

Plural Action

Hans Bernhard Schmid¹

§1 “Intentional Self-Confidence”

In the earliest times of the Apollo Program, John F. Kennedy once went to Florida to visit Cape Canaveral. On his tour through the facilities, he addressed a blue-collar worker and asked him, 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 (or perhaps even funny) about this reply that it made it to an anecdote? It seems to be 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 worker’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, with a grain of salt, 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 within his or her forces. In the briefest (negative) version, the principle spells as follows:

¹ Various versions and parts of this paper have been presented on the following occasions: *Annual American Economic Association/Allied Science Association Meeting*, Chicago (Jan. 2007); *Departmental Philosophy Colloquia*, University of Fribourg, Switzerland (April 2007); *9th Philosophy of Social Science Roundtable*, University of South Florida, Tampa FL (April 2007); *Workshop on Altruism*, University of St. Gallen, Switzerland (May 2007); *Departmental Philosophy Colloquia*, University of Geneva, Switzerland (June 2007). I wish to thank the organizers and audiences there, especially Mark D. White, Gianfranco Soldati, Alison Wylie, Dieter Thomä and Kevin Mulligan. The following people have provided detailed comments and helpful criticism on earlier drafts of this paper: Monika Betzler, David Schweikard, Anita Konzelmann, Juliette Gloor, Caroline Baumann, Rebekka Gersbach and Ulla Schmid. I am greatly indebted to all of them indeed.

Very special thanks go to Arnold G. Simmel.

(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 sometimes, agents have high intentional *ambitions*, and seriously intend to do things they perfectly *know* they might very well turn out to be unable to achieve. One does, in other words, not have to be *certain* to be able to perform what one intends to do. This is to say that intentional self-confidence might be minimal. 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 PM, I cannot *intend* to have lunch there at noon. If I still go there, my intention is a different one: to act *as if* I didn't know about the Restaurant's opening hours, to knock on the closed door as if in surprise, or any other aim along these lines.

Second, 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. The fact that this is compatible with the principle becomes evident from the following specification:

(B) 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 us 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: let us label them "objective" and "subjective". Intentional self-confidence is *objectively irrational* if one intends to do things which one is *generally* and *under normal circumstances* unable to perform. (It is always possible fail at task one takes oneself to be perfectly capable to perform. This does not *per se* 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 *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

² For an early version of this principle cf. Baier (1970). The label "intentional self-confidence" is to be used *with a grain of salt*, because taking oneself to be able to do x is a matter of *belief*. By contrast, self-confidence is an *emotion*.

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".³ 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 seem to show rather clearly. And this very fact that not all intentional self-confidence is rational seems to be precisely what makes the blue-collar worker'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 guideline to quite a 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* to intend to do such a thing as 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 something of the kind of the moon expedition can indeed be described as *one* action. Is this true, and if so: to what particular *kind* of action do such things belong?

§2 The Concept of Plural 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 an 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,

³ It is needless to say that the constraints articulated in the principle of intentional self-confidence apply only to *intentions*, not to desires and other intentional states.

some *goal* is needed, i.e. something the agent *wants*, a state of affairs towards which the agent has some kind of pro-attitude. In what follows, I will use the term “desire” in the wide sense that comprises pro-attitudes of many kinds: wishes, interests, projects, commitments, and so on, or use the term “pro-attitude”, which also contains intentions and valuations. 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 be 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. We do not have to delve deeper into action theory here to answer the question: was man’s travel to the moon something of the kind of 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 goal-directedness and rationality of behavior, it seems to be even 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 the 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 us just assume for the moment that a plausible answer to the open question *had* be given, and have a closer look at the *type* of action to which something of the kind of 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 which 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). There are two ways in which sociality might be required: either *logically* (as in the classic case of marrying, which as a matter of logic 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 kind. Singular actions are social actions in which the participating individuals pursue *different goals*. (By way of an example, consider the case of my taking a plane back home for Christmas, which presupposes, among others, some mechanics' activity, but my spending Christmas at home and his getting the jet engines going are quite different goals.) By contrast to 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.) I call this the *plurality condition*. If Saturn V had been produced in 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 that each 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 be ultimately *one and the same* goal. In brief, the main characteristics of plural actions are these: *many* participants, *one* goal.

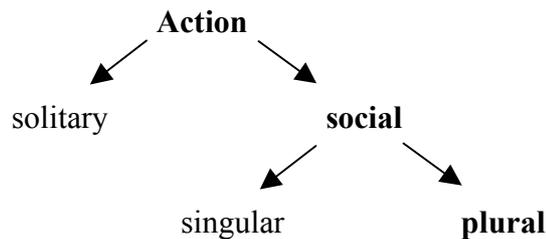


Figure 1: Taxonomy of action types

Plural actions have long been rather shamefully neglected in action theory. It was only in the course of the last two decades that the phenomenon has 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), pushing a broke-down car together (Tuomela 1995) serve as illustrations of the phenomenon. By contrast to this, my choice in this paper 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 choose such an 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 itself to be able to do what requires the *joint forces of many* to be done. So *who can possibly fit that bill? What are plural agents? Are there any plausible candidates to conform to the principle of intentional self-confidence rationally with regard to plural actions?*

In what follows, I shall proceed as follows. In the next section (§2), I shall evaluate three common sense replies to the question concerning plural agents (a collectivist, an individualist and a pluralist reply). These three replies correspond to three models (or perhaps types) of plural actions. Each of these three types meets with serious and sharp criticism, which I shall summarize (§3). In a second step, I shall identify the view that underlies all these objections, and that found its classic expressions in early formulations of the principle of *methodological individualism*. The view that causes so much difficulty in understanding the structure of plural action consists of a combination of three claims or assumptions. First the assumption of *individual intentional autonomy* (a), according to which under normal circumstances, the behavior of each single individual has to be interpreted as his or her own action. Second, the assumption of *individual intentional autarky* (b), according to which the intentional interpretation of each individual's behavior has to bottom out in his or her own pro-attitudes, and third the assumption (or dogma) of intentional individualism, according to which an intentional explanation of behavior has to appeal to *individual intentions* (c). I shall comment on how these assumptions shape our understanding of plural action (§4). My main thesis, which I shall present in the concluding section (§5), is that while b implies both c and a, a neither implies b nor c. I conclude that if this should turn out to be necessary, we should not be all too reluctant to depart from c and b in the course of developing an adequate account of plural action, for this would not mean that we would also have to depart from a. In other words: intentional autonomy without autarky and individualism is possible. I suspect that one of the reasons for our conceptual problems with plural agency stems from the fact that these three claims have always been lumped together.

§3 Three Models of Plural Agency

Let us now return to the crucial question concerning plural action: what are plural agents? Who could ever be so self-confident as to intend to do what requires the joint forces of many

to be done, without being objectively irrational? Common sense offers three types of candidates: the collective agent model, the leadership model, and the teamwork model.

i. The Collective Agent Model. In this view, plural actions are intended (and indeed performed) by *collectives*. The plausibility of this view can easily be illustrated with regard to our example. Just remember Kennedy's words in May 1961, with which he announced the start of the Apollo program: "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, perhaps 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 a kind of agent to match 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 individuals. Such agents, it seems, need not worry about expecting too much of themselves. The most obvious problems of this model, however, are firstly, the very notion of collectives having intentionality. In spite of some recent developments in that direction (Pettit 2002; Pettit/Schweikard 2005; Helm 2007), most philosophers are still rather uncomfortable with this idea, to say the least (I will come back to this below). Secondly, and perhaps even more seriously, the extension of this model seems to be rather limited. It covers only a small range of possible cases. It seems to be suited only for 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 taking a walk together. If the two of us take a walk together, there are only two agents involved in the process – not one, and most certainly not three!

ii. The Influence-Model. The second type of answer to the question of plural agency is this: those individuals who were in control of the project (according to its institutional structure) did it. While the statement "My task is to land a man on the moon" might sound rather silly as coming from a simple blue collar worker's mouth, it doesn't nearly as much so as 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 lower-ranking individuals'. To put it very bluntly: great people can do great

things. In the memory of Kennedy’s recently deceased court historian, I’m tempted to label this *second* model the *Arthur M. Schlesinger-view* of plural agency. 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 current humanities and social science. Among its advantages, however, is the fact that it is deeply rooted in everyday talk. Behind the palace of Versailles is Louis XIV intention; it was Vasco da Gama who successfully searched for the sea passage to India, etc. etc. Among the *problems* of the influence view are, again, an internal or conceptual difficulty, and the limited extension of the model. The conceptual problem is that plural agency is modeled very closely, indeed too closely after singular or even solitary agency. It is as if those leaders had performed their great deeds all by themselves. This model seems to suggest that leaders have many hands, feet and eyes, not just two of each. This simply bypasses the agency of the other individuals involved in the process, and thus fails to account for their contribution. The extensional problem is that there are many plural actions with *no hierarchy* between the participants – consider again the case of taking a walk together. In these cases, the influence model cannot be applied.

iii. The Teamwork-Model. 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 agent(s) are *many*. If one individual, all by him- or herself, cannot perform plural actions, perhaps many individuals *together* can: acting jointly, as it were, or hand in hand, in pursuit of the one shared goal. The most obvious problem of this view is to reconcile the *unity of 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 invoked for. 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 cake, many pieces, one car, many users – but intentions? What can the talk of “sharing” possibly mean in this context?

Model	Plural agent	Problems
Collective agent	Collective	Can collectives have (and act on) intentions? Limited extension
Influence	Individual	Can one individual intend another’s behavior? Limited extension
Teamwork	Individuals (jointly)	How can intentional states be <i>shared</i> ?

Tab. 1: Three models of plural agency and their problems

§4 Plural Agency and Methodological Individualism

I will not address any of the objections raised against the three models of plural agency directly. Instead, I will turn to what I believe *lies behind* the skeptical attitude towards plural agency as such, i.e. to a basic assumption that motivates the objections raised against each of the models or types of plural agency. It is, I claim, *methodological individualism*.

Joseph Alois Schumpeter coined this term exactly one hundred years ago (Schumpeter 1908, 88-98); he did so to label the views he shared with Max Weber. Ever since Schumpeter invented the label, the issue has kept coming up in social philosophy, usually in cycles of about twenty 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 agency*, and about the collective agent view of plural agency in particular. Max Weber's statement is clear: the only agents which social science knows of are individuals.⁴ 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 1988 [1921], 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 ascribing intentions and actions to collectives, one might wonder what reason Weber might have 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

⁴ 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 1988[1921], 13) 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.

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*⁵ – this naturally excludes plural action from the class of *explanantes*, because this kind of action seems to be of the sort that already involves collectivity (cf. the *plurality condition* mentioned above).⁶ 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 of any agents other than single individuals, and in particular with the assumptions of collective agents. For if collectives were agents, the participating individuals would be left with nothing more than the role of mere *instruments* or executing organs of some collective will, and would not be the proper agent's 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. Admittedly, Weber never explicitly says so, but I believe that this is the worry in the back of his mind that leads Weber to opposing the group agent assumption.⁷ Because individuals *are* to be interpreted as agents, *nothing but* individuals can be interpreted as agents on the most basic level of intentional explanation. For any other type of agency would *displace* the participating individual's agency.

⁵ The Weberian project is to explain collectivities as “consequences and organizations of individual actions” (Weber 1988[1921], 13).

⁶ Along this line, Weber's is simply a reductionist view: while it is not necessary to use collectivity concepts to describe individual agency, all collectivity concepts can be translated into 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 state. 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 collectives. There are collectives, because people *think* there are collectives, and not the other way around.

⁷ 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, who to my knowledge went farthest among the 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 motivation for individual actions.

Even though his conclusion might be controversial, I think that Weber's premise is basically right. In the current debate, the commitment to individual agency seems to be almost universally accepted. 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 firm commitment to individual agency. I shall call this commitment the principle of *individual intentional autonomy*. In its shortest formulation, the assumption is the following:

(C) 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 even involved in what I call 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 is *beyond* any of the current controversies revolving around this concept. Intentional autonomy refers to a very basic and elementary way in which individuals are responsible of their behavior as agents.

For further clarification of the term intentional autonomy, I might as well introduce its equally neologistic counterpart, which is *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 *another agent one's*. In other words, intentionally heteronomous individuals would have to be taken to be 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 zombies to the philosophical literature. By contrast to this, the principle of individual intentional autonomy states that individuals are *not* intentional zombies in that they do *not* behave on remote control.

As far as I can see, the principle of individual intentional autonomy is rarely contested. In the literature, it simply seems to be taken for granted, and it is never even really identified as

such.⁸ Most certainly, it should not be dismissed light-heartedly, especially since wide areas of our everyday practices seem to rest on this assumption. Individual intentional autonomy is a central feature of our way of identifying actions and ascribing responsibility to agents. I see no reason why the principle of individual intentional autonomy should not be valid. The problem I wish to address, however, is a deep-seated confusion that permeates the entire debate, and that even starts to have some grip on some of our agency-related everyday notions. On a regular basis, the principle of individual intentional autonomy is mixed up with *two further assumptions*, which I shall call the *assumption of individual intentional autarky* and the *assumption of intentional individualism*, respectively. To clarify the content of these further claims, let me first turn to the distinction between intentional autonomy and intentional autarky. Whereas the principle of individual intentional autonomy states that individuals are (and should be interpreted as) *responsible* for their own behavior, individual intentional autarky is a claim about the *intentional resources* on which we might draw in ascribing agency to individuals. Individual intentional autarky amounts to the claim that in the last resort, *only the individual's own pro-attitudes* are fit candidates to make sense of their behavior.

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

In other words, individual intentional autarky is the view that on the basic level, individuals have to be taken to be acting exclusively on *their own* desires, plans, commitments, intentions and so on; loosely speaking, only the members of the acting individual's own "motivational set" are able to rationalize or explain the individual's behavior.⁹ This needs some further explanation. Individual intentional autarky does not imply that *other individuals'* pro-attitudes can play *no role* in the interpretation of an individual's behavior. People do not normally act *regardless* of what other people want. It is obvious that quite often, people *do* take other people's pro-attitudes into account, and sometimes even *act on* other people's pro-attitudes.

⁸ It is particularly noteworthy that those in the current debate who most fervently advocate the importance of a solid conception of collective agency have a firm implicit commitment to the claim that contrary to what seems to be Weber's view, collective agency is by no means in any sort of conflict with the participating individuals' intentional autonomy.

⁹ In what follows, I will speak of interpretation rather than of intentional explanation, because I wish to avoid a discussion of the role of causality in this paper.

Thus it is clear that an interpretation of an individual's behavior should not methodologically treat that individual as *disregarding* other people's wishes. 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 it imposes the following constraint: if individual A somehow acts on individual B's pro-attitude, A either 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. For lack of a better term, let us call this the *compliant 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: either one has come to want it *oneself*, or one has some other appropriate pro-attitude like 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 compliant pro-attitude.

The second assumption (or dogma, if you wish) that is usually lumped together and mixed up with individual intentional autonomy is *intentional individualism*. This assumption is much weaker than (and indeed implicit in) individual intentional autarky. It constrains the class of possible intentional states to make sense of an individual's behavior to some *individual's pro-attitudes*, either the respective individual's own (individual intentional autarky), or any other individual's.

(E) Intentional Individualism: Any interpretation of an individual's behavior has to bottom out in *some individual's* pro-attitudes.

Before we turn to a somewhat more detailed discussion of the relation between these claims, I wish to examine the effect they have on the theory of plural action. First, the dogma of individual intentional autarky requires us to assume compliant desires or intentions even in cases in which they are not required in the folk psychological view. This is particularly obvious in the case of small scale, spontaneous, low-cost cooperative behavior – we are now finally turning to the opposite extreme in the spectrum of the size of plural actions. What is at stake here are simple patterns such as 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. 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 baby carriage be in the train, or that person x have a seat on the bench). The decisive

question is: do such cases conform to individual intentional 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, 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. In this sense, the folk-psychological intentional interpretation of one's cooperative behavior does *not* bottom out in *one's own* wishes or intentions, but rather in *the other's*. (Some similar seems to be true for certain kinds of acting under other people's influence, especially for obedience to authority, where people seem to experience serious difficulty explaining to themselves why they conform to some other people's wishes.)¹⁰ Another line of argument that seems to suggest that there might be something wrong about 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 to current simulation theory is the claim of 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.

It is not my aim to discuss the arguments for and against the assumption of individual intentional autarky in any more detail. Rather, my concern is with the effect of this assumption on the theory of plural action. The effect is that reduces our view of plural action to cases in which there clearly *are* compliant pro-attitudes. But folk psychology suggests that there are plural actions without compliant pro-attitudes. In these cases, folk psychology seems to allow for what I would like to label *intentional heterarky*. Intentional heterarky is the

¹⁰ Paprzycka (2002) mentions the case of Stanley Milgram's famous psychological experiments. Milgram's test subjects – perfectly decent ordinary people – proved to be 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? Milgram himself explains his stunning results by what he calls an “agentic state”. An agentic state is a condition in which a person sees herself as acting on another person's desires rather than his or her own (Milgram 1974). In an agentic state, the agent's behavior outflanks her moral conscience, since only behavior that is based on the agent's own desires is subject to moral control.

opposite of intentional autarky; the intentional interpretation of such behavior does not bottom out in the acting individual's own pro-attitudes.¹¹

As to the effect of the dogma of intentional individualism on the theory of plural action, it seems that intentional individualism makes it particularly difficult to understand participatory (or teamwork) forms of plural agency, because it forces us to endorse a *distributive* conception of shared intentionality. I label *distributive* those theories of collective intentionality which claim that collective intentions are *aggregates* or *compositions* of the intentions of individuals, which serve as their building blocks. Having done so elsewhere (e.g. Schmid 2003), I will not go into this any further here. Let me just state the following: even though most of the many existing accounts of shared intentionality follow a distinctively *distributive* line, none has been successful in avoiding the *objection of circularity* so far. The objection of circularity points out that whatever *individuals* intend when they share an intention, already *presupposes* the shared intention. 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. There is strong evidence that intentional commonality is indeed *irreducible*, and cannot be captured by distributive concepts of collective intentionality.

The following schema lists the three elements of methodological individualism as distinguished in this section, together with their consequences for the theory of plural action.

¹¹ Views that are closely related to what I call intentional heterarky can be found in Roth (2006), Rovane (1998, chap. IV), Paprzycka (1998, 2002). It is needless to say that this paper owes greatly to all of them. For further references, cf. Paprzycka 2002.

	Claim	Antithesis	Consequences for PA
Intentional Autonomy	Each individual's behavior instantiates <i>his or her own</i> action.	Intentional Heteronomy („remote control behavior“)	The individual participants in plural actions are <i>agents</i> , not just instruments or organs.
Intentional Autarky	Any interpretation of an individual's behavior has to bottom out in that individual's <i>own</i> pro-attitudes.	Intentional Heterarky (other-regarding behavior without compliant pro-attitudes)	Conflict with some folk psychological notions of plural agency.
Intentional Individualism	Any interpretation of an individual's behavior has to bottom out in <i>some individual's</i> pro-attitudes.	Intentional Commonality (irreducible form of intentional jointness)	Vicious circularity of distributive conceptions of collective intentionality.

Tab. 2: Three elements of methodological individualism and their consequences for the theory of plural action

The dogma of individual intentional autarky excludes the possibility of intentional heterarky, which seems to be an essential feature of a wide class of everyday plural actions, and the dogma of intentional individualism excludes the possibility of intentional commonality, which seems to be especially important in teamwork cases. Personally, I believe that there is much to say in favor of both intentional heterarky and intentional commonality. This, however, is none of my concern here. I will neither defend intentional heterarky, nor intentional commonality. The reason for my not doing so brings us back to the starting point of this paper: it is lack of intentional self-confidence. I am aware that such a task could not possibly be achieved in a single paper, and I assume that it would be beyond my forces and abilities anyway. So I set myself a much more modest task for the remainder of this paper, something I do hope to be able to achieve. I will state and defend a claim concerning the *relation* between the three elements of methodological individualism that I have distinguished in this section.

§5 Intentional Autonomy without Intentional Individualism?

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 intentional autarky implies intentional individualism, intentional autonomy implies neither of the two.

(F) Individual intentional autonomy does not imply individual intentional autarky.

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

Even though we have to leave open the question whether or not intentional heterarky and intentional commonality do in fact exist, I defend the *conceptual* possibility of intentional autonomy without intentional autarky and intentional individualism. It is possible to interpret A as acting on B's pro-attitudes without assuming compliant pro-attitudes from A's part and still interpret A's behavior as A's *own* action. Intentional heterarky does not *per se* compromise or displace an individual's agency. And it is possible to assume *intentional commonality* (i.e. a strong conception of collective intentionality) and still interpret each participating individual's behavior as his or her own action. Intentional commonality does not compromise the participating individual's agency. It is not that some group mind *displaces* the participating individuals' agency if individuals were to act on an intention they literally share. The argument for my thesis concerning the relation between intentional autonomy on the one hand, intentional autarky and intentional individualism on the other is very simple. It draws on the analogy between individual actions and those forms of plural agency which are at issue here. Take the park bench case as an example. Consider first a standard individual version, in which A forms the intention to move aside on the park bench a little on some of his or her own desires. Let us assume that a mild spring sunlight has come to fall 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 his desire to cool down a little. It seems that this desire is perfectly enough to make sense of A's behavior. A need not be ascribed any additional desire such as a desire to have his wishes fulfilled. Still, A's moving aside undoubtedly instantiates 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 us now turn 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 base of B's wish to sit down, without an additional desire to conform to B's wishes. It seems hard to see why in this case, the lack of some additional pro-attitude should suddenly compromise A's agency, when it does not do so

in the individual case. It's still *his own action*, only that 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. 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. And similarly for the case of intentional commonality: just as A may form the intention to do x on the base of his own prior intention, he may form an individual contributive (we-derivative) intention to contribute his part to some jointly intended y. The fact that in these (hypothetical) cases, an interpretation of the behavior in question does not bottom out in the respective individual's own intentional resources does not mean that these individuals' agency is somehow impaired or even displaced. Once again, I wish to emphasize that I do not claim that intentional heterarky and intentional commonality do in fact exist; all I do claim here is that it is not necessary *on conceptual grounds* to abandon the principle of individual intentional autonomy if one decides to *depart* from the dogmas of individual intentional autarky and intentional individualism. Intentional heterarky and intentional commonality are, in principle, compatible with individual intentional autonomy.

If this is true, if individual intentional autonomy is conceptually independent from the dogmas of intentional autarky and intentional individualism, the question arises: how come they are always 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's because in *our culture* at least, intentional 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 culture at least, it is a very 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 one of her own pro-attitudes in explanation of her behavior. People, we seem to think, shouldn't be doing things just because other people wanted them to be done – a vivid illustration for this point is provided by Stanley Milgram's (1974) famous experiments.

Thus my conjecture is this: While there might be some *cultural* and *historical reasons* for the fact that under labels such as "methodological individualism", the principles of autonomy, autarky and individualism 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 (such as, for example, normative cultural ideals). As we know, the extent to which human coordination and cooperation is achieved by plural action varies from individual to individual, from society to society, and from time to time. While I am quite convinced that individual intentional autonomy is *universal*, I think that it might well be 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 English language, Joseph Alois Schumpeter 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”. Next to nobody seem to have noted, however, that there is yet another, a much blunter limitation stated in this paper. Schumpeter goes as far as to stating 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 coming up under this label. So what did he mean with 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 *teamwork* kind. On the one hand, it might be true that there was much less teamwork in later socialist societies than Schumpeter imagined. Yet on the other hand, there are, without doubt, much 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.

Remember, I do not claim here, that he is right. All I claim is that *even if he were*, we might still hold on to the central claim that is at the heart of methodological individualism: the principle of individual intentional autonomy.

References

- Arrow, Kenneth J. (1994). Methodological Individualism and Social Knowledge. *The American Economic Review* 84/2, pp. 1-9.
- Baier, Annette (1970). Act and Intent. *Journal of Philosophy* 67, pp. 648-658.
- Bratman, Michael E. (1999). *Faces of Intention. Selected Essays on Intention and Agency*. Cambridge UK, CUP.
- Bratman, Michael E. (2007). *Structures of Agency*. Oxford, OUP.
- Gilbert, Margaret (1996). *Living together. Rationality, Sociality, and Obligation*, Lanham: Rowman and Littlefield.
- Heath, Joseph (2005). Methodological Individualism. *Stanford Encyclopedia of Philosophy* (www.plato.stanford.edu).
- Helm, Bennett (2007). Plural Agents. Forthcoming in *Nous*.
- Milgram, Stanley (1974). *Obedience to Authority. An Experimental View*. New York, Harper & Row.
- Paprzycka, Katarzyna (1998). Collectivism on the Horizon. *Australasian Journal of Philosophy* 76, pp. 165-181.
- Paprzycka, Katarzyna (2002). The False Consciousness of Intentional Psychology. *Philosophical Psychology* 15/3, pp. 271-295.
- Pettit, Philip (2002). Collective Persons and Powers. *Legal Theory* 8, pp. 443-470.
- Pettit, Philip/Schweikard, David (2005). Joint Actions and Group Agents. *Philosophy of the Social Sciences* 36/1, pp. 18-39.
- Rovane, Carol (1998). *The Bounds of Agency. An Essay in Revisionary Metaphysics*. Princeton NJ, Princeton University Press.
- Roth, Abraham (2006). Shared Agency and Contralateral Commitments. *The Philosophical Review* 113/3.
- Schmid, Hans Bernhard (2003). Can Brains in Vats Think as a Team? *Philosophical Explorations* 6/3, pp. 201-218.

- Schumpeter, Joseph Alois (1908). *Das Wesen und der Hauptinhalt der theoretischen Nationalökonomie*. Dunker & Humblot, Leipzig.
- Schumpeter, Joseph Alois (1909). On the Concept of Social Value. *The Quarterly Journal of Economics* 2, pp. 213-232.
- Searle, John R. (1990). Collective Intentions and Actions. In Philip R. Cohen, Jerry Morgan and Martha E. Pollack (eds.) *Intentions in Communication*. Cambridge Mass., MIT Press, pp. 401-415.
- Tuomela, Raimo (1995). *The Importance of Us. A Study of Basic Social Notions*, Stanford, Stanford University Press.
- Udehn, Lars (2001). *Methodological Individualism. Background, History and Meaning*. London, Routledge.
- Weber, Max (1988 [1921]). *Wirtschaft und Gesellschaft*. Tübingen, J.C.B. Mohr.