that rather than each party necessarily voluntarily entering into a joint commitment by allowing her will to be bound, each party can become subject to a joint commitment through the process of coming to see her will as entangled with the plural will. Or to put it in slightly different terms, for one to enter into a joint commitment is for one to come to be bound inasmuch as it requires that one understand part of one's agentic life in terms of the agentic life of the plural subject. The plural will can thus be understood as something that is constructed. It is constructed in the same way that the individual will is constructed: a standpoint is created through our taking ourselves to be jointly constrained by collective intentions. Our being more or less entangled is our having more or less at stake in this understanding.

Allowing-in the messiness of social life is not meant to imply a negation of Gilbert's observations about our experience of constraint by the collective will. The obligation criteria, that each participant has a *pro tanto* obligation to promote fulfilment of the intentions of the plural will, still holds according to PST*. As does the permission criteria, which says that participants understand that they are not (ordinarily) in a position to unilaterally 'by a simple change of mind' remove the constraints imposed on them by the obligation criteria.³³¹ However, there are two important differences between the experience

330 Note that, though being entangled is a product of having a social history it is *not* equivalent to it. Only social interactions that require the existence of a plural subject (directly or through their embeddedness) create the conditions for entanglement.

331 See Gilbert, 2000b, p.17.

of the plural will implied by PST* and that implied by Gilbert's PST. Firstly, rather than entering into a joint commitment being something that an agent – however subtly – must feel themselves to choose to do, it can be something that they can feel non-willingly drawn into. This is explained above in the discussion of the idea of agentive life, i.e. that inasmuch as the agent finds that they have to live a social life, they will find that they have entered into a joint commitment and are part of a plural agent. Secondly, because there is variability in lived agentive life, this creates the possibility of variance in how it feels to rebel against the obligation criteria. I discuss this below.

We might say then that Gilbert's pooling of wills is *binary*, it is all or nothing; you either are part of a joint commitment or you are not. In contrast, my notion of entanglement is *analogue*, it is open to degrees; your embeddedness within the joint commitment is on a sliding scale. For Gilbert, your will is either pooled or it is not pooled, whereas, according to PST*, you can be more or less entangled with the plural will. This is not to say that PST* claims that there is variance in the level of normative requirement towards the collective will. Whatever degree you are entangled, this entanglement still requires you to see the collective will as constraining. What varies is not the fact of obligation; this is constant. Rather, what varies is the *cost of your transgression in terms of loss of self-understanding*. By the cost of loss of self-understanding, I mean the cost to one's ability to understand one's life as the life of an agent. Your experience of your transgression will thus vary along this axis. That is to say, there will be variance in what we might call the *normative pain* you experience in rebelling and this will depend on the level of your entanglement.³³²

What do I mean to invoke by the concept of 'normative pain'? Well, just as we can understand physical pain as a negative feeling, of variable strength, experience of which indicates bodily damage, normative pain is a negative feeling, of variable strength, experience of which indicates normative transgression. Just as Ulysses, who was discussed in

³³² Of course, there are all kinds of reasons why your *total* experience of rebelling may vary; your emotional response tied to such things as feelings of loyalty, how tired or hungry you are, whether the sun is shining etc. may be the overwhelming determiners of your total experience. My claim is just that normative pain, varying along the axis I describe, will form a distinct and important *part* of your experience; it may well not always be the largest part, but – as I will try to show in the examples below – it is a recognisable part.

an earlier chapter as struggling against his binding by rope to his ship's mast, will feel physical pain pulling against the physical constraint of the ropes, one will feel normative pain when one struggles against the normative constraint of the collective intentions that scaffold our social lives.

So, there is a strict requirement to follow the collective will. This does not vary. Rather, it is just that the cost in terms of loss of self-understanding of violating this obligation varies. Variance in normative pain is indicative of variance in this cost in terms of loss of self-understanding.³³³ Admittedly, the concept of variance in cost in terms of loss of self-understanding does give you a variance in your additional instrumental reasons to avoid violation. This is because normative pain, as the name suggests, is something that feels uncomfortable and we have reason not to want to feel uncomfortable, either physically or normatively. So while the strict obligation must remain constant, the slack recommendation (to keep to that obligation) can vary.

In summary, the modified Plural Subject Theory gives the following account of collective action: collective action is the action of a plural subject that is created by the will of members of the collective becoming entangled together to form a plural will through their living of a collective agentive life.

6.2 – PST* as better capturing the phenomenology

In this section I will set out two examples that illustrate the advantages of PST* over PST in capturing the phenomenology of our social lives. PST* does not differ from PST in seeing the base characteristics of the phenomenology of collective action as: unity, collective intentionality, constraint and detachment. Rather, PST* merely modifies how we should expect these base characteristics to be experienced. As explained above; these modifications are (i) the not necessarily voluntary nature of unity and (ii) the variance in the feeling of normative pain of breaking with constraint. The first example will be an

³³³ I take my notion of the possibility of variance in normative pain due to variance in entanglement through lived agentive life to be similar to, though expressed in quite different terms, Elisabeth Pacherie's notion of the possibility of variance in what she calls our "*sense of agency*" (2012).

expansion on the canonical example – which I set up in Chapter One – of a romantic couple. The second will be that of political obligation as Gilbert sets up in *On Political Obligation*.³³⁴

6.2.a – Romantic Couples

In Gilbert's example of the romantic couple, Tony and Celia, sharing a cake, we saw that a romantic couple is a form of plural subject.³³⁵ Marriage, with its formal vows and declarations, might be thought to be the clearest form of relationship of this sort. However, especially in our modern world, becoming a romantic couple need not be a matter of making any formal vows or declarations.³³⁶ Thus far, this description of romantic couples may sit fine with Gilbert's vision of the formation of a plural subject. As has been explored, she does not require that partners must make official declarations. Rather, she believes that our expressing our joint willingness can be informal and even implicit in our general behaviour. Thus, though unmarried romantic couples may never have formally made a joint commitment, they will have lived as a *we*; as Gilbert notes they will have done many things jointly such as "... creating and maintaining a comfortable home, raising a family ... maintaining a joint bank account, buying a car, visiting the parents and taking the kids to the zoo."³³⁷ These collective actions will be part of the joint life that the couple live together.

This is all correct. However, a stronger claim also seems to be true: not only is it the case that we can find ourselves in a romantic couple without making a formal declaration; further, we may well find ourselves becoming drawn into being in a romantic couple against our will. Think of a casual 'just seeing each other' pair of individuals who deny that they love each other until it dawns on them that they are, de facto, a couple. This speaks

334 Gilbert, 2006.

335 Or to put it less strongly, the types of relationships that people who are in love commonly form are plural subject ones. I do not mean to imply that the *only* way to be in love is to be in such a relationship.

336 Gilbert's claim that, though ongoing romantic union can exist "...without the benefit of marriage", marriage remains "... its usual locus in culture as things stand" (Gilbert, 1996d) seems increasingly anachronistic, at least in the UK where according to European Union statistics nearly half of babies are now born to people who are not married. (Eurostat, 2013)

337 Gilbert, 1996d, p.222.

fundamentally against Gilbert's voluntaristic vision of forming a joint commitment. In defence of Gilbert, it might be complained that here I am mixing two things up: having romantic feelings and becoming united. The first, it might be argued, is the thing that we feel we have no control over. That, however, does not capture the force of love: to say that we cannot help falling in love is not simply to say that we cannot help having romantic feelings towards someone, it is to say that we cannot help finding ourselves as being united with them. This is because, by virtue of *being* and *acting* as a plural subject, lovers find that their agentive identities become entangled; a where-*we*-stand emerges and we cannot help but invest more and more of our sense of where-as-individuals-we-stand in it. This is partly because of the inescapability of romantic feelings that compel us to be together, but also just because of the practical realities of negotiating a social world that treats romantically engaged individuals as a single unit. While this does not fit with PST, it is a perfect example of the idea of joint commitment as arising through entanglement that we find in PST*.

The second aspect of romantic love, which I want to argue fits better with PST* than PST, is that there is a certain anguish to going against the collective will (that is, going against the will of your joint couple-self) which is separate from the strength of your emotional attachment to the other person. We can see this in the following example: imagine that a romantic couple, Mary and Claire, are on their way to the airport to go on holiday. Suppose that things do not go smoothly and that Mary turns around, before they get to the terminal, and starts walking back home. Her partner Claire would be justified to say, "You can't go home, *we* said *we* were going to Majorca!" This fits with Gilbert's

conception of the obligation criteria.³³⁸ However, what I want to suggest is that, though this is correct, it leaves out the possibility of the variability of normative pain.

To demonstrate this possibility, let us think about the following three different couples. Firstly, we have a newly formed couple, Bill and Ted. Let us say that they have been together for a few months, seeing each other once or twice a week. They see themselves as very much in love; they are full of the strong feelings of lust and desire, but live in separate houses, have separate groups of friends and different hobbies. Secondly, our Claire and Mary. Let us say that they have been together for twenty-five years, bought a house together, adopted and raised a child together, share friends and have the same hobbies. They feel still very much in love. Finally, imagine a couple, Patrick and Madeleine who have been in a relationship longest of all, thirty years. They do live together, and pay bills jointly and have raised children. But they do not really feel much romance toward each other and often find themselves attracted to other people. Now suppose that each of these couples is on their way to the airport and one of each of them turns round and starts walking back home. Each rebel will be open to a rebuke of Gilbert's form: "You can't go home, *we* said *we* were going to Majorca!" However, the act of rebelling will feel different *in a relevant way* to each.³³⁹ The relevant difference (for our discussion of joint commitment) is as follows: when Ted turns round to Bill and announces that he is not going the wrong that he is committing by violating the collective will feel less normatively

338 Andrea Westlund raises the worry that any account of love that allows individuals to be bound to the collective will must be incorrect, because it cannot respect the reciprocity that she argues is at the heart of any loving relationship. Against this I argue (in Kisolo-Ssonko, 2012) that while it is true that when a relationship is healthy each party tries to take the needs of the other into account, this does not mean that the collective will itself must be *automatically* reflective of any change in each individual's perspective. So, for example, in the holiday case: if Mary and Claire are in a well-functioning romantic couple then each will try to accommodate the other. Even once they have come to their collective decision they will be open to the possibility of further deliberation over its content. If Mary comes to realise that she does not want to go to Majorca, then this *does* mean that Claire *ought* to be open to changing their collective intention, but it *does not* imply that the existing collective intention automatically ceases to hold sway. Westlund's confusion stems from the fact that she idealises romance, ignoring the fact that unhealthy romantic couples still count as plural subjects. Given the reality of sexism in our society, a more realistic picture of actual marriage could well be that given by Susan Moller Okin, who says that "... gender structured marriage involves women in a cycle of socially caused and distinctively asymmetric vulnerability" (1989, p.138). A similar point regarding the idealising of love (though not directed at Westlund, nor acknowledging the structural influence of sexism in quite the terms used by Okin) is noted by Gilbert (1996d).

339 It will of course also feel different for lots of reasons that are irrelevant to the case in hand. For example, the different emotional temperaments of the individuals.

painful than the same wrong will feel to Mary. Similarly, the wrong of Madeleine's doing so will feel even worse than Mary's. The differences between the three couples exist because of differing levels of entanglement, and thus differing levels of cost in terms of loss of self-understanding in abandoning the collective as plural subject; the more their understanding of themselves is bound up with their being in the couple that they are in, the more discomfort (of the sort in focus) they will feel.

The idea is not that any of the couples experience no obligation towards their collective wills. Rather, the claim is that the degree of normative pain of breaking these obligations will vary. Of course, variation in their *overall* experiences of rebellion from the plural will could be a function of variance in other factors; the different level of moral, romantic, or practical obligations felt by each party. Moreover, it will be affected by their different emotional sensitivities, where these are shaped by their individual histories.³⁴⁰ However, I have tried to set the example up so that variance in normative pain does not correlate with these other factors; thus the rebuke may feel more normatively painful to Madeleine than it feels to either Mary or Ted, even though Madeleine no longer has any emotional feelings of lust or desire for Patrick. This is because, though she is no longer emotionally attuned with Patrick, her will is nonetheless highly entangled with his, due to the length of time they have lived together as a couple. In contrast, though Ted is infatuated with Bill, their agentive lives are less entangled in their plural subject-hood and thus the specifically normative pain will feel less, though the emotional pain may well feel greater. The difference between the couples' experiences lies in the nature of the *we* rather than the personal feelings of each party. This interpretation is made all the more plausible given that the rebuke in question makes *direct* reference to the will of the *we*.³⁴¹

340 See footnote 332 above.

341 The phenomenology of the variance of normative pain can also be seen not only when we look at the phenomenology of rebelling against a particular obligation but also in that of totally breaking with the plural subject in question. Nozick puts it, in a way that fits very neatly with my argument, “[a] willingness to trade up [i.e. to find a new partner], to destroy the very *we* you largely identify with, would then be a willingness to destroy yourself in the form of your own extended self.” (1995, p.235) We might think of this as the possibility of an “*Us! There is no us!*” scenario as the ultimate (but still not without cost) way of escaping from the obligation of the plural subject.

6.2.b – Political Obligation

Whilst Margaret Gilbert, in line with others in the collective intentionality field,³⁴² often focuses on small scale examples of collective action, she envisions her theory as being able to explain even the world of the grand social scale. In *A Theory of Political Obligation* she sets out to achieve the ambitious goal of using her account to explain the nature of our experience of political obligation, in particular, of political obligation towards one's nation state.³⁴³ In this section, I will argue that the way in which we experience political obligation fits better with my Modified Plural Subject Theory than it does with Gilbert's. In particular, I will argue that we can see this in the not necessarily voluntary nature of political obligation and the variance in the felt cost of rebellion against our political state.

As I explored in depth in Chapter Three, we can think of the construction of plural subjects as the creation of an alliance.³⁴⁴ The forming of such an alliance need not be political, at least not in the strong sense.³⁴⁵ However, the language of alliance (as with that of related terms such as partnership, commitment, unity and such like) appears particularly apt to capture the world of politics; it is common to hear people talk about such things as “our party's views”, “our duty to defend our nation's religion”.³⁴⁶ This match between Plural Subject Theory and the way we conceptualise our political world was evident in the comparisons I made between Gilbert's notion of unity between people and types of social contract theory.³⁴⁷ Whilst the political realm might be phenomenologically ripe for interpretation in plural subject terms, those of an individualist bent, sceptical as they are about Gilbert's claims to have captured the phenomenology of small scale groups, are even more likely to claim that there is nothing as grand as a plural subject on a large scale. The

342 See Chapter One, Section 1.3 for a survey of the field.

343 Gilbert, 2006 and also pre-empted by an earlier short paper (1996e) where she gives a brief sketch of the application of PST to political obligation.

344 See Gilbert, 1996c, p.271.

345 Though we might think that inasmuch as any social activity involves interaction between people, it is at least political with a small 'p'.

346 Gilbert notes the ease with which “... people think in terms of 'our' government, 'our' constitution and so on” (2006, p.294).

347 See Chapter 3, Section 3.1. As Thomas Smith notes “... there is a formal parallel” between Gilbert's approach and contract for if we “... substitute ‘have agreed’ for ‘are jointly committed’ ... and ‘agreement’ for ‘joint commitment’ ... Gilbert’s argument comes to resemble a more or less familiar contractarian approach” (2007).

problems they point to, however, as I suggest below, can be understood not as defeating Plural Subject Theory, but rather as supporting PST* over PST.

One of the reasons that people are sceptical about explaining political obligation on the plural subject model is a variant on the '*no actual contract*' criticism of contract theories of political obligation. Contract theories of political obligation hold that we have obligations towards others in our society because we have formed a contract with them. Critics of this view point out that no actual agreements to such contracts have been made by the general populace.³⁴⁸ A similar criticism might be made of Plural Subject Theory as applied to the state, i.e. it might be claimed that it is apparent that there are no actual expressions of readiness to be jointly committed to the state. Gilbert agrees with the no-actual-agreement criticism of contract theory. However, rather than seeing it as a complaint that can also be raised against her theory, she sees it as an important advantage of her theory over actual contract theories. The validity of her doing so is grounded in the fact that she does not believe that joint commitment requires explicit agreements; rather, her theory allows for the possibility that the expressions of readiness to be jointly committed can be much more subtle and elongated over time. Such subtle expressions are evident in the ways we communicate in the everyday social world. As Gilbert puts it:

“... it is common knowledge ... that in face-to-face conversations, letters, and so on everyone speaks without hesitation of 'our country', 'our constitution', 'our laws', and so on in relation to the population as a whole. They speak of what 'we' are doing in terms of both international relations and internal issues. They evince guilt, pride, and other such emotions over such things. And they give no indication that they do not wish the plural subject interpretation to be made.”³⁴⁹

So, Gilbert's theory can allow for the lack of explicit agreements in a way that contract theory cannot. However, there is a tension in Gilbert's writing between, on the one hand, her desire to allow joint commitments to arise in non-explicit ways and, on the other, her need – required by the formal structure of her theory – to say that “... [e]ntry into a joint commitment is [always] at some level voluntary”.³⁵⁰ This has the outcome that her theory of political obligation struggles to deal with the *possibility of hold outs*. Gilbert

348 See Gilbert (2006, p.70 -75) for a concise summary of the 'no-actual agreement' objection.

349 Gilbert, 2006, p.244.

350 Gilbert, 2006, p.168.

allows that some people may be 'hold outs'; they may reject the political society they live in. She conceives of these as people who live reclusive lives; explicitly rejecting society's invitation to join in joint commitments.³⁵¹ The question we might ask, however, is: what is to stop us thinking about people who *do* play along with society as also being hold outs? What is to stop people who do not isolate themselves, but live with others, simply choosing to do so without signing up to Gilbertarian joint commitments? Recall that I raised this question in a previous chapter when considering the possibility that our rugby team member might simply pretend to be seeing himself as part of the collective action of pushing the bus up the hill. The rather grander variant of that question we can ask here is: *what blocks the possibility that every member of a political society is simply playing along with things?*

Gilbert seems right to think that we find ourselves in political societies that we feel ourselves to be obligated to, even though we may not feel ourselves to have explicitly signed up to these societies. However, her theory seems to be at a loss to explain why it is that we cannot just all be secret hold-outs, merely pretending to play along with the national plural subject. In contrast, our actual experience of political obligation seems to be of feeling it to be something non-voluntary in a way that Gilbert's account does not capture. Here the advantage of PST* is that it can rule out the possibility of hold outs who simply play along with social life. This is because for PST* it can be the very process of living a social life that creates joint commitment and, just as we have seen in the example of the romantic couple above, the fact that one can feel oneself to be forced to live such a social life means that one can feel oneself to be forced to enter into a joint commitment at the level of the nation state.

The tension generated by the possibility of non-voluntarily entering into plural subjects becomes all the more sharp when we move from consideration of liberal democracies, where people are assumed (whatever the actual realities) to have some sort of participatory possibility to more unsavoury political setups. For example, if we suppose that the German state in the 1940s constituted a plural subject, we might feel uneasy about

351 Gilbert, 2006, p.296.

assigning obligations to play their part in invading France to all members of the German state. Had all the citizens taken part, even if just implicitly, in the creation of a joint commitment? As John Searle notes, any description of the time must acknowledge the complexities and the differing levels of involvement and motivation. He says that “[a]t the time of the Nazi regime ... members of the Nazi Party enthusiastically endorsed the institutional structure of the Third Reich. But there were lots of people in Germany at the time, who, while not endorsing the institutional structure went along with it as a matter of nationalism, indifference, prudence, or even just apathy.”³⁵² Gilbert must surely struggle to incorporate such states into her theory, even though she does appear to want to allow the possibility of cohesion. PST* however does not struggle, for rather than require explicit endorsement, or even implicit endorsement, we can just say that it is the experience of living a social life that entangles the citizens into the plural will of the German state.³⁵³

The example of the alliance to the German state of the 1940s also highlights a second key point at which PST* succeeds in reflecting our phenomenology better than Gilbert’s PST: the possibility of *variance of experience of political obligation*. It seems to me that at least a large part of the reaction against extending Gilbert’s PST to political obligation, and the scale of states, has to do with the feeling that Gilbert’s binary notion of joint commitment (i.e. you either have committed or you have not) implies a binary experience of political obligation that does not fit with our real life experience. Against this, it seems to me that Gilbert is right to think of our obligation as all or nothing. Obligation is, after all, a strict normative requirement, rather than a slack one. However, PST* allows us to see that there is a *different space* from which we can explain the variability of how we experience our obligation. According to PST* the variation can be in our level of

352 Searle, 2010, p.57.

353 An alternative way to face down the unease such examples create is to claim that only liberal democracies count as plural subjects. This is problematic. Firstly, because we might question the extent to which even so called ‘liberal democracies’ live up to the promise of genuine democratic participation. But more than this, such a move seems to commit the same error of idealising plural subjects as I claim Westlund commits in the case of conceptualising love. As I noted above (see footnote 338), an ill-functioning romantic couple can still count as a plural subject. The same should also be true for an ill-functioning state for, as Gilbert says, “[c]oercive circumstances need not prevent me from entering into a joint commitment. If I am party to such a commitment I am obligated and that is that” (1996, p.373) and thus “[p]olitical obligations ... offer practical support to tyranny, as they do to any form of political society.” (Gilbert, 2006, p.286) It is this that leads Gilbert to claim that there is something “tragic” about group membership. (1996f, p.387).

entanglement and thus in the level of normative pain that we would feel in rebelling against the collective will. This will give rise to a very different experience of being in conflict with the collective will; however, this need not be a difference in the actual obligations that we have.

The plausibility of such a way of understanding the variance in experience of political obligation can be seen in imagining the following contrasting examples: suppose that Karl lives in a rough part of town and that he reasons that his personal interests will be best protected by joining the local Nazi group chapter. He joins it and shapes his life around his membership of it. He sings Nazi songs and he thinks of his interactions with others in terms of their place within the Motherland. Contrast Karl with Bob, a rather more indifferent citizen. Let us suppose that Bob lives out in the countryside. He has less contact with the state apparatus, perhaps there is not even a local chapter of the Nazi party in his tiny village. He is a member of the German state, and he knows it exists, but it has only a marginal place in his general life. It seems obvious that the committed nationalist Karl will experience rebelling against the plural will differently to the indifferent citizen Bob. There thus appears to be variance in the extent that we feel political bodies, states or parties to obligate us, which mirrors the variance in our political leanings. According to Gilbert's theory, the obligation to the state should be experienced by all as being equally binding upon them. An individualist might claim that this shows that variance in experience of national collectivity is really just variance in the extent to which one is being deceived by political rhetoric. PST* offers an alternative to both of these; an alternative that allows us to claim that our experience of the state as a plural subject is not a mere illusion, but at the same time allows for a variance in how people experience the obligating force of national collectivities. This is because we can see variance in political leanings as changing the way we live our social lives, and this in turn as changing the extent to which we are entangled in the plural will.

Bob and Karl are, according to PST, both obligated to the state and its collective action. However, if they rebel against this obligation, the normative pain of doing so will differ for the two of them in a way that tracks their level of entanglement. Remember that I

am not claiming that this will explain the entirety of their experience of rebellion. Human beings are complex creatures and other factors, such as for example differing attitudes towards risk, will colour their experiences in all sorts of complex ways.³⁵⁴ My claim is only that there is an element of their experience of the plural will that will vary in its painfulness in the way I describe. According to PST*, what we sacrifice in rebelling against the collective will is our ability to understand the agency of the group and our part in it and, as I have said above, the normative pain of this is variable depending on how much we have conceptually invested in the agency of the collective. Thus, Karl will experience defecting from the national will in a different, more normatively painful way than Bob. It will feel worse for him to reject the political obligations he feels towards his fascist party because a greater number of the things that he has done will be tied up in understanding that plural agent as an agent. In this sense, the nationalist may not just be living an illusion when he feels himself to be bound to the nation state in a way that the indifferent citizen does not – rather the nationalist *because he has lived and understood his life as a nationalist*, will *actually* suffer more normative pain in rejecting the national will than Bob would.

6.3 – Chapter Summary

In this chapter I have developed those aspects of PST* that distinguish it from Gilbert's theory, in particular the experience of the sometimes non-voluntary nature of forming a plural subject and the possibility of variance in normative pain. PST* is not intended to be a new theory replacing PST; rather, it should be seen as a refinement of it – a refinement that can be summarised as coming from supplementing Gilbert's notion that the creation of a plural will requires the voluntary pooling of wills, with the insight that wills can become entangled simply through the process of a lived social life. I have argued that this refinement is exemplified in the phenomenology of our actual social lives, such that we experience a variance in the normative pain of breaking with the collective will that corresponds to the level to which we are entangled. In the next and final chapter I will give a broad sketch of how PST* impacts on our general understanding of our social worlds.

³⁵⁴ Again, see footnote 332.

7 – Conclusion: Capturing the messiness of social life

What does it mean to be an agent? This question is fundamental to our understanding of ourselves. To properly address it, we must recognise that we are not lonely Robinson Crusoe figures. Rather, we are social creatures. Acknowledging this requires more than just seeing the interconnection between individual acts. It means recognising the fact that *we* do things together. *We* raise families; *we* conquer crevasses; *we* win leagues; and *we* save our green community spaces from development; *we* do all these things and more. In doing so, *we* embody a kind of agency which has only recently come to be seen as important in analytic philosophy: *plural agency*.

Pitting our understanding of agency against social reality undermines any attempt to fully understand our lives using a limited conceptual schema according to which individuals are the only valid agents. Even sophisticated individualist outlooks, such as Michael Bratman's Shared Cooperative Activity Account, fail to do justice to the phenomenology involved in our social existence. This forces us to expand our outlook. It forces us to consider the possibility that to collectively act is to act *together as plural subjects*. Free from the straightjacket of individualism, we can embrace the fact that by uniting together *we* create *intersubjective intentionality*.

Margaret Gilbert succeeds in mapping out the general contours of this terrain. Most importantly, she captures the possibility that the plural will can be in conflict with the individual will. This, however, creates a conceptual challenge: finding a way to make sense of *intersubjective constraint*. That is, finding a way to make sense of individuals being limited, in what they can rationally do, by the collective agency of which they are part. Overcoming this challenge requires the use of the same conceptual tools we can use to understand the constraint of individuals by their own wills; in particular, the conceptual tools provided by Bratman's planning theory of agency. In deploying these tools, Gilbert's notion of the plural subject becomes grounded in a substantive conception of agency. It

becomes grounded in the fundamentals of what it means to be a planning agent in a social world.

This taking of a scalpel to the innards of plural agency reveals that the relationship between the individual and the collective is neither clean nor neat. Rather, it is messy and tangled. Instead of a unity of individuals voluntarily pooling their wills, the actual process of living a social life may entangle wills together. Letting the messiness of real social life into our theory will impact on our general understanding of our social worlds, and, to conclude, I will highlight some of the ways it might do so. This will both encapsulate the difference between my proposed modified theory and Gilbert's original, and point towards interesting areas for potential further research.

In Chapter Five I asked: what bars the possibility of someone merely pretending to be an authentic participant in collective actions? This question was, and is, poignant. Such a possibility would undermine Gilbert's vision of our actual social life being scaffolded by the constraint of collective wills. From an individualist perspective, merely pretending to live social lives is attractive because it seems that an individual who did so would gain greater autonomy. The challenge for the collectivist view is that if such pretence is attractive, then why not think that the whole operation of society is nothing but a sum of such inauthentic pretences? PST* can face off this danger because, according to it, the pressure to live authentic social agentive lives means that one almost inevitably becomes entangled in the plural wills that populate our social worlds. So, for example, even if our member of Leicester Tigers rugby club is *attempting* to merely pretend to live a social life,³⁵⁵ they will come to *find themselves entangled* with the plural will that they are trying to fake involvement with. On this picture, entering into joint commitments is just not as fundamentally voluntary as Gilbert makes out. It is, in contrast, something that we may find we cannot help but do. Being an agent in a social world thus means not being fully in control of the agency we are part of.

355 See discussion in Chapter 5, Section b.2.

This point about the *inevitability of social life* is the key insight that must inform our view of plural agency. One of its consequences is that it muddies the water when we try to set up examples of coordinated acts that are not quite collective actions. So much can be seen in one of the key examples from Gilbert's foundational work on plural agency, *On Social Facts*.³⁵⁶ Gilbert introduces a thought experiment that has us start by imagining a set of individuals who clearly have no social life. She then has us consider what minimal characteristics must be added to make the set social. She starts with a set of individuals who survive by picking mushrooms alone in the woods, never interacting. She then adds various conditions of ever-increasing interpersonal interaction. In doing this, she claims to show that, even where we allow that our mushroom pickers regularly interact, it is still logically possible that they fail to be united as a plural subject.³⁵⁷ Gilbert is correct to present such a possibility as a *logical possibility*. However, she fails to recognise quite how hard it is to imagine that in the *actual world* this set of circumstances could exist, whilst still no social group be formed. She says that, "... it does not really matter whether or not it is true that after finding themselves in this position that people would naturally get together to form a social group as long as they do not yet constitute one".³⁵⁸ This is correct as far as it goes. However, the inevitability with which social life, and hence the generation of a plural subject, will flow from this situation seems to be underplayed by her understanding of joint commitment. If we accept that the mere living-of-a-socially-agentive-life can create plural subjects, then, inasmuch as the *material conditions* of the mushroom pickers are such that they are likely to come to live a social life, those material conditions will drive them to being a plural subject.³⁵⁹ The question of when, and where, our agency is social, or is not social, is thus hard to answer precisely.

356 Gilbert, 1992.

357 Gilbert, 1992, p.36- 43.

358 Gilbert, 1992, p.38.

359 Allowing for the notion that the material conditions of people's lives can drive construction of plural subjects opens up the possibility that we might be able to think of the Marxist theory of historical materialism in plural subject terms. Thus, for example, we can see that when Gilbert says "[w]hen Marx and Engels urged the workers of the world to unite ... the message behind the call to unity was that workers should set themselves up to act jointly." (Gilbert, 1992, p.228) This need not mean that they should set themselves up to share in readiness to jointly act but rather just that they should live collectively agentive lives. *Calling on the workers to unite is, thus, calling on them to live socially agentive lives* and this is a process that Marx saw as being made historically inevitable by the change in the means of production that pulled people away from the countryside and isolated individual labour and towards the town and mass factories with collectivised social labour.

As I also highlighted in the first chapter, the question of the general rationality of social cooperation will be affected by how we understand collective action. I cannot give here a full treatment of what the Modified Plural Subject Theory might tell us about the nature of this problem. However, a broad outline of what the results of such an application might look like is easy to sketch. As I detailed in Chapter One, Martin Hollis suggested that we might think that the problem of trust is solvable "... if [we] can trust one another to *play as a team*".³⁶⁰ This is because, from the perspective of the team, cooperation is rational. It is rational in that, even when pulling out is the best option for the individual, cooperation can remain the best option *for the team* – as each person properly playing their part is what is needed for the completion of the collective goal. For individualists, the puzzle of such an account is that "as a team" is always to be read as "as *if* there were a team". That is, thinking from the team perspective is always, in essence, just a pretence. By accepting the possibility of plural agency, the nature of the question changes. The problem is no longer one of asking why it is that the individual might pretend to think as a fictional amalgam. Rather, it is one of asking what the relationship is between the individual and the – *actually existing* – plural subject, of which they are part.

With the above in mind, we can see that according to Gilbert's picture, the team solution should be cashed out as follows: we make a joint commitment with the other, and then we are bound by that joint commitment as neither of us can then unilaterally remove it. On the picture presented by the modified theory the obligation and permission criteria still hold. However, there is a difference in how the individuals will experience their normative constraint. Furthermore, this difference tells us something instructive about the way in which we experience social pressure to cooperate (or indeed not to cooperate). While reneging on the promise will be normatively objectionable, for the entangled individuals, the normative pain of rebelling will vary according to the levels of this entanglement. Given this, we can say that those social situations where individuals have lived more of their lives working as a team will guard more strongly against the possibility of default. *The more entangled we are the less the problem of trust is a problem.* Conversely,

360 Hollis, 1998, p.137 [*emphasis mine*].

the less entangled we are, the more the problem of trust is a problem. Established and ongoing plural subjects will, thus, have a greater power to facilitate a smoothly running social world than *ad hoc* groups. This tells us that what it means to be a cooperating agent does not have a single answer; rather, it is particular for different individuals depending on the extent to which they are entangled.

Recognising the existence of entanglement allows us to realise that, in terms of an agent's vision of their preferred self, the formation of actual plural subjects is likely to be a disorderly and contradictory affair. We need not feel fully comfortable with the social situations we find ourselves in, yet we find ourselves in them all the same. This means that social life is a powerful force for individuals, but a force that they may not always feel to be on their side. Such an observation fits well with Christopher Kutz's point that:

“... many cases of collective action involve contexts where agents are *alienated* from the end to which they contribute, whether because of coercion, wilful ignorance, or moral qualms. A pacifist takes a job at the nuclear weapons plant, because it is the only job available; an accountant processes the astonishingly large receipts of a pizza parlour, not inquiring too carefully into their explanation”.³⁶¹

In addition to not feeling comfortable with the plural subjects we find ourselves entangled within, we may even commonly find ourselves entangled in *multiple plural subjects, the standpoints of which conflict with each other*. Using Kutz's sense of alienation, and extending Abraham Roth's notion of intersubjectivity,³⁶² we can say that this is the possibility of *intersubjective-alienation*. Such alienation is evident in the experience of Tenzing Norgay, one of the mountain climbers in the example used to illustrate collective intentionality. In his autobiography, Tenzing describes many times where he felt himself to be torn between his existence as a part of the Sherpa community, in particular the Sherpa crew who were performing a collective job of service for the mountaineers, and his existence as a 'true' mountaineer, as part of the group of men primarily motivated by the goal of getting to the top of the mountain for its own sake. After he had achieved the climbing of Everest, the expedition on which arguably he first truly became a 'real climber' in addition

361 Kutz, 2000, p.26 [*emphasis mine*]

362 Roth, 2003. As discussed in Chapter 4.

to being an employed leader of the Sherpa team (a Sirdar), he said that, being “... both a high climber and a Sirdar, with the two different responsibilities, is too much for one man.”³⁶³ Tenzing lived multiple social agentive lives with different components. We might say that he felt *dual loyalty*; that is, that he had, as Ebrahim Mohammed Alwuraafi puts it, the “... emotional experience of being pulled in two different directions”.³⁶⁴ Moreover, we can see that it is his feeling of belonging to two different collectives that pulled him in these competing directions. In the language of PST* we can say that, he found himself in situations where, whichever side he put himself on, he would experience the normative pain of rebelling against one collective perspective. Such tension is a very real part of the phenomenology of social life. As I noted above, Gilbert recognised the possibilities of such felt sense of dilemma but characterised it only as arising from the contradictory pull between external normative pressures such as morality, and normative pressure internal to the collective will. On PST*, we can recognise that there can be a felt sense of dilemma that is *internal to the conflicting collective wills*.

The possibility of a sense of dilemma, which is internal to the conflicting collective wills, can be seen in situations of political rebellion where individuals jump alliance from one collective identity to another. For example, the members of the German resistance, through their acts of opposition to the authority of the Nazi party, set up a new plural subject. In doing so, they may have found themselves entangled in both plural identities, and thus faced conflicting obligations. Such a situation is contradictory, but such contradictions are not theoretical problems. Rather, they are just part of the complexity of political life. We can give a similar reading to Daniela Tagliafico's rendering of the example of the Libyan revolution: Libyan citizens, she notes, faced a situation in which it is unclear what collective political entities really exist; the Libyan Arab Jamahiriya, with its leader, Gaddafi, or the National Transitional Council that claimed to be the true legitimate government; “... one could then ask which of the two competing governments was really existing; was it the Libyan Arab Jamahiriya or the National Transitional Council, or

363 Ullman, 1956.

364 Alwuraafi, p.28, 2013.

perhaps both?”³⁶⁵ On Gilbert's PST, the question we can ask is: which plural subject have the citizens collectively committed to? On PST*, things can be much messier. We need not look only to voluntarily commitments, but can also investigate the extent to which the agent is entangled through their lived social life in *both* plural agencies. We can thus see that the changes in their social experiences will alter this entanglement and therefore result in variance in the normative pain of the move, from recognising the authority of one plural subject, to recognising the authority of another. It will thus always be a challenge for the individual, living in a complex social world, to understand the agency they are part of and the contradictions within it.

One might worry that the case of the two competing governments is better understood as one where entanglement and commitment come apart, for it seems possible that an agent's social life indicates entanglement with the Libyan Arab Jamahiriya but that that agent nonetheless supports the National Transitional Council. However, PST* allows for this as a possibility. It allows for it because supporting the Council, through a voluntary act of will, *is* a form of placing oneself into an entangled situation. You are entangled because you have expressed your understanding of the National Transitional Council as something that you are now part of. We can compare this to the discussion of the couple sharing a cake and the stranger wishing also to share. In Chapter One, I had said that Celia is free to take up (the stranger) Bernard's offer. I said that she can choose to perform a collective act with him, and accept the unity this involves, without having had any prior relationship with him. However, rather than say that plural agency is possible where entanglement can be absent, we can see *voluntarily pooling your will as a way of becoming entangled*. To extend the metaphor of physical entanglement: it is akin to intentionally throwing a ball of string down onto the floor to tangle it up. Thus, entanglement does not require a prior relationship. Rather, it can be generated by the mere making of an explicit commitment, i.e. a pledge of alliance to the Transitional Council, or such like. This does not negate the fact that if the former life as engaged with the Libyan Arab Jamahiriya has entangled you, then in rejecting it you will face normative pain. Further, merely making a

³⁶⁵ Tagliafico, p.84, 2012. Tagliafico's discussion of what she calls the paradox of collective acceptance is framed in Searle's terminology rather than Gilbert's – nevertheless, her insights into the messiness of social life fit equally well into considerations of how plural wills are formed.

commitment to start supporting the other side may well be a weaker way of being entangled than that of having an extended social agentive history. Questions of political allegiance, social life and the nature of agency are thus intimately linked.

In this thesis, I have sought to give a theoretical backing to the phenomenology of our social experience; that is, to our strong sense of doing things together. I have hopefully shown that collective action is a real phenomenon; it cannot be explained away. The framework of individual action is a blinker; *our agency really is extended*. Thus, as we leave our two walkers, satisfied with their day's achievement, at the top of Scafell Pike, we can see their achievement as not only being one of reaching the summit of the hill. Rather, they have also created something that is bigger than their individual selves; they have constructed a plural subject of which they are both members. The existence of this plural subject enriches their lives as it allows them to move beyond their individual agencies and interact with the world as a *we*. The price of this enrichment, this harmony of wills, is the possibility of discord and constraint. As they swap sandwiches and arrange their next endeavour, the hillwalkers might feel they have lost some of their individual autonomy in not having full individual control over their conquest of the hill terrain. However, their real loss is to be weighed against their real gain: by being entangled in the intersubjective plural agent, they obtain the ability to engage in enlarged social projects. The possibility of plural agency is part and parcel of what it means for us to be the kind of agents we are.

Bibliography

- Alwuraafi, E. M. (2013), "Dual Loyalties in Arab American Novel: A Case Study of Scattered Like Seeds by Shaw J. Dallal", *International Journal of Literature*, Vol. 3, IAMURE Multidisciplinary Research
- Anscombe, G. E. (1963), *Intention*, Oxford: Blackwell
- Bacharach, M. (2006), *Beyond individual choice: teams and frames in game theory*, Princeton: Princeton University Press
- British Broadcasting Corporation (BBC), (2010), "China faces growing gender imbalance", *BBC Online*, <http://news.bbc.co.uk/1/hi/8451289.stm> [accessed 16/9/2013]
- (2013), "Iran 'successfully sends monkey into space'", *BBC Online*, <http://www.bbc.co.uk/news/world-middle-east-21230691> [accessed 16/9/2013]
- Brand, M. (1984), *Intending and Acting : Toward a Naturalized Action Theory*, MIT Press
- Bratman, M. (1981), "Intention and Means-ends Reasoning", *The Philosophical Review*, Vol. 90, pp.251 – 265
- (1999a), "Introduction: Planning agents in a Social World", in his *Faces of Intention: Selected Essays on Intention and Agency*, Cambridge University Press, pp.1–15
- (1999b), "Shared Cooperative Activity", in his *Faces of Intention: Selected Essays on Intention and Agency*, Cambridge University Press, pp.93–108
- (1999c), "Shared Intention", in his *Faces of Intention: Selected Essays on Intention and Agency*, Cambridge University Press, pp.109–130
- (1999d), "I Intend That We J", in his *Faces of Intention: Selected Essays on Intention and Agency*, Cambridge University Press, pp.142–164
- (1999e), "Davidson's Theory of Intention", in his *Faces of Intention: Selected Essays on Intention and Agency*, Cambridge University Press, pp.209–225
- (1999f), *Intention, Plans, and Practical Reason*, *The David Hulme Series*, CSLI Publications. Originally Published (1987) Harvard University Press
- (2000), "Reflection, Planning and Temporally Extended agency", *The Philosophical Review*, Vol. 109, No. 1 (Jan.), pp. 35 – 61 Also reprinted in Bratman 2007, Chapter 2
- (2004), "Shared Valuing and Frameworks for Practical Reasoning" in *Reason and Value*, ed. Wallace, Pettit, Scheffler, and Smith, Oxford: Oxford University Press, pp.1–27
- (2007), *Structures of Agency: Essays*, Oxford

- (2009a), “Intention, Practical Rationality and Self-Governance*”, *Ethics*, pp. 411–443, University of Chicago
- (2009b), “Shared Agency”, *Philosophy of the Social Sciences: Philosophical Theory and Scientific Practice*, Ed. Mantzavinos, C., Cambridge University Press pp. 41–59
- (*forthcoming*), *A Theory of Shared Agency*
[published after the submission of this thesis as (2013) *Shared Agency: A planning theory of acting together*, Oxford University Press
- Broome, J.** (1999), “Normative Requirements”, *Ratio (new series)*, Vol. XII, Oxford: Blackwell, pp. 398–419
- Buss, S.** (2008) "Personal Autonomy", *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), <http://plato.stanford.edu/archives/fall2008/entries/personal-autonomy/>
- Chisholm, R. M.** (1964), 'The Descriptive Element in the Concept of Action', *The Journal of Philosophy*, Vol. 61, No. 20, American Philosophical Association Eastern Division Sixty-First Annual Meeting, pp. 613–625
- CIVIII (8th Conference on Collective Intentionality) Website**, (2012)
<https://sites.google.com/site/collintviii/> [accessed 16/9/2013]
- Constant, B.** (2003), *Principles of Politics Applicable to all Governments*, trans. Dennis O’Keeffe, ed. Etienne Hofmann, Indianapolis: Liberty Fund <http://oll.libertyfund.org/title/861> [accessed 16/9/2013]
- Copp, D.** (1979), ‘Collective actions and secondary actions’, *American Philosophical Quarterly*, Vol. 16, No. 3, Basil Blackwell
- Davidson, D.** (1971) “Agency”, in *Agent, Action and Reason*, Ed. Binkley, R. Bronaugh R. & Marras, A.
- (2001a), “Actions, Reasons and Causes”, in *Essays on Actions and Events*, Oxford: Oxford University Press, pp. 3 – 20. Reprinted from (1963) *Journal of Philosophy*, no. 60
- (2001b), “Intending”, in *Essays on Actions and Events*, Oxford: Oxford University Press, pp. 83 – 102. Reprinted from (1978), in *Philosophy of History and Action*, eds. Yovel, Y. & Reidel, D. The Magnes Press
- (2005), 'Truth Rehabilitated', in *Truth, Language and History*, Clarendon Press. First published in (2000) *Rorty and his Critics*, ed. R. Brandom, Blackwell
- Demeulenaere, P.** (2009), “Comment: Where is the Social?”, A 'comment' on Bratman's “Shared Agency” (same volume), in *Philosophy of the Social Sciences: Philosophical Theory and Scientific Practice*, Ed. Mantzavinos, C., Cambridge University Press pp. 41–59

- Dufoe, D. (1994), *Robinson Crusoe, A Norton Critical edition*, ed. Shinagel M., Norton & Company. Main text originally published (1719) as *The Life and Strange Surprising Adventures of Robinson Crusoe of York, Mariner*, Taylor
- Ebert, P. & Roberson, R. (2010), "Mountaineering and the value of self-sufficiency", *Climbing: Philosophy for everyone*, Wiley-Blackwell
- Eurostat. (2013), "Live births outside marriage", *Eurostat: Your Key to European Statistics, European Commission*,
http://epp.eurostat.ec.europa.eu/portal/page/portal/product_details/dataset?p_product_code=TPS00018 [accessed 16/9/2013]
- Frankfurt, H. (1971), "Freedom of the Will and the Concept of a Person", *The Journal of Philosophy*, Vol. 68, No. 1, pp. 5–20.
- (1997), "The Problem of Action", in A. Mele (ed.) *The Philosophy of action*, Oxford: Oxford University Press pp. 42–52
- (2006), "Taking ourselves seriously", in *Taking Ourselves Seriously & Getting It Right*, Ed. Satz, D. Stanford University Press
- Fricker, M. (2012), "Group Testimony: The Making of A Collective Good Informant", *Philosophy and Phenomenological Research*, Vol. 84, Issue 2, pp. 249–276
- Gilbert, M. (1992), *On Social Facts*, Princeton University Press. Reprinted with minor alterations from Hardback edition (1989), Routledge
- (1996a), "Introduction: Two Standpoints – The Personal and the Collective", in her *Living Together: Rationality, Sociality & Obligation*, Roman & Littlefield, pp.1–20
- (1996b), "Walking Together: A Paradigmatic Social Phenomenon", in her *Living Together: Rationality, Sociality & Obligation*, Roman & Littlefield, pp.177–194 Reprinted from (1990), *Midwest Studies in Philosophy*, Vol. xv, University of Notre Dame Press, pp.1–14
- (1996c), "More on Social Facts", in her *Living Together: Rationality, Sociality & Obligation*, Roman & Littlefield, pp.263 – 278 reprinted from (1991), "The Author responds: More on Social Facts", *Social Epistemology* 5, pp. 233–244
- (1996d), "Fusion: A sketch of the contractual model", in her *Living Together: Rationality, Sociality & Obligation*, Roman & Littlefield, pp. 215–228
- (1996e), "Group Membership and Political Obligation", in her *Living Together: Rationality, Sociality & Obligation*, Roman & Littlefield, pp.361–375 Reprinted from (1993) *The Monist*, Vol. 76, pp.119–131
- (1996f), "On Feeling Guilt for what one's group has done", in her *Living Together: Rationality, Sociality & Obligation*, Roman & Littlefield, pp.375–390
- (2000a), "Introduction: Sociality and Plural Subject Theory", in her *Sociality and Responsibility*, Roman & Littlefield pp.1–13

- (2000b), “What is it for us to intend?”, in her *Sociality and Responsibility*, Roman & Littlefield pp.14–37. Originally printed in (1997) *Contemporary Action Theory*, ed. Holmstrom-Hintikka G. & Tuomela, R. Vol.2, Kluwer Academic Publishers
- (2000c), “The Idea of Collective Guilt”, in her *Sociality and Responsibility*, Roman & Littlefield, pp.141–153
- (2000d), “Reconsidering the 'Actual Contract' Theory of Political Obligation”, in her *Sociality and Responsibility*, Roman & Littlefield pp.97 – 122. Reprinted (with minor revisions) from (1999), *Ethics*, Vol. 109, University of Chicago Press, pp.236–260
- (2001), “Social Rules as Plural Subject Phenomena”, in *On the Nature of Social and Institutional Reality*, eds. Lagerspetz, E., Ikäheimo, H. & Kotkavirta, J. SoPhi: University of Jyväskylä
- (2006), *A Theory of Political Obligation*, Oxford University Press
- (2007), “Collective Intentions, Commitment, and Collective Action Problems”, *Rationality and Commitment*, ed. Peter, F. & Schmid H. P., Oxford University Press
- (2008), “Two Approaches to Shared Intention: An Essay in the Philosophy of Social Phenomena”, *Analyse & Kritik*, Vol. 30, pp.483–514
- (2009), “Shared intention and personal intentions”, *Philosophy Studies*, Vol. 144, pp.167–187
- Gopal, P.** (2011), 'Trafalgar has much in common with Tahrir', *Comment is Free, The Guardian*, <http://www.guardian.co.uk/commentisfree/2011/mar/25/trafalgar-tahrir-protests-middle-east> [accessed 16/9/2013]
- Goldman, A.** (2010), “Why Social Epistemology Is Real Epistemology,” in A. Haddock, A. Millar, and D. Pritchard, eds., *Social Epistemology*, Oxford University Press
- Graham, K.** (2007), *Practical Reasoning in a Social World*, Cambridge: Cambridge University Press
- Gruner, R.** (1976), ‘On the actions of social groups’, *Inquiry*, Ed. Hannay, A. Vol. 19, Universitetsforlaget
- Hollis, M.** (1998), *Trust Within Reason*, Cambridge University Press
- Hudson, V. M. & Boer, A. M.** (2005), *Bare Branches: The Security Implications of Asia's Surplus Male Population*, Cambridge, Mass.: The MIT Press
- Hulme, D.** (1739), *Treatise on Human Nature*.
- Jenkins, S.** (2011), 'Britain has long been a poor venue for protest – Saturday won't change this', *Comment is Free, The Guardian*, <http://www.guardian.co.uk/commentisfree/2011/mar/24/protest-trafalgar-square-tahrir-square> [accessed 16/9/2013]

- Keeley, M. (1981), 'Organizations as Non-persons', *Journal of Value Inquiry*, Vol. 15, pp.149–155
- Kisolo-Ssonko, J. (2013), “How where I stand constrains where we stand”, *Institutions, Emotions, and Group Agents, Studies in the Philosophy of Sociality*, Springer, Dordrecht
- (2012), “Love, Plural Subjects & Normative Constraint”, *Phenomenology and Mind: The Online Journal of the Research Centre in Phenomenology and Sciences of the Person*, Vol. 3 IUSSpress
- Kolodny, N. (2005) “Why Be Rational?” *Mind*, Vol. 114, pp.509–63
- Kosgaard, C. M. (2009) *Self-Constitution: Agent, Identity, and Integrity*, Oxford University Press
- Krauss, L. M. (1995), *The Physics of Star Trek*, Flamingo
- Kutz, C. (2000), “Acting Together”, *Philosophy and Phenomenological Research*, Vol. 61, No.1, pp.1–31, International Phenomenological Society
- Lewis, D. (1969), *Convention: A Philosophical Study*, Harvard University Press, Cambridge,
- List, C. & Petit, P. (2011), *Group Agency: The Possibility, Design, and Status of Corporate Agents*, Oxford University Press
- Locke, J. (1748), *An Enquiry Concerning Human Understanding*
- Ludwick, K. (2007), “The Argument from Normative Autonomy for Collective Agents”, *Journal of Social Philosophy*, Vol. 38, No. 3, pp.410–427.
- (2013), “Proxy Agency in Collective Action”, *Noûs*, Wiley Periodicals, Inc
- Marx, K. (2000), *Karl Marx: Selected Writings*, (Compiled by McLellan, D.), Oxford
- Meijers, A. W. M. (2003), “Can Collective Intentionality Be Individualized?”, *American Journal of Economics and Sociology*, Vol. 62, No. 1, *Special Invited Issue: John Searle's Ideas about Social Reality: Extensions, Criticisms, and Reconstructions*, pp.167–183
- Mele R. & Moser P. (1997), “Intentional Action”, in *The Philosophy of Action*, Oxford Readings in Philosophy, ed. Mele R. Oxford
- McKinlay, T. (2006), “Wrong to let climber die, says Sir Edmund”, *The New Zealand Herald*, http://www.nzherald.co.nz/nz/news/article.cfm?c_id=1&objectid=10383276 [accessed 16/9/2013]
- National Institutes of Health, (2012), *NIH Human Microbiome Project defines normal bacterial makeup of the body*, <http://www.nih.gov/news/health/jun2012/nhgri-13.htm> [accessed 16/9/2013]
- Nozick, R. (1995), “Erotic Love”, in *Philosophical Perspectives on Sex & Love*, ed. Stewart, R. Oxford University Press

- Pacherie, E.** (2012), “The Phenomenology of Joint Action: Self-Agency vs. Joint-Agency”, in Seemann (ed.), *Joint Attention: New Developments*, MIT press : Cambridge MA, pp.343–389
- Pettit, P.** (2007), “Constructing Sen on Commitment”, in *Rationality and Commitment*, ed. Peter, F. & Schmid H. P., Oxford University Press
- Okin, S. M.** (1989), *Justice, Gender and the Family*, Basic Books
- Raz, J.** (2005), “Instrumental Rationality: A Reprise,” *Journal of Ethics and Social Philosophy*, Vol. 19, pp. 1–19
- Roth, A. S.** (2003), “Practical intersubjectivity”, in *Socializing Metaphysics: The Nature of Social Reality*, Ed. Schmitt, F., Rowman & Littlefield
- Rousseau, J.** (2004), *The Social Contract*, Penguin. First published (1762)
- Schmid, H. B.** (2007), “Can brains in a vat think as a team?”, *Philosophical Explorations: An International Journal for the Philosophy of Mind and Action*, Vol. 6, Issue 3, pp.201–217
- (2009), *Plural Action: Essays in Philosophy and Social Science*, Springer
- Schmitt, F.** (2003), “Socializing Metaphysics: An Introduction”, in *Socializing Metaphysics: The Nature of Social Reality*, Ed. Schmitt, F. Rowman & Littlefield
- Searle, J. R.** (1996), *The Construction of Social Reality*, Penguin. First published hardback (1995)
- (2001), 'Social Ontology and The Philosophy of Society', in *On the Nature of Social and Institutional Reality*, eds. Lagerspetz, E. Ikäheimo, H. & Kotkavirta, J. SoPhi, University of Jyväskylä
- (2010), *Making the Social World*, Oxford
- Smith, T.** (2005), “What is the Hallé?”, *Philosophical Papers*, Vol. 34, No. 1, pp.75–109
- (2007), “A Theory of Political Obligation: Membership, Commitment, and the Bonds of Society”; Book Review, *Mind*, Vol. 116
- (2011), “Playing One’s Part”, *Review of Philosophy and Psychology*, Vol. 2, *Special edition: Joint Action; What is Shared?*, Springer, pp. 213–244
- Stern, R.** (1999), “Introduction”, in *Transcendental Arguments*, Ed. Stern, R. Oxford
- Sugden, R.** (1993), “Thinking as a team: Towards an explanation of nonselfish behaviour”, *Social Philosophy & Policy*, Vol. 10, no. 1, pp.69–89.
- Tagliafico, D.** (2012), “The Paradox of Government: Explaining the Life and Death of a State”, *Phenomenology and Mind: The Online Journal of the Research Centre in Phenomenology and Sciences of the Person*, Vol. 2, IUSSpress

- Taylor, C. (1985a) "Atomism", in *Philosophy and the Human Sciences: Philosophy Papers 2*, Cambridge University Press
- (1985b) "Interpretation and the science of man", in *Philosophy and the Human Sciences: Philosophy Papers 2*, Cambridge University Press
- Thatcher, M. (1987), "Interview for Woman's Own ("no such thing as society")", *Thatcher MSS (digital collection)*, <http://www.margaretthatcher.org/document/106689> [accessed 16/9/2013]
- Tollefsen, D. (2004), 'Collective Intentionality', *The Internet Encyclopedia of Philosophy*, ISSN 2161-0002, <http://www.iep.utm.edu/coll-int/> [accessed 16/9/2013]
- Toumela, R. (2007), *The Philosophy of Sociality: A Shared Point of View*, Oxford University Press
- Velleman, D. (1997), "How to share an intention", *Philosophy and Phenomenological Research*, Vol. 57, No. 1, pp. 29-50
- (2000), *The Possibility of Practical Reason*, Oxford: Oxford University Press. Republished online by author, <http://homepages.nyu.edu/~dv26/Possibility/> [no longer online]
- Verbeek, B. (2007), "Rational self-commitment", *Rationality and Commitment*, ed. Peter, F. & Schmid H. P., Oxford University Press, pp.150-174
- Vincent, A. (1989), "Can groups be persons?", *The Review of Metaphysics*, Vol. XLII
- Quinton, A. (1975/1976), 'Social Objects', *Proceedings of the Aristotelian Society*, Vol. LXXVI, Compton Russell Ltd, pp.1 – 27
- Ullman, J. R. (1956), *Man of Everest: The Autobiography of Tenzing*, Transcribed from conversations with Norgay, T. The Reprint Society London. First published (1955), Harrap & Co.
- Webber, M. (1907), "R. Stammlers "Überwindung" der materialischen Geschichtsauffassung", in his *Gesammelte Aufsätze zur Wissenschaftslehre*, Tübingen, J.C.B. Mohr
- Westlund, A. (2006), "Love and the Sharing of Ends", *Twenty-First Century Papers: On-line Working Papers from the Center for 21st Century Studies*, University of Wisconsin, Milwaukee
- (2008), "The Reunion of Marriage", *The Monist*, Vol. 91, No. 3 & 4, Illinois
- (2009), "Deciding Together", *Philosophers' Imprint*, Vol. 9, No. 10, MPublishing, University of Michigan Library