Chapter 1
Plural Action

Concepts and Problems

I call *plural* those actions that require the participation of at least two individuals (sociality condition) acting in pursuit of one and the same goal (plurality condition). Examples of plural actions are: going for a walk together, jointly writing a paper, or playing a symphony. Even though plural actions abound in our lives, they have been somewhat neglected in philosophical analysis. Part of the reason for this is that plural actions do not seem to fit easily into our standard philosophical conception of agency. Whereas any singular action can be attributed to a single individual agent – the only kind of agent standard theory of action knows of – plural actions seem to require a different kind of agent (the plural agent problem). In the first section of this chapter, I shall use the intuitive idea that one cannot intend what one takes oneself to be unable to perform to approach the plural agent problem, and situate plural actions within a taxonomy of action types (§1). I then turn from action theory to common sense. In contrast to action theory, common sense seems to have no difficulty whatsoever in coming up with suitable agents for plural actions. There are at least three different common sense solutions to the plural agent problem: plural actions are either attributed to *collective agents* (such as in the case of Parliament's passing a law), to *powerful individuals* (such as in the case of Caesar's defeating the Helvetii), or to a plurality of individuals jointly intending an action (such as in the case of a bunch of friends going for a walk together). These three replies correspond to three different models (or perhaps types) of plural agency. I propose to call them the collective agent model, the influence model and the teamwork model, respectively, and I shall argue that the teamwork model is the most fundamental of these.

I shall then turn to a somewhat more detailed discussion of each of these models, and examine the reasons why they are met with so much reservation (or even resistance) from the side of action theory. As to the *collective agent view* (§2), it is obvious that many authors are still reluctant to ascribe intentions and actions to collectives. This reluctance found its classical expression in methodological individualism of the Weberian kind. I examine a reason Max Weber might have had for not admitting collective agents to the basic level of intentional interpretation, and I conclude that he believed collective agency to be incompatible with what I suggest to call *individual intentional autonomy*. I propose this label for the view that each individual is an *agent* of his or her own, i.e. that his or her behavior should

H.B. Schmid, *Plural Action: Essays in Philosophy and Social Science*, Contributions to Phenomenology 58, © Springer Science+Business Media B.V. 2009

be interpreted as *his or her own actions*. Looking at current conceptions of collective agency, I argue that, contrary to what Weber seems to have thought, collective agency is compatible with individual intentional autonomy, so there is no reason not to accommodate a robust conception of collective agency in action theory.

The following section (§3) examines the *power or influence model* of plural agency, according to which the leaders and authority figures have a claim on the ownership of a plural action. Its main intuitive problem seems to be that, by ascribing the plural action in question to one single individual, it bypasses the other participants' own individual agency. According to this view, all participating individuals have a claim on their own individual *contribution* to the plural action, and no individual has a claim on more than that, so that the plural action as a whole cannot be attributed to one individual, however powerful he or she might be. It might seem that this view is a direct consequence of the *intentional autonomy* of the participating individuals, but I shall argue that this is mistaken, and that there is a *further* and *more problematic claim* involved in this view, which I shall call *individual motivational autarky*.

The assumption of motivational autarky is that each individual acts on a motivational agenda of his or her own, i.e. that the interpretation of each individual's behavior has to bottom out in *his or her own volitions or pro-attitudes* (rather than in some other individual's volitions or pro-attitudes). I argue that intentional autonomy does not imply motivational autarky, and that "non-autarkical" behavior might actually play an important role in many cases of plural action (this issue is further pursued in Chapter 8 below). I conclude this section with a discussion of why the autonomy assumption and the autarky assumption have always been mixed up. My thesis is that while intentional autonomy is an universal feature of human agency, and indeed an essential feature of what it means to be an agent, motivational autarky is something very different: a very strong *normative ideal* in our particular culture.

Turning finally to the *most basic form* of plural action in §4, I argue that the main problem *teamwork models of plural agency* have to cope with is that of reconciling the unity of action with the plurality of agents. This is precisely what, in the current literature, the concept of *collective intentionality* is invoked for: many individuals can intend and claim shared ownership of a plural action insofar as they share the respective intention. Most philosophers of collective intentionality, however, are very reluctant to admit a *straightforward understanding* of the sharedness of intentional states. The predominant view is a distributive reading of collective intentionality, according to which individuals cannot *literally* share an intention, and that each individual has his or her own intention when they intend to do something together.

Behind this view lies what I propose to label *intentional individualism:* the view that all intentionality is *some individual's*. I shall argue that this *distributive* reading leads to circular conceptions of collective intentionality. I believe that the reason why most philosophers of collective intentionality endorse a distributive conception is that they believe that intentional individualism is implied in individual intentional autonomy. I shall argue that it is not, and that the theory of plural agency will make great headway by dropping intentional individualism, and endorsing a stronger conception of *intentional commonality*.

The concluding §5 wraps up the line of argument developed in this chapter, and closes with an observation on the occasion of the first centenary (2008) of the term "methodological individualism".

§1 The Plural Agent Problem

In the earliest stages of the Apollo Program, John F. Kennedy once went to Florida to visit Cape Canaveral. On his tour through the facilities, he asked a technician what his task was. The anecdote has the man giving the following answer: "To land a man on the moon and returning him safely to the earth before this decade is out, Mr. President!"

What is so unusual here that people bother to retell the story? The clue, it seems, is the huge gap between what the worker claims to be his task, on the one hand, and what he is actually able to do on the other. If the man is really *serious* about what he claims to be his task, sending a man to the moon is what he in fact *intends* to do; if this is the case, however, he expects *far too much* of himself. Sending a man on the moon simply exceeds a single technician's possibilities.

Thus this anecdote sheds some light on how we *normally* think about the relation between intentions and abilities. What seems to be at stake here is something we might call the Principle of *Intentional Self-Confidence*. This principle puts some constraints on the range of things one can intend to do. It states that an agent's intentions must be *in tune* with what he or she takes to be *possible*, given her abilities and the opportunities at hand.[1] In the briefest (negative) version, the principle reads as follows:

(A) The Principle of Intentional Self-Confidence: One cannot intend to do what one takes oneself to be unable to carry out.[2]

This needs some explanation. First, the principle of intentional self-confidence is perfectly compatible with the fact that "intend" does not imply "can". People might well intend the *objectively impossible* just as long as they don't *take it* to be impossible. Also, the principle does not rule out the possibility of certain forms of aiming at the *subjectively impossible*.[3] And naturally, the principle does not entail

[1] For an analysis of the concept of ability cf. Kenny 1976. Useful ideas – especially the distinction between the subjective and the objective components of ability – can be found in Löwenstein 1911.

[2] For an early version of this principle cf. Baier (1970). The main difference between the usual way of putting the limitation claim and mine is this. Taking oneself to be able to do x is usually assumed to be a matter of *belief*, i.e. a cognitive intentional state. Self-confidence, by contrast, is an *emotion*, i.e. an affective intentional state.

[3] In his "Impossible Doings" (1992), as well as in some later papers on the topic, Kirk Ludwig contested that claim. Ludwig discusses the following example. P assumes (with certainty) that the battery of his car is dead. Upon another person's request, he turns the ignition key. Contrary to what he expects, the engine starts. Ludwig claims that it would be wrong to say that P started the engine unintentionally. I agree with Ludwig that there are some cases of trying the subjectively impossible

the claim that, in order to intend to do A, one has to believe that one *will* A.[4] In his *change in view*, Gilbert Harman (1986) discusses the example of a sniper who takes himself to be a terrible marksman, but still intends to kill the ambassador, in spite of his self-doubts. This is not in conflict with the principle. Sometimes, agents have high intentional *ambitions*, and seriously intend to do things they perfectly well *know* they might be unable to achieve after all (indeed, such ambitious intentions are common at the start of any project, such as the intention to write a book on plural agency). In other words, one does not have to be *certain* to be able to perform what one intends to do. Life would be boring indeed if we limited our intentions to things we perfectly well know we can do. This is to say that intentional self-confidence might well be *minimal*. Even the faintest hope of achievement is enough. All the principle states is that intentional self-confidence cannot be *zero*, for one cannot intend to do what one is perfectly *sure* of being unable to perform. If I *know* that the restaurant opens only at 6 p.m., I cannot *intend* to have lunch there at noon. If I still choose to go there, and if this isn't a case of conflict between what I know my abilities to be in my head and what I feel able to do in my guts, my intention is a different one: to act *as if* I didn't know the restaurant's opening hours, to knock on the closed door as if in surprise, or any other aim along these lines.

I have already emphasized that the principle of intentional self-confidence is not in conflict with the possibility that one might be *mistaken* in what one takes oneself to be able to perform. One can always *misjudge* one's forces and abilities, and expect too much or too little of oneself. This is perfectly possible, but to the degree that this happens *systematically* and *under normal circumstances*, it renders intentional self-confidence *irrational*. Intentional Self-Confidence is *rational* to the extent that under normal circumstances one's intentions are *in balance* with one's *actual* forces and abilities.

Let's call this the *rationality specification* of the principle. It allows for two directions of imbalance: one can either overrate or underrate one's forces and abilities. In other words, intentional self-confidence can be irrational in two ways – for lack of better terms, let's simply label them "objective" and "subjective", respectively. Intentional self-confidence is *objectively irrational* if one intends to do a thing which one is *generally* and *under normal circumstances* incapable of performing. (It is always possible to fail at a task one takes oneself to be perfectly capable of performing. If this happens *by chance* or *due to unusual and unforeseeable circumstances*, this does not render one's intentional self-confidence irrational.) Conversely, intentional self-confidence is *subjectively irrational* if one fails to form an *intention* to do something one wishes to be done *for the sole reason* that one takes oneself to be

---

where "trying" does not function as a proper action term. I argue, however, that in such cases, the agent must take himself to have a chance at success, however minimal, which might be in conflict with his conscious assessment of the situation. Sometimes an agent's intentional self-confidence is not in tune with his or her beliefs concerning his or her ability. If this is true, Ludwig's point does not prove that it is not the case that people cannot intend to do what they take themselves to be unable to do. The question is how to understand the "taking": insofar as it is belief, Ludwig is right; insofar as it is self-confidence, he is not.

[4] This claim is often ascribed to Paul Grice (1971) and J. David Velleman (1989).

unable to do it, when it is actually well within one's forces and abilities. This second form of irrationality consists in an *understatement* of one's forces and abilities.

Again: neither of these two kinds of irrationality is incompatible with the principle as such. The principle of intentional self-confidence is a *conceptual* principle. It is part of how we use the term "to intend".[5] As such, however, the principle does not say that intentional self-confidence is always *rational*. Rationality, in other words, is a *normative standard* for intentional self-confidence, not a *conceptual requirement*, as those pervasive cases of both varieties of *irrational* intentional self-confidence show rather clearly.

The fact that not all intentional self-confidence is rational seems to be precisely what makes the NASA technician's reply in the above anecdote funny. The man's reply takes intentional self-confidence to its *objectively irrational extreme*. By taking his task to be to send a man to the moon, he takes himself to *be able* to do such a thing, which he is clearly not, because what single workers can do is limited to such things as wiping factory floors, assembling parts of rocket stages, etc. Thus the worker grossly and grotesquely overestimates his forces.

I will not delve any deeper into an analysis of intentional self-confidence here, but rather use the principle as a guide to a quite different issue at stake in the anecdote. If sending a man to the moon couldn't rationally be *a single worker's intention*, because what single workers can do is limited to much more moderate tasks such as assembling rocket parts, the question arises: *whose task was it, then?* Who could *rationally* ever be so *self-confident* as to intend to land a man on the moon? This is the question I wish to address in the following. Before looking at possible candidates for this role, however, I should first make sure that this question does indeed make sense. It does so only if intention is conceived of in action-referential terms, which some authors claim is not necessary. And it does so only if something like the moon expedition can be described as *one* action. Is this true, and if so: to what particular *kind* of action do such things belong?

First a remark on the question of whether or not intention should be conceived of in action-referential terms. Intention is action-referential insofar as it is an intention *to* A. This seems the most natural way of putting intention, but this places tight restrictions on the possible objects of intention. A needs to be an action, and as the only actions one can intend *directly* seem to be one's own, it appears that the only objects of intention are *one's own actions*. In action theory, there is a tendency to claim that possible objects of intention extend beyond one's own actions (cf., e.g., Bratman 1987; Vermazen 1993). These authors claim that the objects of intentions are *propositions* rather than actions. Thus intention should be conceived of in *propositional* rather in *action-referential* terms. In their view, intention should be seen as intention *that p*, where p can be any state of affairs, and even an action whose subject is not the subject of the intention. This considerably widens the scope of intention. The question is: should the technician in the above anecdote have said "I intend *that* a man be landed on the moon and returned safely to earth"?

[5] Needless to say, the constraints articulated in the principle of intentional self-confidence apply only to *intentions*, not to wishes and other intentional states.

Wilfrid Sellars (1992: 183ff.) has claimed that propositional intentions are expressions of practical commitments only by virtue of their conceptual tie to action referential intentions. Sellars says that the intention that X, when made explicit, is the *intention to do whatever is necessary to make it the case that X*, which is action self-referential, because again, one has to perform the doing oneself. I think he is right. But it is completely sufficient for present purposes to accept action referential intention as *one important kind* of intention, and that the action-referential mode of expressing intention should not be abandoned completely, even though it might not be the only way of thinking about intention.

What about the second question concerning the unity of action? Let me start by mentioning four fairly uncontroversial features of the concept of action. First, for there to be an action there has to be some kind of agent, i.e. somebody to whom the action is attributed, and who can be held responsible for its consequences according to our normative practices. Second, action requires some kind of *behavior*, of which the agent is in a certain degree of control (typically consisting of the agent's own bodily movements). Third, some *goal* is needed, i.e. something the agent *wants*, a state of affairs towards which the agent has some kind of *pro-attitude* or "desire" in the wide sense of the word. In the context of action, goals are conceptually tied to *intentions*. If a complex of behavior is taken to be an action, it is assumed that the agent is in fact *trying to achieve* his or her goal, i.e. that the goal is the condition of satisfaction of an intention. And fourth, there has to be some connection between the agent's goal and the complex of behavior in question. The behavior has to be *minimally rational*, i.e. the agent has to show at least some minimal degree of *concern* about the behavior's being suited as a means to the end (however successful or unsuccessful she might be at this task).

It goes without saying that although these characteristics may be necessary conditions for actions, they are certainly not sufficient ones. We do not have to delve any deeper into action theory here, however, in order to answer the simple question: was man's travelling to the moon an action? If we leave aside for the moment the open question concerning the agent, it seems that the moon expedition meets all conditions. As to the goal-directedness and rationality of behavior, it seems to even be a *paradigmatic* case of an action. There clearly was a goal, and not only was the goal obviously *intended*, but the achievement of the goal was also permanently monitored, with a constant effort to choose suitable means to the end. Thus it seems clear: the moon expedition *was* an action – *if* the one open question can be answered: *if* an agent can be identified, i.e. *if* it is possible to answer the question *whose* action it was.

Before we come to that, let's just assume for the moment that a plausible answer to the open question *had* been given, and have a closer look at the *type* of action to which something like the moon expedition would belong. I propose the term *plural action* for the kind of action in question. Plural actions are *social actions*. I call social those actions that require the participation of more than one individual (I label this the *sociality condition*). Let's call the class of non-social actions *solitary* actions (it is still controversial whether or not this class contains any elements, i.e. whether or not a hypothetical lifetime Robinson Crusoe could be an agent). There

are two ways in which sociality might be required: either *logically* (as in the classic case of marrying, which as a matter of conceptual necessity takes at least two), or *contingently* (as in the case of building a house, which one could do all by oneself if only one was a little stronger). Accordingly, there are two kinds of social actions. Sending a man to the moon belongs to the class of contingently social actions. I'm not concerned with *this* distinction here, however, but with the one between *singular* and *plural* actions, which is independent of the distinction between logically and contingently social actions.

The distinction between singular and plural actions is a matter of the *goals* pursued in each case. Singular actions are social actions in which the participating individuals pursue *different goals*. By way of example, consider the case of my taking a plane back home for Christmas, which presupposes, among other things, some mechanic's activity, but my spending Christmas at home and his getting the jet engines going are quite different goals. In contrast to examples such as this, *plural actions* are social actions in which the individuals pursue the *same goal*.

To give an example: our playing a duet together requires that we have this goal in common. If you simply aim at performing your part (as long as I perform mine), and *vice versa* – if our goals are, in other words, *different* goals – we may make our way through the score, but we certainly won't be playing a duet. If our playing is to be a duet, our goal has to be the *same*. I call this the *plurality condition*. If Saturn V had been produced with the sole aim of selling it to the highest bidder, the production of Saturn V would not have been part of the plural action of sending a man to the moon. To be as explicit on this point as possible: plural actions require more than the participants' having *similar* goals. Just because each individual in a group has a similar or the same *type* of goal, or even goals with the same *content*, does not make the activity in question a plural action. The goal must ultimately be *one and the same* goal. In brief, the main characteristics of plural actions are these: *many* participants, *one* goal (Fig. 1.1).

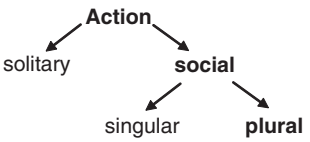

Fig. 1.1 Taxonomy of action types

Plural actions have long been rather shamefully neglected in action theory. Only in the course of the last 2 decades has the phenomenon started to attract any attention. In the meantime, however, a small, but rapidly growing debate on the structure of plural action has developed. It is characteristic of much of this debate that *small-scale examples* are used to discuss the structure of plural action. Activities such as going for a walk together (Gilbert 1996), jointly operating a water pump (Bratman 1999), preparing a Sauce Hollandaise by one pouring the oil and one stirring the sauce (Searle 1990), or pushing a broken-down car together (Tuomela 1995) serve as illustrations of the phenomenon. By contrast to this, my choice in this chapter is a large-scale example; apart from wars and military expeditions, the

Apollo program was probably among the most extended plural actions in the entire history of mankind (I will turn to smallest-scale examples below).

I have chosen this example because the large scale helps to illustrate what I see as the *crucial problem in the theory of plural action*. It is this: If the principle of intentional self-confidence is valid, plural actions require a particular *kind of agent*: one that can *rationally* take him- or herself to be able to perform an action which can only be carried out with the joint efforts of many? But *who can possibly fit that bill? What are plural agents?* Are there any plausible candidates that conform to the principle of intentional self-confidence *objectively rationally* (in the sense defined above) with regard to *plural actions?*

Common sense offers no less than three types of candidates: in everyday parlance, we routinely ascribe plural actions to *collective agents*, to *influential individuals* (or leaders), and to *jointly intending individuals*. These three types correspond to three commonsensical models of plural agency: the collective agent model, the influence model, and the teamwork model. In the following sections I will characterize each of these models in turn, and make some comments on why these intuitive notions have not usually been well received in action theory (to say the least), and on the main obstacle standing in the way of a deeper and more adequate analysis of plural agency.

§2 Collective Agents and Individual Autonomy

According to what I propose to call the collective agent view, plural actions may be intended (and indeed performed) by *collectives*. The plausibility of this view can easily be illustrated with our example. Just remember Kennedy's choice of words when he announced the start of the Apollo program in May 1961: "I believe that *this nation* should *commit itself to achieving the goal*, before this decade is out, of landing a man on the moon." So maybe the nation really did it after all, or perhaps NASA did it. Another, somewhat less plausible candidate for that role would be *mankind* – the collective Neil Armstrong invoked when he was setting his foot on the moon. The advantage of this model is that it *avoids* the problem of rational intentional self-confidence by invoking an agent that matches the size of the task in question. The agent is not an individual agent, but a super-agent over and above the heads of the participating individuals. Such agents, it seems, need not worry about expecting too much of themselves when they form an intention to carry out a plural action. Given their size, they can be rationally self-confident in their ability to do such things. The only question is: do such agents really exist?

Most prominently, the forefathers and founding fathers of *methodological individualism* rejected the assumption of collective agents. Thus Max Weber famously stated that the only agents which social science recognizes are individuals.[6]

[6] Weber articulates this central precept of his methodology in the following way. When discussing social phenomena, we often talk about various "social collectivities, such as states, associations, business corporations, foundations, as if they were individual persons" (Weber [1921] 1980: 13).

What is the reason for Weber's view? It has repeatedly been pointed out that Weber's rejection of collective agency should be seen in the context of his commitment to the method of *intentional interpretation* (Heath 2005). And this is indeed what Weber says: *only individuals* "can be treated as agents in a course of subjectively understandable action" (Weber [1921] 1980: 6). Collectives, Weber seems to think, are simply *not suited* as objects for intentional interpretation. Since, in everyday life, we often seem to have no difficulty whatsoever in ascribing intentions and actions to collectives, however, one might wonder what reason Weber might have had for this claim.

Unfortunately, Weber does not expand on this any further. But one can think of a whole series of arguments for this view. First, it is well known that Weber defines action by meaning (*Sinn*) and behavior, both of which – in Weber's view at least – seem to be essentially individual. Another reason might be that, within the framework of Weber's analysis, action is supposed to play the role of the *explanans*, with the collectivity being the *explananda*[7] – this naturally excludes plural action from the class of *explanantia*, because this kind of action seems to be the sort that already involves collectivity (cf. the *plurality condition* mentioned above).[8]

I certainly do not underestimate the role of any of these arguments for Weber, but I think that his *basic concern* is yet another one. The worry is this. If we were to treat collectives as agents, individual agency would be somehow conceptually *compromised* or *impaired*. The point of departure of *this line* is Weber's firm commitment to the view that individual behavior *is* the proper object of intentional interpretation. Individuals are *agents*. Their behavior instantiates *their actions*. This commitment to the agency of individual persons, Weber seems to think, is *incompatible* with the assumption that there are any agents other than single individuals, and in particular with the assumptions that there are collective agents. For if collectives were proper agents, the participating individuals would be nothing more than the mere *instruments* or executing organs of some collective will, and would not be the proper agents behind their behavior anymore. Thus it seems that, insofar as individuals are to *be treated as agents* in the interpretation of social phenomena, collectives simply cannot be so treated. Admittedly, Weber never explicitly says so, but I believe that

---

While Weber does not take issue with any such everyday talk at all (he even admits that for other epistemic purposes, the assumption of collective agents might indeed be quite "useful"), he stoutly opposes its use in scientific interpretation.

[7] The Weberian project is to explain collectivities as "consequences and organizations of individual actions" (Weber [1921] 1980: 13).

[8] Along this line, Weber's is simply a reductivist view: while it is not necessary to use collectivity concepts to describe individual agency, all collectivity concepts can be translated into descriptions of aggregates of individual actions. The only reason why social science cannot fully *do away* with collectivity concepts on the lowest level is, according to Weber, that collectivities are part of the *content* of individual intentional mental states. People happen to *believe* that there are collectives, and they act on this belief. Insofar as a certain type of individual action is the object of social science, collectivity concepts cannot completely be ignored. But clearly, the order of explanation goes from the individual to the collective. There are collectives, because people *think* there are collectives, and not the other way around.

this is the worry in the back of his mind that leads him to oppose the group agent assumption.[9] Because individuals *are* to be interpreted as agents, there can be *nothing but* individuals on the most basic level of intentional explanation. For any other type of agency would *displace* the participating individual's agency. Even though his conclusion might be controversial (we shall come back to this shortly), I think that Weber's premise, i.e. his firm commitment to the individuals' own agency, is basically right. In the current debate, the commitment to individual agency seems to be so universally accepted that it is not even identified as such. Even those who reject the *conclusions* drawn by Weber, and believe that plural agents are an important feature of the very basic structure of social reality, seem to take it for granted that this is compatible with a robust notion of individual agency. I shall call this commitment, which I believe to be at the heart of methodological individualism, the principle of *individual intentional autonomy*. In its shortest formulation, the assumption is the following:

(B) Individual Intentional Autonomy: Under normal circumstances, each individual's behavior instantiates *his or her own actions*.

"Normal circumstances" exclude such cases as mere reflex behavior, which does not instantiate any action at all. Admittedly, the use of the term "autonomy" is somewhat unusual in this context. In the current debate, autonomy is normally taken to involve such highly complex and elaborate structures as self-transcendence, motivational hierarchies, and reflective self-management (cf. e.g. Bratman 2007: 162ff., 195ff.). None of these is presupposed or involved in what I call individual intentional autonomy, even though I dare to claim that, conversely, intentional autonomy in the sense defined here is one important presupposition of all of these more ambitious and richer philosophical concepts of autonomy. In other words, my use of the term autonomy underlies any of the current controversies revolving around this concept. Intentional autonomy refers to a very basic and elementary way in which individuals are responsible for their behavior as agents, in which their behavior can be ascribed to them *as actions*, and in which they can – to introduce a metaphor which I will use repeatedly below – claim *ownership* of their action.

For further clarification of the term intentional autonomy, I introduce its equally neologistic counterpart, *intentional heteronomy*. An intentionally *heteronomous* individual's behavior, were it to exist – which I doubt – would instantiate none of the respective individual's *own* actions, but rather that of *another agent*. In other words, intentionally heteronomous individuals would have to be taken as behaving on another agent's *remote control*, as it were. They would in fact be what we might call *intentional zombies*, to add yet another sort of zombie to the philosophical literature. In contrast to this, the principle of individual intentional autonomy

[9] It is not easy to feel the threat of collectivism now, and perhaps collectivism was never much more than a specter that haunted this debate, but I assume that, in this role, it has been quite effective. Even Ludwig Gumplowicz (1928), who to my knowledge went farthest among early social scientist in asserting the independence of the intentionality and agency of collectives from the intentionality of individuals, asserted that any explanation of social phenomena ultimately bottoms out in motivations for individual actions.

states that individuals are *not* intentional zombies because they do *not* behave on remote control. I assume that this claim is uncontroversial. Intentional zombies abound in philosophical thought experiments (e.g. Mele 2003), in some radical interpretations of the possible effects of hypnotism, in self-reports by schizophrenic patients (Spence 2001; Marcel 2003), and in American Sci-Fi. It seems generally accepted, however, that the everyday social world is not populated by intentional zombies.

So much for the defense of Weber's core claim. The critical question, however, is: is Weber right in assuming that individual intentional autonomy is incompatible with a solid conception of collective agency? I think that he is not, and I take it that this has been sufficiently established by the recently renewed interest in the agency of collectives. There are robust conceptions of collective agency on the market which are perfectly compatible with individual intentional autonomy. This is particularly obvious in Philip Pettit's recent work. In his analysis of the *discursive dilemma*, Pettit has developed the view that (certain types of) collectives can be interpreted as intentional subjects. He even ascribes to them some sort of personhood (Pettit 2002; cf. Rovane 1998), such that, groups can have "a mind of their own" (Pettit 2003).

In his analyses, Pettit's concern is with the *rational unity* of the groups' perspective, which under some circumstances, requires some measure of *discontinuity* with the participating individuals' own perspectives. The phenomenon to which Pettit draws our attention is that the rational unity of a group perspective sometimes requires that this perspective be *distinct* from that of any of the participating individuals'. These collectives are genuine agents. But it becomes more than clear in analyses such as Pettit's that, contrary to what I think Weber's worries were, this does not compromise or displace the agency of the individuals. In Pettit's conception of collective agency, plural subjecthood is solidly grounded in the volitions of the participating individuals. Groups have a sort of agency of their own based on the participating individuals' *insight* into the problems of aggregating individual decisions to collective decisions, and on the participating individuals' *choice* to get their collective act together in avoiding the pitfalls of the *discursive dilemma* and to act consistently and rationally *as a group*. Forming a collective agent does not compromise or displace, but rather *presupposes*, individual intentional autonomy.

Thus, contrary to a worry that still seems to be in the back of the mind of many action theorists, genuine collective agency does not compromise individual agency. There is no reason why action theory should treat conceptions of collective agency with so much reserve. This does not mean, however, that *any* kind of plural agency can be interpreted along the lines of the collective agent model. Collective agency is but one kind of plural agency. The extension of the collective agent model is limited to those cases where the distinction between the collective agent on the one hand and the participating individuals on the other has some intuitive plausibility because this is the way the participating individuals *themselves* interpret their situation. This is particularly true of Hobbesian *personae fictae* – but there is no *persona ficta* involved in smaller cases of plural actions such as going for a walk together. If the two of us go for a walk together, there are only two agents involved in the process – not one, and most certainly not three. Thus the collective

agent model provides a solution to the plural agent problem only in some selected cases. For other plural actions, we seem to need different plural agents. Yet there is an even farther-reaching reason for doubt as to the scope of collective agent-explanations of plural actions. It seems clear that for a collective agent to emerge from the agency of intentionally autonomous individuals, there has to be some sort of *agreement* between those individuals. Such agreements usually presuppose communication. Acts of communication, however, are plural actions. Thus it seems that collective agents presuppose plural agents which are not collective agents. Let us therefore turn to the remaining two common-sense replies to the question of plural agency.

§3 The Dogma of Motivational Autarky

The second type of answer to the "whodunit" question concerning plural agency is this: those individuals who were in control of the project (according to its institutional structure) did it. Let's again take the Apollo program as an example. While the statement "My task is to land a man on the moon" might sound rather silly coming from a simple technician's mouth, it doesn't nearly as much coming from, say, NASA's chief administrator in his leather armchair, or indeed from the President's own lips. It seems that such people's claims to rational self-confidence are simply much better substantiated than those of lower-ranking individuals. To put it as bluntly as possible: great people can do great things. In the memory of Kennedy's recently (2007) deceased court historian, I'm tempted to label this *second* model the *Arthur M. Schlesinger-view* of plural agency, but for the sake of brevity, let's stick to the label *influence model of plural agency*. Admittedly, this is a somewhat patriarchal notion, and, to say the least, it is not very popular in the current humanities and social sciences. Among its advantages, however, is the fact that it is deeply rooted in everyday talk. Behind the erection of the palace of Versailles was Louis XIV intention; it was Vasco da Gama who successfully searched for the sea passage to India, etc. etc. This view models plural agency very closely on the paradigm of singular agency, which makes plural agency look somewhat less unfamiliar. But therein lies the central problem of the model. In attributing plural actions to single individual agents, it makes it look like these leaders had performed their great deeds all by themselves.

Thus the model seems to suggest that leaders have many hands, feet and eyes, not just two of each. The agency of the other individuals involved in the process is simply bypassed, and their individual contribution remains completely unaccounted for in this view. This cannot be right, and it goes against the grain of a deep-seated "democratic" conviction in the theory of social action. This conviction has it that individuals – however powerful they might be – can be attributed only their *individual contribution* to plural actions, and not the plural action as a whole.

Here is a conjecture concerning the line of reasoning that might be behind this "democratic" view. Influential individuals might perform such actions as *giving*

orders, or *bringing others to make their goals their own*, or any such acts. But they cannot simply *do* what requires the joint forces of many to be done, because this would require that the intentionality of the leader extend *directly* to the behavior of the subordinates. The other participants would be quite literally reduced to the leader's hands, feet, and other limbs: the behavior (body) being the subordinate's, the intentionality (mind) being the leader's. The subordinates would then, it seems, not be seen as acting on *their own* intention, in the execution of their *own* plans (however conformist they might be), since the intentional explanation of their behavior would point to the leader's wishes and intentions rather to their own. This is at odds with the conviction that all participants in plural actions, not just the leaders, have to be interpreted as agents. Therefore, the influence model cannot *literally* be true.

This worry, which seems to be quite widely shared in the relevant literature, closely resembles the commitment to individual intentional autonomy. But it is essential to be as careful as possible here. As I shall argue, the critical claim goes one step beyond the claim that each participating individual's behavior has to be interpreted as his or her own action. There is a further claim involved in this line of argument. To highlight both the proximity and difference to the assumption of individual intentional autonomy, I propose to call it the assumption of *individual motivational autarky*. I shall turn to the relation between the autonomy assumption and the autarky claim shortly. To introduce the idea of individual motivational autarky, let me just highlight the difference in focus. Whereas the principle of individual intentional autonomy states that individuals are (and should be interpreted as) *responsible* for – or *owners* of – their behavior, the autarky claim is a claim about the kind of *volitional resources* on which we might draw in ascribing agency to individuals. The term "autarky" usually refers to a closed economy, especially to the mercantilist ideal of an empire with no outside trade. The Greek word is composed of "autos", the self, and the verb "arkein", "to suffice", meaning self-sufficiency. This captures nicely what is at stake here on the level of *intentional* rather than economic resources. Individual motivational autarky amounts to the claim that in the last resort, *only the individual's own wishes, desires, projects, volitions, or whatever pro-attitudes he or she might have* are fit candidates to make sense of their behavior.

(C) Individual Motivational Autarky: Any interpretation of an individual's behavior has to bottom out in that individual's *own* pro-attitudes.

In other words, motivational autarky is the view that, on the basic level, individuals should be taken as acting exclusively on *their own* desires, plans, commitments, intentions and so on; loosely speaking, only in terms of the members of the acting individual's own "motivational set" is it possible to rationalize (or make good sense of) the individual's behavior.[10] This needs some further explanation.

[10] In Donald Davidson's words, "R is a primary reason why an agent performed the action A under the description d only if R consists of a pro attitude *of the agent* toward actions with a certain property, and a belief of the agent that A, under the description d, has that property" (Davidson 1963: 687, my emphasis).

Individual intentional autarky does, of course, not imply that *other individuals'* pro-attitudes can play *no role* in the interpretation of an individual's behavior. This would be obviously wrong, because people do not normally act *regardless* of what other people want. Quite often, people *do* take other people's pro-attitudes into account, and sometimes even *act on* other people's pro-attitudes. Thus it is clear that an interpretation of an individual's behavior should not methodologically treat that individual as *disregardful* of, or neutral towards, other people's wishes. Indeed this would be a great mistake, resembling some outdated economic models of human action. The idea of individual intentional autarky has nothing to do with such narrow-minded selfishness. Individual intentional autarky is perfectly compatible with the fact that other people's wishes can play an important role in how we act, and that we sometimes act *in accordance with* and even *on the base of* other people's wishes.

To put this differently, individual intentional autarky is not in conflict with the fact that action can be *other-regarding*. But – and this is the essential point – it imposes the following constraint: *if* individual A acts on individual B's pro-attitude, either of the following has to be the case: A has made B's wish *his or her own*, or A has some other appropriate pro-attitude, such as the wish to conform to B's wishes, or the wish to conform to the social norm of accommodating other people's wishes, or some such. For lack of a better term, let's call this the *other-directed pro-attitude condition*. It is always possible to do what the other wants, but if one does so, either of the following has to be the case: one has come to want it *oneself*, or one has some other appropriate pro-attitude such as the wish to conform to the other's desires, or the desire not to violate the appropriate set of rules of conduct, or some other kind of other-directed pro-attitude. In brief, motivational autarky is the claim that people cannot act on other people's wishes without having a volitional agenda of their own.

Just as with intentional autonomy, I introduced motivational autarky as a *methodological precept* rather than as an assumption about the ontology of action. I think, however, that there is a close link between methodology and ontology. To put it in a catch-phrase: methodology follows ontology. The question of whether or not we should stick to the rule of basing all intentional interpretations of an individual's behavior on a pro-attitude which we ascribe to that individual *herself* is ultimately settled by the question of whether or not there *are* such pro-attitudes at the base of the intentional infrastructure of the action in question. How is this question to be settled? Philosophers, as well as many non-philosophers, seem to take it for granted that the issue at stake is a *conceptual* one, and that motivational autarky is just as *essential* a feature of action as intentional autonomy (from which it is never clearly distinguished). This is to say that, according to the predominant view, an individual cannot be an *agent* without being motivationally autarkical. Motivational *heterarky* (i.e. the opposite of intentional autarky), just as intentional *heteronomy*, would displace that individual's agency. The reasoning behind this thesis seems to be the following: if the intentional interpretation of individual A's behavior were to bottom out in some of individual B's pro-attitudes rather than in any of A's own (other-directed) pro-attitudes, A's behavior would have to be interpreted *as B's action* rather than A's. Thus it seems that, insofar as A is an agent, she needs to be motivationally autarkical.

My claim is that this view is mistaken. I will not argue that there really *are* cases of motivational heterarky, even though I will present some evidence that this might actually be the case. Rather, my main aim is to show that while motivational autarky implies intentional autonomy, *the converse is not true*. Motivational autarky involves a *further claim*. Moreover, some forms of motivational heterarky are compatible with intentional autonomy. Let me first focus again on plural action. I have introduced the assumption of motivational autarky as the reason why most philosophers seem to think that the influence model of plural agency cannot literally be true. I certainly do not wish to deny that many (perhaps most) cases of action under the influence of another individual conform to the assumption of individual autarky. In these cases, the interpretation of these individuals' behavior has to bottom out in the respective individual's own (other-directed) pro-attitudes. But it is also true that there are some folk psychological views according to which no such additional other-directed pro-attitudes are needed in order to interpret an individual's behavior. This is particularly obvious in altogether unassuming cases of influence, especially in spontaneous, low-cost cooperative behavior. (We are now finally turning to the opposite extreme in the spectrum of the size of plural actions: from the Apollo program to smallest scale everyday cooperation). What is at stake here are simple patterns such as the following: holding a door open for a stranger, spontaneously helping a stranger to lift a baby carriage into the train, or moving aside a little on a park bench so that another person can find a seat, too (cf. Chapter 8 below). These are *social* actions, because they require cooperation, and they are *plural actions* exactly *insofar* as the helper's goal is the same as the individual's who is being helped (i.e. that the stranger pass the door, that the baby carriage be in the train, or that person P have a seat on the bench).

The decisive question is: do such cases conform to motivational autarky? I do not claim to have any conclusive evidence, but there are some reasons for doubt stemming from two sources: folk-psychology and the theory of empathy. From a pre-theoretic perspective at least, it does *not seem implausible* at all to assume that there need not be some wish to have another person sitting beside oneself, or a desire to conform to other people's wishes, or even just a particular disposition to conform to some set of rules, or *any such* pro-attitude, in order to move aside a little on the park bench (Paprzycka 2002). If I move aside, it might seem from a folk-psychological perspective that I do not do so because of anything *I* want, but I do so because of what *she*, the other, wants – and similarly for the other examples I have given. In this sense, the folk-psychological intentional interpretation of one's cooperative behavior does *not* bottom out in *one's own* volitions or pro-attitudes, but rather in *the other's*. Something similar seems to be true for certain kinds of acting under other people's influence, especially for some forms of obedience to authority, where people do not just give in to submissive desires of their own, but seem to have serious difficulty explaining to themselves why they conform to some other person's wishes.[11] Another line of argument that seems to suggest that there might

[11] Paprzycka (2002) mentions the case of Stanley Milgram's famous psychological experiments. I shall say more about this below.

be something wrong with the assumption of individual intentional autarky is the analysis of *empathy*. An important element of the *history* of the concept of empathy, from Theodor Lipps (1903) to current simulation theory, is the claim that there is a *direct connection* between the *understanding* of another individual's intentions, on the one hand, and *action tendencies* that are geared towards the same goal, on the other.

To sum up this argument, the effect of the dogma of motivational autarky is that it reduces our view of human interaction to cases in which there clearly *are* other-directed pro-attitudes. But folk psychology suggests that there are plural actions without such motivation. In these cases, folk psychology seems to allow for what I would like to label *motivational heterarky*. The intentional interpretation of such behavior does not bottom out in the acting individual's own pro-attitudes.[12] Remember that the "bottoming out" clause in the autarky claim allows for the fact that individuals often take other individuals' pro-attitudes into account, but requires that, in doing so, individuals act on a volitional agenda of their own. By contrast motivational heterarky is the claim that people may sometimes act on other people's wishes without having any volitional agenda of their own. I believe that there is much to say in favor of the assumption of motivational heterarky, and I hope that my remarks have been successful in raising some doubts concerning the universality of motivational autarky.

I will not defend and present any more evidence for motivational heterarky in this chapter. Instead, I set for myself a much more modest task for the remainder of this section, something I do hope to be *rationally* self-confident about. I will state and defend a claim concerning the *relation* between the assumption of intentional autonomy and the dogma of motivational autarky. My thesis is the following: our deep-seated conviction that each individual should be regarded as a responsible agent, and the widely shared assumption that the only intentional resource that can explain an individual's action are that individual's *own* pro-attitudes, are *two different claims*. In other words, it is possible to treat an individual as an *agent* without claiming that the interpretation of his or her behavior has to bottom out in his or her own pro-attitudes. Or, more precisely: while individual motivational autarky implies intentional autonomy, the converse is not true.

(D) Intentional autonomy does not imply motivational autarky.

I suspect that the main obstacle in the way of an adequate understanding of the role of intentional autarky in human action is that motivational autarky is mixed up with intentional autonomy. And, at first glance at least, it might indeed seem that the autonomy-claim and the autarky-claim amount to the same thing. After all, it does seem plausible that if an intentional interpretation of individual A's behavior were to bottom out in individual B's pro-attitudes rather than in any of individual A's own, the action in question would have to be attributed to B. A would be left no more than

[12] Views that are closely related to what I call motivational heterarky can be found in Roth (2006), Rovane (1998: Chap. IV), Paprzycka (1998, 2002). Needless to say, this paper owes a great deal to all of them. For further references, cf. Paprzycka 2002.

the role of a kind of a Manchurian Candidate, as it were, or an intentional zombie. But this view is mistaken (and the confusion of intentional autonomy and intentional autarky lies at the heart of our difficulty to understand the role of influence in plural action). It *is* possible to interpret A as acting on B's pro-attitudes without assuming other-directed pro-attitudes on A's part, and still interpret A's behavior as A's *own* action. Motivational heterarky does not *per se* compromise or displace an individual's agency.

The argument for my thesis is simple. It draws on the analogy between individual actions and those forms of plural agency which are at issue here, and on the distinction between motivation and intention. Take again the park bench case presented above as an example. Consider first a standard singular (or even solitary) version, in which A forms the intention to move aside on the park bench on the basis of some of his own desires. For the sake of the example, let's assume that a mild spring sun has come to shine on the side of the bench on which A is sitting. After a while, A is a little warm; pondering about whether to take off his jacket or to move out of the sunlight, he decides that he prefers the latter. So he moves to the shadowy part of the bench on the basis of his desire to cool down a little. It seems that this is perfectly sufficient to make sense of A's behavior. The intentional interpretation of A's behavior bottoms out in A's desire to cool down a little. A need not be ascribed any additional desire such as a desire to have his wish to cool down a little fulfilled. Still, A's moving aside undoubtedly is A's own action, even though A might not have a particular wish that he do what he wishes to do, but simply wants to cool down a little.

Let's now turn from the solitary to the plural action case. Assume for the sake of argument that it were in fact possible for A to form the intention to move aside on the bench on the basis of B's wish to sit down, without an additional underlying desire to conform to B's wishes. It seems hard to see why, in this case, the lack of some additional pro-attitude should now suddenly compromise A's agency, when it does not do so in the individual case. All that is needed to form an intention to move aside is some form of *awareness* of the other person's wish. It's still *his own action*, only now the intentional resources going into it extend beyond the range of A's own pro-attitudes. It's not that B somehow acts *directly* through A's behavior, bypassing and displacing A's agency. A's behavior does not have to be attributed to B's agency, rather than to A's; A does not behave on B's remote control. Rather, A's behavior still instantiates A's own action. A does not become B's *intentional zombie*, as it were, just because he acts on B's pro-attitude without there being any conforming pro-attitude from A's side involved. Behavior such as moving aside on park benches even when one does not have any particular wish to have another person sitting beside oneself, or to conform to other people's wishes, or even to conform to the norms of propriety, is not a form of intentional zombieism – not even a mild one. Rather, it is a matter of simple *politeness* (even though not all heterarkical behavior is of the nice, beneficial kind, as we shall see shortly). Heterarkical agents *do intend what they do* (e.g., move aside on the bench), but the chain of intentional interpretation leads beyond what's in their own solitary motivational set. These agents are intentionally autonomous, but not motivationally autarkical.

If this is true, if individual intentional autonomy is conceptually independent from the dogma of motivational autarky, and if there is such a fundamental difference between the two, the question arises: how come they have always been lumped together? *Why do we tend to mix up the idea of being the agents responsible for our own behavior with the apparently very different idea that in the last resort, only* our own desires *are fit candidates to make sense of our behavior?* In short, my answer is this: it is because, in *our culture* at least, motivational autarky describes the way people are *supposed to be* (and see themselves). Being the one and only ultimate source of the intentional infrastructure of one's own behavior may not be a conceptual feature of agency, but in our particular culture at least, it is a very basic and extremely strong *normative ideal*. While a person's explaining her actions in terms of another person's intentions is quite frequent in everyday talk, we tend to press for "deeper" explanations, and even to react *embarrassed*, if a person fails to come up with some pro-attitudes of her own in explanation of her behavior. It is as if such a person had failed to conform to our idea of selfhood, and it is very tempting to blame this on her way of describing her action, rather than on the structure of her action. People, we seem to think, really *shouldn't be doing things just because other people wanted them to be done*, without thereby conforming to any of their own wishes – and insofar as we regard them to be fully developed selves, they just *can't*. Therefore, we tend to believe that in such cases, there has to be something wrong with their interpretation of their action.

This negative evaluation of heterarkical behavior might surprise, especially since most of the presumptive cases of motivational heterarky discussed above are of a rather beneficial, pro-social kind (think of the park bench example). But even in such cases, heterarky it is not well regarded. People are welcome to assist other people, but in our culture at least, it is believed that they should be performing such acts *because they wanted to be of help*, and not just because other persons wanted those acts to be done. Moreover, there are distinctly negative examples of motivational heterarky. A vivid illustration is provided by Stanley Milgram's (1974) famous psychological experiments (here, I follow a hint given by Paprzycka [2002]). Remember that Milgram's test subjects – perfectly decent ordinary people – proved willing to administer deadly electroshocks to innocent others, just because they were told to do so by some authority figure. There were neither financial incentives nor sadistic inclinations involved. So how come those people did what they did? Of course, it is always possible to assume that people acted on some *desire* to conform to the authority figure's wishes, or some desire to be a "good" and obedient collaborator, or some such pro-attitude. Based on repeated interviews with his subjects, however, Milgram himself gives another explanation for his stunning results. He explains his test subjects' behavior with what he calls an "agentic state".

An agentic state, Milgram explains, is a condition in which a person sees herself as *acting on another person's desires rather than on any of his or her own* (Milgram 1974). The most convincing evidence for the existence of agentic states is the fact that in the interviews carried out immediately after the experiments, and again some months after the event, many test subjects proved to be utterly unable to *explain to themselves* why they acted in the way they did, and did not come up

with any compliant inclinations in explanation of their behavior. The reason might be that there really *were* no such compliant pro-attitudes. It strikes the reader of the subjects' statements printed in Milgram's book that this utter cluelessness concerning the deeper motivation for their action is even true of some of those test subjects who explicitly accepted full *responsibility* of what they proved capable of doing during the experiment, and were not just looking for excuses. It might well be that this cluelessness stems from the fact that the subjects were looking for the reason for their action in the wrong place: in their own "motivational set", instead of in the authority figure's.

In his book, Milgram tends to dismiss agentic states as some sort of illusion; moreover, he depicts agentic states as an *unusual* condition that requires the presence of authority. As is well understandable from his experiments, he sees agentic states as morally utterly condemnable. Thus the normative ideal of intentional autarky becomes very clear in Milgram's depiction of the fatal consequences of agentic states. By contrast to Milgram, I propose to consider three things: first, agentic states might not be simply a matter of self-deception; second, motivational heterarky might be a *normal* condition rather than an exception, which, third, may lead to morally disastrous consequences under conditions such as those examined by Milgram, but can also be very beneficial under such circumstances as those found in public parks, airports, and railway stations (think of the beneficial and cooperative examples of motivational heterarky mentioned above). In short, I will not pass any judgment on whether or not we should hold on to our ideal of motivational autarky. What is certain, however, is that we cannot even start to discuss the question of whether or not motivational autarky is indeed an ideal worthy of defense, if we continue mixing it up with intentional autonomy. Because intentional autonomy is a constituent of *any* action, it is not to be changed. By contrast, motivational autarky is a cultural ideal, which we may or may not want to uphold.

With this result, let us finally come back to the question of the role of influence in plural agency. If we accept the possibility of motivational heterarky, it seems that the problem diagnosed above simply vanishes. We do not have to deny the possibility that sense can be made of plural actions *as a whole* in terms of the intentionality of the leading individual, as is so often done in everyday parlance. If we discontinue mixing up intentional autonomy with motivational autarky, it becomes obvious that to base an intentional interpretation of the participating individual's behavior in question on the leading figure's volition does not mean to bypass or to compromise the other participating individual's agency. We may ascribe the leader an *intentional authority* over what is going on by way of an intentional interpretation of the behavior in question that bottoms out in the leading figure's volitions, without thereby divesting the other participants of their own individual agency. The other contributors can still be seen as *agents*, with their behavior instantiating *their own action*. The leading figure's claim to the entire action does not necessarily disregard the other participants' own individual agency, because it does not interfere with the other participants' ownership of their individual contribution to the plural action. That it really was *Caesar* who defeated the Helvetii in the battle of Bibracte (in the sense that an intentional interpretation of the movements of the roman legions

bottoms out in no other than his attitudes), does not contradict the fact that even the most obedient and servile soldier of his beloved tenth legion threw his javelin *himself*. Thus there is ample room to take the commonsensical notion of the role of influence, power, and volitional openness more seriously in the theory of plural agency, without letting go of the idea of the fundamental intentional autonomy of *all* participants, not just the leaders. Here, as in the above case of the collective agent model of plural agency, action theory should be more accommodating towards common sense.

§4 Intentional Individualism

In many cases, some collective agent provides the solution to the plural agent problem. In others, some individual's influence and authority does the job. While both models should be taken more seriously in action theory, it seems clear that the extension of either of these models is limited, that there are cases of plural actions that can be fitted into neither of them, and that both types of plural agents presuppose plural action of another kind.[13] Consider again the following example. If you and I go for a walk together, this is clearly a case of a plural action – the action requires more than one participant, and the two of us will pursue the same (token) goal, i.e. walking together. I have argued that this case cannot be fitted into the collective agent view, because collective agency involves a kind of agency that is *different* from that of the participating individuals, which does not seem to be the case here: if you and I go for a walk together, there are *two* subjects involved in the case, not one (there is no collective agent walking all by himself), and not three (there is no additional collective subject escorting the two of us through our walk). Thus the collective agent view does not cover this case. It also seems to be futile to try to fit it into the influence model. Each of us will be walking with the other, but none of us has a claim to ownership with regard to our walking. Our walking is something *we own together:* in such cases, *ownership is shared*. This brings us, finally, to the third and last common-sense concept of plural agency. In this last view, the plural agent is not *one* agent – neither *one individual*, as in the influence view, nor *one collective*, as the collective agent view has it. Rather, the plural agent(s) are *many*: acting jointly, as it were, or hand in hand, in pursuit of the one shared goal.

I believe that this model of plural agency – I shall call it the *teamwork view* – is *the most basic one*. Teamwork is *presupposed* in the collective agent view, insofar it is only by virtue of teamwork that there are any collective agents at all; for there to

[13] I have argued above that, for collective agents to emerge from the agency of individuals, there has to be some communication going on between these individuals, which is plural action. A similar point can be made with regard to power and influence. In most cases, power is based in – or generated by – some form of *collective acceptance.* This is a shared intentional attitude, which is typically expressed in some form of declaration, or affirmed in some other form of public expression. This in turn is a plural action of the third type, i.e. teamwork.

be a collective agent, individuals have to *act jointly* in pursuit of the goal to *create* and *maintain* a collective agent. Also, it seems that most cases of the influence type of plural agency can also be modeled on the teamwork view. If it seems correct to ascribe the Helvetii's defeat in the battle of Bibracte to Cesar, it is no less correct to ascribe this action to *the Romans*, or to those Romans active in the course of the events, acting jointly as a team under Cesar's leadership. Thus it seems that the teamwork view is much more than just one view of plural agency among others. It is the bedrock of plural agency, and should therefore be the main focus of any theory of plural action.

But this model, too, has its difficulties. The most obvious problem any theory of teamwork will have to cope with is that of reconciling the *unity of the action* on the one hand with the *plurality of agents* on the other. In the current debate, this is *precisely* what the *concept of shared or collective intentionality* is meant to do. The claim is this: many people can intend *one and the same action* precisely insofar as they *share* the respective intention. The problem, however, is that it is somewhat unclear what it means to share an intention. Cakes and cars can be shared – one (token) cake, many pieces, one (token) car, many users – but intentions? What can talk of "sharing" possibly mean in this context?

Looking at the debate on collective intentionality that has evolved over the last 20 years, it becomes obvious that most authors tend to understand the sharing of intentionality not in the straightforward sense, but rather as a metaphor. According to authors such as Raimo Tuomela (as read by John Searle), John Searle himself, and Michael Bratman, there is no single (token) shared intentional state that is behind the joint intentional activity, but many intentional states instead, individual intentional states that are marked out from those involved in the case of solitary or singular agency in that they are either of a special *form* (Searle), *mode* (Tuomela), or *content* (Bratman), providing the "glue" for collective intentionality. In other words, the existing accounts of collective intentionality tend to be of a *distributive* kind. I call *distributive* those conceptions of collective intentionality which claim that, whenever people share an intention, each individual has his or her own intention, and that there is no such thing as one (token) intentional state that is shared by the participants in the straightforward sense of the term.[14] In other words, distributive conceptions of collective intentionality are marked by what I propose to call *intentional individualism*.

(E) Intentional Individualism: Any interpretation of an individual's behavior has to be given in terms of *individual* intentional states.

[14] In the second chapter of his *Analysis and Metaphysics* (1992), Peter F. Strawson introduces a distinction which is important to correct John Searle's influential misunderstanding of Tuomela's position. Strawson distinguishes between *reductivist* and *connectivist* analysis. In contrast to analyses of the reductivist kind, connectivist analyses do not identify independently existing "building blocks", or "atoms", but rather elements that might, for their existence, be dependent on each other. Insofar as this is true for Tuomela's analysis of shared intention, his position is not distributive, contrary to what Searle's reading suggests. In the meantime, Tuomela has repeatedly endorsed a non-reductivist reading of his position (e.g. in Tuomela 2007).

I have put intentional individualism like this so that it is a much weaker claim than individual motivational autarky. It constrains the class of possible mental states required to make sense of an individual's behavior to some *individual subject's pro-attitudes* (which could be *either* the respective individual's own – this is the case of individual motivational autarky – or any other individual's). Needless to say, most philosophers of collective intentionality implicitly accept not only intentional individualism, but individual motivational autarky, too, claiming that the pro-attitudes in terms of which sense should be made of the behavior of any individual participating in a plural action should not only be *individual* pro-attitudes, but that respective individual's *own* pro-attitudes. Having discussed some of the problems of the dogma of motivational autarky at some length in the last section, however, let's focus on intentional individualism here. It might seem that intentional individualism is so obviously true that it is not in need of further substantiation. It might appear that to assume that there is collective ownership of an action that cannot be ascribed to a separate collective subject, but that is *shared* among the participants intending the action *jointly* in the straightforward sense, would seem to amount to some implausible *fusion of mind*. Most philosophers of collective intentionality think that the idea of a non-individual mind is so terribly and obviously mistaken, that there is no need for further argument.[15]

Before examining the hidden background of this almost universal endorsement of intentional individualism, let me first say a word about why I think it might be problematic. As I have said, intentional individualism forces us to adopt a distributive conception of collective intentionality. The problem with distributive conceptions of collective intentionality – at least with those that have been put forward so far – is that they tend to be *circular*. The objection of circularity points out that whatever *individuals* intend when they share an intention, already *presupposes* the shared intention.[16] In other words, the dogma of intentional individualism makes it impossible to understand the element of *intentional commonality* that seems to be *presupposed* whenever people form an intention to participate in joint intentional activities. The circularity issues of the existing distributive conceptions of collective intentionality lends some plausibility to the conjecture that intentional commonality is indeed *irreducible*, and cannot therefore be captured by a distributive conception of collective intentionality. Intentional commonality, as I propose to use the term, is incompatible with intentional individualism; it implies sharing an intention in the straightforward sense of the word: one (token) intentional state, many participants.

---

[15] Thus Searle – without providing any further argument – shuns such "perfectly dreadful metaphysical excrescences" (Searle 1998b: 150). Tuomela, in turn, dismisses non-individualistic conceptions of the mind as "spooky" (Tuomela 1995: ix, 5, 353, 367). For a closer examination of the role of the group mind in collective intentionality analysis cf. below Chapter 2.

[16] *I* can intend to do my part in a plural action x *as my part* only if *we* intend to do x; thus my intending to do my part as my part is no independent "building block" of collective intentionality, as a superficial reading of Tuomela/Miller's (1988) account has the authors claiming, but rather an element of a holistic intentional structure. Similar points can be made with regard to Michael Bratman's and John Searle's distributive accounts of collective intentionality (for a detailed analysis cf. Schmid 2005c).

If this is true, if there really are some more or less obvious problems understanding the structure of collective intentionality along the distributive line, and if this is due to intentional individualism, the question arises: why do most philosophers of collective intentionality simply take intentional individualism for granted, and not even think that it is necessary to provide an argument for its validity? My conjecture is the following: just as motivational autarky is usually mixed up with intentional autonomy, the possibility that intentional individualism might not be implied by individual intentional autonomy is overlooked. I believe that the worry in the back of the mind of the distributive philosophers of collective intentionality is that, if intentional individualism were not true, the individual participants would be deprived of the ownership of their contributive action, thus comprising their agency. This worry, however, is unsubstantiated.

(F) Individual intentional autonomy does not imply intentional individualism.

It is possible to interpret several individuals as sharing one intentional state (in the straightforward sense of the term), and still interpret these agents as the owners of their contributive action. The argument follows precisely the same line as the one we used to establish the independence of intentional autonomy from motivational autarky in the last section. The upshot is this. Even within an intentional interpretation of a given behavior that appeals to a *shared* intentional state rather than to any of the participant's own *individual* intentional state, we may still interpret the participating individuals as intending their contribution *individually*, and as owners of their individual contributive actions. The fact that their individual contributive intention is *derived* from a shared volition does not undermine the individuals' ownership of their contribution.

Intentional commonality does not compromise the participating individuals' agency. It is not the case that some group mind *displaces* the participating individuals' agency if those individuals were to act on a shared intention. Each individual still intends to do his or her own part *individually*, and is thus the owner of his or her contribution, but this participatory individual intentionality is derived from an intention that is not individual, i.e. from the *shared* intention to carry out the plural action in question together. Thus the chain of intentional interpretation of the individual behavior in question leads beyond what is intended individually – without thereby flying in the face of the fundamental idea that each individual participating in the process is an agent in his or her own right.

I think that once it becomes clear that we might drop intentional individualism without letting go of intentional autonomy, the urge to exorcise the group mind from collective intentionality analysis vanishes. We can be more relaxed with regard to the non-individualistic conceptions of the mind that are so pervasive in much of the earlier history of thought, and from which we might still learn a lesson or two about what it means to act jointly, as a team.

§5 Plural Agency and Methodological Individualism

Let me now wrap up the line of argument developed in this chapter. Where does the resistance against conceptions of plural agency that is so pervasive in action theory come from? I suggested that the attitude underlying much of this resistance is methodological individualism. The term "methodological individualism" was coined 100 years ago (Schumpeter 1908: 88–98). Schumpeter introduced the term to label the views he shared with Max Weber. Ever since Schumpeter coined the term, the issue has kept coming up in social philosophy, usually in cycles of about 20 years or so (cf. Udehn 2001). In each round, the controversy had a somewhat different focus. Under the title 'methodological individualism', issues as different as the limits of social planning, the relation between social action and social structure, and the role of collectivity concepts in social explanation have been discussed. I believe that the right way to celebrate the centenary of methodological individualism would be finally to come back to the heart of the matter.

At its (historical) heart, methodological individualism is about *plural action*, and more precisely, it is the claim that there are no plural agents. At the basic level of intentional interpretation, all agency has to be treated as singular agency: this is how one might summarize the central precept of methodological individualism. In other words: plural actions, as we encounter them in social life, should be ascribed to singular agents. This does, of course, not mean that there is no *social action*, i.e. that all action is (in the terminology developed above) of the solitary kind. Methodological individualists are well aware that many actions presuppose for their possibility the actions of other individuals. And it does not mean that there are no plural actions. Methodological individualists do not have to deny that sometimes, individuals act jointly in pursuit of the same goal. All that is claimed by methodological individualists is that these actions do not require any particular notion of agency, i.e. that it is enough to assume singular agents for the purpose of the interpretation of plural actions.

I have used the principle of rational intentional self-confidence to cast some doubt on this view above. If a plural action can indeed be legitimately interpreted *as one action* (and not just as an aggregate of many actions), singular agency simply will not do: we need a notion of plural agency (I labeled this the "plural agency problem"). Common sense has no difficulty providing suitable candidates for this role. The problem, however, is that these common sense interpretations of plural actions are not well received in action theory, and I suspect that a wide-spread, more or less tacit commitment to methodological individualism is to blame.

In this chapter, I argued that methodological individualism (qua singularism about plural agency) rests on three mistaken conclusions drawn from one valid insight. The basic insight is *individual intentional autonomy*: however plural an action might be, each participating individual's behavior has to be interpreted as *his or her own action*. We are, in other words, not intentional zombies. The three mistakes are the following: first, contrary to what methodological individualists seemed to think, individual intentional autonomy does not rule out the irreducibility of collective agency. Second, methodological individualists are mistaken in assuming that

individual intentional autonomy rules out what I called *motivational heterarky*, i.e. behavior whose intentional interpretation bottoms out in pro-attitudes that are not the respective individual's own. Some plural actions can be ascribed to influential individuals or authority figures, without thereby bypassing the other participating individuals' ownership of their contributive actions. And third, methodological individualism is wrong if it amounts to endorsing the view that individual intentional autonomy is in conflict with robust notions of *intentional commonality*. We need a solid conception of intentional commonality to solve the circularity problems encountered by the existing distributive notions of collective intentionality. Such a conception can be compatible with individual intentional autonomy. This insight should help to overcome the widespread fear of non-individualistic conceptions of the mind, and lead to more adequate theories of *teamwork*.

While there might be some *cultural* and *historical reasons* for the fact that the assumptions of autonomy, autarky and individualism usually come as a package, there is *no reason* why we shouldn't *unpack* methodological individualism and start to think about which items to keep and which to throw out. Well understood, this is *not to say* that those ingredients of methodological individualism which we might find unfit for the purposes of the theory of plural action might not turn out to be useful for some other purpose (e.g. as cultural ideals). As we know, the extent to which human coordination and cooperation is achieved by plural action varies from group to group, from society to society, and from time to time. While I am quite convinced that individual intentional autonomy is *universal*, I think that the same is not true for autarky and individualism. It might turn out that there is much more autarky and much more individualism in *some* societies than in *others*.

This brings me to my final point. It is well known that in the paper in which the term "methodological individualism" was first introduced into the English language, Joseph Alois Schumpeter himself limited its validity. It is often quoted – by Kenneth Arrow (1994), among others – that Schumpeter (1909) says that the social can sometimes be considered "as if" it were an "independent agency". Nobody seems to have noted so far, however, that there is yet another, much blunter limitation stated in this paper. Schumpeter goes as far as to admit that methodological individualism is a relative principle which should not be applied to a certain type of society: it should not be applied to *communism*. Writing in 1908 (the paper was published in 1909), there was no way for Schumpeter to know what was to come under this label. So what did he mean with the term communism? He meant this: a society in which there are not just individual wants, but shared wants, too, and where there is joint action based on these "social wants". Let me quote a passage from Schumpeter's paper:

> The only wants which for the purpose of economic theory should be called strictly social are *those which are consciously asserted by the whole community*. The means of satisfying such wants are valued not by individuals who merely interact, but by all individuals *acting as a community consciously and jointly*. (Schumpeter 1909: 216)

This means plural agency, and indeed it means plural agency of the fundamental *teamwork* kind. On the one hand, it might be true that there was much less teamwork in later socialist societies than Schumpeter could ever imagine. Yet on the

other hand, there are, without doubt, many more genuine teams at work in capitalist societies than individualists like to think. This is not just the case in large-scale ventures such as the Apollo-program, but, above all, in altogether unassuming everyday interactions. Plural Action is an important part of life. And if Schumpeter is right, it cannot be adequately understood within the framework of methodological individualism.