

AGENTS IN ACTION

Ralf STOECKER
Universität Bielefeld

Summary

I offer a justification for the received view that the characteristic feature of agents is to be found in the particular way their behaviour is explainable. Agents are people who have acquired three skills: (i) to act in accordance with inner or public deliberation; (ii) to do many things almost as if they had deliberated; and (iii) to recognize situations where it is worthwhile to switch from the second to the first skill. We can therefore assume that agents behave as if they were accompanying their behaviour by constant thinking although they don't actually deliberate all the time. This view allows for some attractive solutions for notorious problems in action theory but has the surprising ontological consequence that, although there are agents in action, there are no actions.

Agents are beings who exhibit agency. But even agents do not always act. They do many other things, too, that are not instances of agency, e.g. they stumble, faint or catch a cold. It is therefore legitimate to ask, where the difference lies: What is the characteristic feature of agents in action? It is this basic question of action theory that I shall try to answer.

The answer is based on two, rather uncontroversial presumptions. First, acting is (in a very broad sense) *being responsible for something*. Or to put it slightly differently: when an agent is acting something is *due* to the agent. And, secondly, what is peculiar to agency is to be found in a *special way*, in which the agent is responsible. The question is only how to characterize this special way.

In my paper, I shall start with a paradigmatic case of agency, give

a first tentative account of what makes it one, and then try to broaden the account sufficiently to cover all instances of agency.

1.

The paradigmatic case derives from A.I Melden's influential book "Free Action" (Melden 1961, p. 18): Imagine a driver (presumably in the early days of the 20th century) who is signalling a turn by raising her arm. Raising the arm certainly is an instance of agency, but why is it so? The following information (which isn't to be found in Melden) might help in answering this question:

Immediately before the driver's hand rises she says to herself:
 "Oh, there is the crossing, where I have to turn right. I had better raise my arm and notify the other drivers that I am about to change direction, so that they don't bump into my car",
 and then the arm goes up.

In this case it is quite obvious what it is that makes the driver's raising her arm an instance of agency. It is the fact that she has gone through a process of deliberation, of thinking, and, moreover, that the thinking in some way has led to her arm's rising. So what is peculiar to an agent's agency, it might seem, is that there is a deliberation which leads to her bodily movement. Accordingly, the following is a first tentative proposal:

- (I) An agent acts, if her *deliberation* leads to a movement of her body.

Yet this proposal is obviously not satisfactory. On the one hand, there are cases where deliberation leads to a bodily movement, without its being an instance of agency, e.g. when my knees tremble, because what I think about is so frightening. Agency must therefore be bound to a special way of bringing about bodily movements. (I) needs to be modified in this respect to be true.

On the other hand, even if (I) were true, it would only provide us with a sufficient condition for agency, not a necessary one. The pro-

posal isn't necessary in two respects. First it is restricted to bodily movements, and agents need not move their bodies. They clearly move around a lot of other things, too, and also sometimes seem to exhibit agency by moving nothing at all, letting things happen instead.

The second reason why (I) doesn't provide a necessary condition for agency is that we only very rarely think before we act. No tolerably experienced driver would go through a deliberative process like the one in the example before signalling a turn, and the same is true for most other things we do. Normally, we don't waste a thought on what to do, we simply do it. And after all, that couldn't possibly be different. Deliberation itself presumably exhibits agency, hence any attempt to tie agency to a preceding deliberation would probably lead into a vicious regress.

So agency is not bound to being deliberate, but deliberate action is at least a paradigm case of agency. It is therefore worth taking up the first objection and asking in which special way our deliberations lead to our doing what we do in the case of action (by contrast to a case of, say, our knees trembling). To answer this question I shall engage in some plain folk psychology first, and ask *how children learn to think*.

Learning obviously happens in a social context, mostly in interaction between the children and their parents. What the children learn is to behave in accordance with something the parents tell them to do. Parents say: "Don't touch the oven", and the child – hopefully – keeps her fingers away from the oven. But parents usually say more, they don't confine themselves to such a meagre order, they supplement their advice with reasons. They say for example: "Don't touch the oven, you will hurt yourself." Hence, the children learn that things have to be done *because* there are reasons for it.

From here it is only a small step to learning that usually there are reasons not only for one type of behaviour, but for a number of different, alternative types, and at this point the children quite naturally join into the decision process themselves. They start to argue with their parents, protest against their advice, come up with new ideas of what to do or what to achieve, etc. That is, the children learn to participate in a common deliberation, which aims at leading them to do whatever it is that emerges from that joint deliberation.

In a final step the children learn to take over the whole procedure and run it on their own. This step is important, because now they need no longer consult others, to find out what to do, they can talk it over with themselves. (At first this will be done audible, but then, as usual, silently, in their heads.) That is when the children have learned to think.

I take it that all of us are acquainted with the course of such a learning process. It shows that reasoning ultimately depends upon a social practice which people have learned to engage in and then to obey. But this learning process in turn allows us to understand in what special sense an agent's behaviour is due to her deliberation. The agent has acquired a *disposition*, the disposition to act in accordance with a decision process, whether this process occurs in public, together with others, or internally in the agent's inner monologue. Hence, the behaviour is due to the agent in the sense that it is the actualisation of a very special disposition of her (while the trembling of my knees, although it may also be explainable with recourse to a disposition, is certainly not due to the disposition to act in accordance with a decision process). Agents are skilled reasoning internalizers as well as reasoning executors, this is the general justification for explaining what they do on the basis of their preceding deliberation.

This leaves us with the question which behaviour may be explained in recourse to which thoughts, since obviously not every kind of thought explains every behaviour. If on my way home from the university I rack my brains over a smooth solution for the mind-body problem this will hardly explain why I stop at a red traffic light. So the thoughts at least have to provide a good argument for the behaviour. But unfortunately this answer is not easy to understand in itself, first because it is not always clear what a good argument for a certain piece of behaviour looks like, secondly because of the general logical problem of matching arguments with behaviour, and thirdly because actual thinking usually is a very sketchy and gappy matter, which makes it even more difficult to regard e.g. an inward "Oh hell, there is that crossing" as a good reason to turn off. Nonetheless I shall for the moment postpone these three problems and go on the assumption that we know what we mean when we say that a thought provides a good reason for some behaviour.

We are then still left with the problem that most things we do are definitely not preceded by a reasoning process, and are therefore not covered by answer (I). The answer only provides a sufficient condition for agency, not a necessary one.

2.

One way to react to this problem is to turn the tables and, instead of regretting that the condition is not broad enough to cover all instances of agency, put the blame on the agents themselves. Answer (I), one could argue, would fit very well for people who always think before they act, yet alas we aren't such perfectly deliberate beings. Real, perfect agency, according to this suggestion, is to be found in what answer (I) expresses, everything else being just a low, deviant variety.

What makes this proposal interesting is that it is so clearly silly. It is no wonder and far from being regrettable that there aren't any of those super-deliberate beings around, since in any way they would not have survived long, at least not in a challenging, rapidly changing environment like ours. Thinking takes a lot of time, energy and attention, so that we would be in practice (and quite apart from any suspicion of circularity) totally stuck if we tried to consider everything we are going to do beforehand. It is far from being a deficiency of our reasoning faculty that we employ it rather sparingly. What it shows instead is the fact that our reasoning faculty is merely part of a broader skill characteristic of agency.

To understand this broader skill let us again look at children's learning history. The story I have told earlier was not complete. Children do not only learn to do whatever results from parents' advice or from discussion with others, what they also have to learn in most instances is to do the same thing again the next time without consultation. Children should not only for example obey our present warning not to touch the hot oven, they should also learn not to touch it tomorrow either, when we aren't around to warn them again, and in the end they should learn not to touch it, without wasting a thought on the matter themselves. It is necessary that they simply, automatically don't touch it. Thinking, as I said, is an expensive device that

should not be squandered for avoiding hot ovens.

So when children learn to think, they learn at the same time to make thinking superfluous, i.e. to behave as if they were thinking, but without actually thinking.¹ This is all very familiar to us. Imagine in our traffic example a second driver who is so well versed, that she simply signals a turn, without wasting her time in going through the reasoning process that crossed the mind of her inexperienced colleague. Thinking is a waste of time for her because she anyhow acts as if she had entertained these thoughts before, so that there was no need really to think them. In fact, it is a significant mark of her advanced driving skill that she is able to behave as if she were thinking without normally being forced to go through actual thinking processes – just as it is a mark of the children's skill of dealing with ovens that they behave as if they had reasoned about touching them or not.

Such a skill is not only indispensable for an agent herself, it also has consequences for others, who deal with the agent. When they know that an agent is disposed to act as if she were thinking about what to do, they need not worry about her actual thinking processes any longer, but can ground their expectations on a consideration of what the agent would have thought if she had thought at all. So in the example we don't bother about what actually ran through the experienced driver's mind, but treat her as if she thought the same things as the driving novice. This is so natural for us that we frequently even use the same word and say that she *thought* that she should signal a turn, although there was in fact no such thought. Yet we also have another way to express the same thing less misleadingly: We say that the agent *believed* that she should signal a turn.

Believing, as well as desiring, hoping, expecting and all the other so called intentional attitudes are part of the familiar explanatory strategy which Daniel Dennett has called the *intentional stance* (Dennett 1971). What I want to suggest now is that this strategy should best be understood as the strategy of treating agents as if they

1. This is certainly not the only way children learn to behave reasonably. They acquire a lot of important skills directly, via conditioned response, i.e. without first learning to reason about the matter (e.g. when they learn their hot-oven-behaviour from the experience of touching one). My point is that in order to acquire many (though not all) important types of behaviour they have to learn to behave as if they were thinking without actually thinking.

were thinking a lot more than they actually do, while the different kinds of attitudes correspond to different kinds of thoughts. So e.g., to believe that one has reached a particular crossroad amounts to being someone who behaves as if she would entertain the thought "I am approaching this particular crossroad" in her deliberation about what to do, and desiring to turn the car right amounts to being someone who behaves as if she were thinking "This is where I should turn right".

To be sure, this is only a first step into a full-fledged theory of intentional attitudes. It has for example to be explained why we can attribute the belief that one has reached a particular crossroad to someone who can't speak English and therefore could not entertain English thoughts. But I am confident that this and other problems could be dealt with by an adequate theory of meaning, truth, and interpretation, which alas is definitely beyond the scope of this paper.

The story isn't complete yet in another respect, since it may seem that so far I haven't done justice to the importance of actual thinking for our life. It looks as if I have degraded it to a learning aid for small children, like counting with fingers, which they leave behind as soon as they acquire the ability to act well without thinking. And this would certainly be inadequate.

To understand the role of thinking in our adult life it is useful to look more closely at the ability that is supposed to ground our intentionality ascriptions, the ability to act as if one had thought beforehand. It is after all amazing how a being ever could have the ability to mimic deliberate behaviour, without actual deliberation. There are so many new and unforeseeable situations calling for very sophisticated practical deliberations that the ability to simulate this behaviour must display an incredible flexibility. How can this ever be done habitually? The answer is surprisingly simple: Part of our skill to behave as if we were constantly thinking is the ability to recognize situations in which we had better stop acting habitually and *really think*. We shut off the automatic pilot, so to say, and fall back to our original cognitive ability of deliberating about the situation. That is, we have in fact no ability always to act as if we were thinking without actually thinking. We have the ability to employ actual thinking very selectively to make sure that our behaviour generally is in tune with what we would think. This is again already part

of our learning history. Children do not simply learn to do well without thinking, they also learn when it is necessary to switch back to real thinking. Accordingly, the experienced driver of our example is not only normally able to find her way automatically around through the traffic, she is also able to realize situations where it is better for her to give a real thought to what to do.

To sum up, this is what I want to offer as the very core of agency. There are three skills that mark an agent:

- (i) The skill to act in accordance with inner or public deliberation.
- (ii) The skill to do many things almost as if the agent had deliberated.
- (iii) The skill to recognize situations where it is worthwhile to switch from the second to the first skill.

Agents are people who behave as if they were accompanying their behaviour by constant thinking processes, though they are far from actually doing this. What they have learned is to employ their deliberative faculty only in special situations which do not fit into the learned patterns or where it is more promising to think, instead of proceeding automatically. And since agents have these skills, there is a powerful strategy for explaining what they do, namely, the intentional stance. As in the special case of deliberate action, these explanations are instances of dispositional explanations, since the three skills mentioned are specific dispositions of the agent. In this sense the action is due to the agents intentional attitudes and hence to the agent herself.

We can therefore give another, more promising answer to the question of what is characteristic for an agent's agency:

- (II) An agent acts, if there is a movement of her body which is explainable with recourse to her within the *intentional strategy*.

I take this answer to be basically correct, though in need of further clarification in order to overcome some difficulties. This is what I shall do in the rest of the paper.

3.

At first, (II) still seems to be insufficient to cover a host of paradigmatic instances of agency. Imagine that in the traffic example the driver's raising her arm gives a shock to her timid passenger, who hadn't expected the sudden move. Obviously, the driver acted, when she alarmed her passenger, but none the less she did not want to alarm him, hence what she did wasn't appropriate in view of her intentional attitudes. It wasn't something she would have decided to do, if she had thought about it.

Still, this example fits well into (II), we only have to understand the explainability condition broad enough. After all, what happened was that the timid passenger was alarmed, because the driver raised her arm, and the driver raised her arm because she wanted to. Hence we can give a correct explanation of the passenger's getting alarmed, with recourse to the driver's intentional attitudes, we only have to understand that the intentional stance explanation referred to in (II) can be complex. To be sure, there is still a difference between the driver's raising her arm and her alarming the passenger, namely, that only the former can be explained on the basis of the assumption that the agent would have decided to act in that way if she had thought about it. But this isn't a difference between correct and incorrect ascriptions of agency. Instead it marks the difference between what the agent does *intentionally* and what she doesn't do intentionally. The driver intentionally raises her arm, but she doesn't alarm her passenger intentionally (only inadvertently, unpremeditatedly, involuntarily or the like). The alarming isn't intentional, yet it is an instance of agency, since it is done by doing something else intentionally.

This leads to another, connected objection. If being intentional simply means being explainable with recourse to beliefs and desires, it seems that one could not understand what agents do when they form an intention for the future, i.e. in cases of what is usually called *pure intending*. As many authors have emphasized, there is a crucial difference between the driver's intending to turn right and her wanting to turn right, because intending is incompatible with other intendings, but wanting isn't. The driver can easily want to turn right (because of the beautiful landscape over there) and at the same time want to turn left (e.g. because that is the shortest way home), but she

can't simultaneously intend to turn right and intend to turn left. So pure intending isn't simply reducible to wanting. But if pure intending therefore has to be taken as an attitude of its own, it seems very natural to suppose that it is the distinguishing mark of intentional action, so that acting intentionally should be something other than acting on an appropriate pair of beliefs and desires.

Yet in fact there is no contradiction here. To see this, let us return to practical reasoning again. Our picture of its role is still, in an important respect, incomplete. We usually don't reason with others or with ourselves just about what to do right here and now, we also try to get clear about what has to be done later on, we consider whole courses of actions, in short we make *plans*. As Michael Bratman in particular has pointed out, intending has to be understood through its role in action plans (cf. Bratman 1987).

Plans are as important for the coordination of the different things that one agent does as for the coordination of what several agents do. Their importance lies in the fact that they mark a boundary, at least provisionally, between what to think about and what could be taken for granted. In social discourse these limits are set by explicit *decisions*, on which the participants can usually rely until the decision is revised again. The agents can adjust their behaviour to the decision, which is often profitable for all parties. If e.g. the driver and her passenger have decided to take the street to the right and then to stop at a vantage point to take some pictures of the scenery, the passenger may already prepare his camera before they actually turn right and not left. It is reasonable for him to do this, because he can expect the driver to turn right or else to resume the dialogue and revise the former decision.

But decisions in advance aren't only useful for the coordination of social action, they also play an important part within the intentional stance. If we want to know whether the driver will turn left or right, her ambivalent desires may give us no clue as to what she will do. But when we learn that she has agreed with her passenger to turn right, we can for the time being go on the assumption that she will do this, i.e. we can assume that she *intends* to do it. The ascription of an intention is grounded in the characteristics of decisions in social deliberation. We are justified in ascribing such an attitude to an agent because the agent has learned to behave according to public decision

procedures which entail that decisions reached by these procedures bind further thought and behaviour. Thus we are normally justified in ascribing to agents the intention to carry out such decisions.

Yet only some intentions actually derive from a public decision procedure. Usually we ascribe an intention, because the agent herself, in making plans, came to a decision. We can do this because, as I have maintained, these solipsistic deliberations are internalizations of the social practice and basically follow the same rules and carry the same commitments. No agent would be able to carry out even moderately complex projects if she would not be disposed to stick to her previous private decisions and to treat them as boundaries of what to worry about. Hence, we can take advantage of this disposition and build it into our explanatory strategy, by including intentions as separate attitudes into the intentional stance.

Nonetheless, we don't have to maintain that intentional behaviour always entails that the agent has an intention to do it. What counts against this "simple view", as Bratman has called it, is the fact that we often act without having reached a decision in advance. The experienced driver did not first decide to raise her arm and then raised it, hence we need not ascribe to her a special attitude (an intention) towards raising it. She acted *instead* of deciding to act. So the assumption of the simple view, that in intentional agency the agent has an appropriate intention, is unnecessary and can be discarded.

The orthodox view according to which explainability by an appropriate pair of beliefs and desires is the mark of agency is correct. But the explanatory value of the intentional stance is not confined to these so-called primary reasons. We also want to learn why these reasons were effective, and here the additional information that the agent had formed already an appropriate intention enhances our understanding. In this way intentions and plans play an important and irreducible role in our intentional strategy.

4.

Still, there are other difficulties for (II). In this section I return to the question which actual thoughts could explain which behaviour. As I have already mentioned, it is obvious that thoughts have to provide

good reasons for the behaviour they are to explain. But there were three problems with this move: first, to determine what, under given circumstances, a good reason for action is, secondly to explain how reasons could fit to actions at all, and finally to deal with the fact that real thoughts are often so sketchy and gappy. I shall start with the third problem and then discuss the second one. With regard to the first one I shall simply assume that there is a way to answer it.

The third problem can also be put as follows: How could we ever base our intentional strategy on the assumption that people act as if they were constantly deliberating about what to do, if actual deliberation frequently is confined to thoughts like: “Oh hell, there is that crossing”? – The answer is to be found in the observation that not only agents who act without deliberating can be regarded as if they were thinking, but even agents who actually think should still be regarded as if they were thinking *a lot more*. Our inexperienced driver for example, who reflects on what to do, still does not entertain every relevant thought, but only those that people with little driving experience tend to have. She does not, for example, think that it would be expedient not to have a car accident or that the raising of her arm won’t damage the car. But she acts as if she were thinking these thoughts, i.e. she wishes not to have a car accident and believes that raising her arm will not damage the car. So, employing the intentional strategy *always* consists in treating agents as if they were entertaining a lot more thoughts than they actually do.

Remember the third skill mentioned earlier, the skill to employ actual thinking selectively so as to secure that the behaviour as a whole corresponds to what the agent would have thought to be best. After all, the situations where it is necessary to think aren’t normally such that everything runs off the rails, in most cases there is only one specific aspect of the situation that has to be reconsidered. Consequently in all other respects the agent does very well to act without thinking. Hence we can also treat the agent who has entertained the thought “Oh hell, there is that crossing” as someone who acted not only on this thought but also on many others like: “I should turn off at that particular crossing”.² That is, despite the rudimentary thought that ran through the agent’s mind, we can explain her behaviour with

2. The fragmentary character of actual thinking only mirrors the highly ellip-

recourse not only to the belief that this particular crossing has come, but also to the desire to turn off there.

This is why we can generally demand that an explanation for an agent's behaviour mention good reasons for it. Although the agent usually only thinks about what seem to her critical or doubtful aspects, we can treat her as being in command of an abundance of relevant thoughts in the light of which she behaves. Ideally, we do this by forming a string of reasons that almost inevitably lead the deliberator to a conclusion, i.e. we form a so-called *practical syllogism*.

This is now a good point to pick up the other problem, the apparent logical mismatch between reasons and action. Aristotle left us with the thesis that the conclusion of a practical syllogism is an action, but actions are no linguistic entities and cannot be squeezed into logical relations. They can't be true, cannot contradict anything and have no implications. Or at least they can't if we follow the orthodox view in action theory, namely, that actions are a special kind of events, where events are assumed to be spatio-temporally located particulars in the world. How could such worldly entities ever fit into a syllogism?

To solve this problem it is necessary to say something more on the place of *actions* in the account of agency proposed. Actions have not turned up in my exposition yet, but it seems easy to locate them in proposal (II). Actions, one is tempted to add, are exactly those bodily movements that are intentionally explainable with recourse to the agent. The rising of the driver's arm is her action of raising the arm.

This view, that an agent's action is to be identified with a particular bodily movement of her, goes back to Anscombe and Davidson. It has notorious difficulties with the spatio-temporal localization of some actions, e.g. killings, which consist in producing or achieving something.³ But there is another, more interesting reason why this view would not fit well with the account of agency sketched in (II). There is a looseness in (II) that has to be tidied up now. According to

tical character of mutual deliberations with close friends and relatives, which also does not render these deliberations less stringent or conclusive.

3. Cf. Stoecker 1993, 1998.

(II) the driver's raising her arm is an instance of agency, when her intentional attitudes explain her arm's rising (within the intentional strategy). But this could not literally be true, because the rising of the driver's arm is an event, and events are not the right kind of entities to be explained at all. Therefore proposal (II) cannot be correct.

This difficulty is not obvious because it is so common for us to talk about the explanation of events. Evidence that nonetheless events cannot be explananda (nor explanantia) is to be found in the observation that explanatory sentences otherwise should allow for substitution of event descriptions *salva veritate*, which they don't. Consider an example by Alvin Goldman (Goldman 1970, p. 3): a man's anger can explain that he is answering the telephone loudly, although it doesn't explain that he answers the telephone at all. Goldman took his example to show that events are very slim, fine-cut entities. But in fact it is much more plausible to take his example as showing that it is a misconstrual of the logical form of explanatory sentences to regard them as expressing relations, one of the relata of which is an event like the arm's rising.⁴

A better account of their logical form becomes visible once we note that paradigmatic explanatory sentences relate the explanans to a "that"-clause.⁵ We don't explain the event of the arm's rising, but we explain *that the arm rises*. Hence, it is not surprising that these sentences display the same semantical anomaly as other "that"-clauses, in particular in ascriptions of propositional attitudes and indirect speech. As we all know, we had better not infer from "P believes that Venus is the morning star" to "P believes that Venus is the evening star" (although in fact the evening star is the morning star). What this shows is again that these sentences don't simply express relations between persons and e.g. stars, but have a more complicated ('opaque') logical form instead. Fortunately, there is an at-

4. There are several good reasons not to accept Goldman's fine-grained event individuation, e.g. the reason that it is sometimes merely a question of emphasis to turn a correct into an incorrect explanation ("His anger didn't cause him to say *hello* loudly, it caused him to say hello *loudly*."), which Goldman could only counteract with a further dramatic increase of still more fine-cut events. (Cf. Achinstein 1975 and Stoecker 1992, pp. 118 ff.)

5. In what follows I rely on a more elaborate account I have given in Stoecker 1992, chpt. 12.

tractive account of their logical form, put forward by Donald Davidson (cf. Davidson 1968), which can be employed for explanatory sentences, too. According to Davidson sentences with a “that”-clause should not be regarded as being single logically compound sentences, but as consisting of a sequence of two ‘paratactically’ attached sentences related by a demonstrative element in one of them (the “that”) which refers to the utterance of the other sentence. Davidson originally put this idea to use for the logical form of ‘saying that’-sentences, but it applies equally well for ‘explaining that’-sentences. Accordingly, explanatory sentences have the form: “X explains the following: *p*” (where “*p*” stands for a sentence like “the arm rises”). As one might put it, explanations don’t explain events, they explain facts. Certainly, the fact might be that an event of a certain type occurred, but it is not the event (‘as such’) that is explained, but the fact that such and such an event took place.

Proposal (II), then, needs to be revised:

- (III) An agent acts, if the fact *that* her body moves in a certain way is explainable with recourse to her within the intentional strategy.

By attributing agency to a person we relate her to the fact that her body moves, not to the movement itself. Agency isn’t understood as producing events, but as affecting the way the world is, just as believing is having an idea of how the world is and desiring is having an ideal of how the world could be. This is, after all, not surprising, since (intentional) agency is the actualisation of a disposition to behave in accordance with beliefs and desires. And it allows for a neat solution of the problem how actions could be conclusions of a practical syllogism. If action sentences display a structure parallel to the structure of belief- and desire-sentences, it is no longer a logical problem to relate them to the conclusion of a syllogism. As we ascribe beliefs and desires with the help of sentences that could have been premises in the agents reasoning process, we ascribe agency with the help of a sentence that could have been the conclusion of such a reasoning process. In both instances we use actual reasoning procedures merely as a device, a measure for explanatory purposes. What we affect when we act is therefore as much a conclusion of an

argument as what we believe is one of its premises.

Despite these advantages proposal (III) still shares one shortcoming with its predecessors, one that was already mentioned with respect to proposal (I). Prima facie, neither (I), nor (II) or (III) state necessary conditions for agency because they are confined to bodily movements, while agents do not merely move their bodies, but affect a lot of other things, too. Thus it is tempting to move from (III) to a still broader condition:

- (IV) An agent acts, if *something* is explainable with recourse to her within the intentional strategy.

Consider another famous example from action theory: Jones' buttering his toast slowly, deliberately, in the bathroom, with a knife, at midnight (Davidson 1967). The toast's being buttered is certainly no bodily movement, still Jones is acting when he is buttering the toast. According to (IV) this is due to the fact that the toast's being buttered is explainable by Jones's intentional attitudes. Hence, (IV) seems to be a more adequate proposal for understanding agency than (III).

On the other hand, though, the move from (III) to (IV) has a disturbing side-effect, the diminished role it assigns to actions. According to (I) and (II) agency was understood as a relationship between an agent and her bodily movements, which allowed for the identification of actions with those movements. According to (III) agency was characterized as a relationship between an agent and a fact, but it still was a fact about the agent's bodily movements, which, after all, could be taken (though slightly misleadingly) to be her actions. But according to (IV) there is no longer anything that could be taken as the agent's action. (Certainly, neither Jones himself, nor the being buttered of the toast are candidates for being an action.) Hence, if there isn't anything more to agency than (IV) then actions aren't essential to agency. And this sounds implausible, or at least premature. After all, one might protest, the butter doesn't get onto the toast by pure wishing, obviously Jones had to do something in order to get the toast buttered, e.g. take a knife and spread the butter onto the toast. So it seems that there still have to be bodily actions fitting into (III) without which there couldn't be agency, although they aren't mentioned in every ascription of agency. (III) may therefore, despite

appearances, state a necessary condition for agency.

This objection deserves a thorough reply. As a preliminary step it asks for a general explanation of the “by”-locution (which has already played an important part in my discussion of unintentional agency and is in any case one of the standard topics in action theory that has to be handled by an adequate account of agency). What do we mean, when we say that an agent does something *by* doing something else, e.g. that Jones buttered the toast by spreading butter with a knife onto it?

As long as actions are taken to be events it is almost inevitable to assume that in employing the “by”-idiom we sometimes relate different descriptions of one and the same event, e.g. of what Jones did with the toast. But on this assumption we have the notorious difficulty to explain why “by”-sentences are irreflexive, asymmetrical and intransitive, although the identity-relation is reflexive, symmetrical and transitive, i.e. why we say that Jones buttered the toast by spreading butter onto it, but do not say e.g. that he buttered it by buttering it, nor that he spread butter onto it by buttering it.

If on the other hand action sentences are taken to express the existence of an explanatory relationship between an agent’s intentional attitudes and a state of the world, as I have suggested, the problem disappears. Seen from this perspective, connecting action sentences with a “by” relates two claims about the existence of explanations, such that the second claim provides a justification for the first one. For instance, in the bread-and-butter example, if we learn that the agent has buttered the toast by smearing butter onto it with a knife, we know that the agent’s intentional attitudes explain the fact that there is butter on the toast, *because* they explain that the knife has smeared butter onto the toast. To put it very general and roughly: In action sentences we say that something in the world is due to the agent’s wishes, and in “by”-sentences we (at least frequently) justify this claim by supplementing parts of the story which tell how it could be that in this particular respect the world obeyed the agent’s wishes.

Taking this explanation of the “by”-locution for granted, we can now reformulate the objection against (IV). As is well known, “by”-regressions stop somewhere. This is where (in Arthur Danto’s words) the *basic actions* are. Jones butters the toast by spreading

butter onto it, and he does so in turn by moving the knife up and down the slice of bread, but there isn't anything he does in order to move the knife, he simply moves it. Moving the knife is therefore Jones' basic action. What might be guessed now is that every "by"-regression stops at a bodily movement, or in other words: that for every case of agency there has to be a basic action that is a bodily movement. If this assumption turned out to be true, we should better not abandon (III) in favour of (IV), which means that there would still be a good sense in which actions are essential to agency, namely as the respective basic bodily movements. (We would then still have to keep in mind that, as we have seen in the example of Jones' buttering the toast, action sentences frequently don't refer to actions in this sense.)

Yet the assumption that every instance of agency is grounded in a bodily basic action is false, and hence actions aren't essential to agency (although we still can use the word "action" for those 'basic actions' that are bodily movements). The reason why it is false is the topic of the next, final section.

5.

Prima facie it seems to be obvious that an agent's wishes couldn't influence the world except by way of bodily movements. But at a closer look, this isn't true. Imagine that the driver in our example decides, instead of turning off at the crossroad, to drive on straight ahead. Obviously this is an instance of agency. The fact that the driver goes on straight ahead can be explained with recourse to her intentional attitudes as easily as if she had turned off. But driving on isn't something that is done *by* doing something else, particularly not by moving any of her limbs. Hence, the example shows that the world may very well accord to our wishes although we don't even lift a finger. This is the case, when the facts are due to our intentional attitudes, not because we have changed something, but because we have *let things go* without changing them.

Letting happen or abstaining from action usually are discussed in ethics, with respect to extraordinary situations (like those in which one might let someone die). But in fact our everyday agency is

soaked with instances of letting happen and non-doing which smoothly mingle with instances of active agency. Consider what we do when we care for an ill friend. We e.g. serve him delicious dishes, keep it warm in the room, let him sleep as long as he wants to, keep bad news from him, sit at his bedside etc.. These are things we do, but only part of it are done by moving around our limbs. So although it may be that what we first learn as children is to move (and also to stay put) according to advice and reasoning, in adult life agency is no longer necessarily bound to bodily movements. A large part of it doesn't need any bodily effort.

But what then is it bound to? Look again at the traffic example. What makes the driver act when she drives straight ahead is not her bodily influence on her environment, but her being in *control* of the environment. Her intentional attitudes explain the fact of her driving straight ahead, because it was at her command to intervene and turn off instead. This shows that we don't always need to engage in real action in the rudimentary sense of bodily movement in order to be responsible for what is the case in the world. Sometimes it is sufficient to have it in our power to prevent or modify the course of the world in this respect.

Yet sometimes isn't always. While actually fiddling with the world is not necessary for agency, simply being able to change things is not sufficient either. There are plenty of things we could change, although leaving them as they are isn't something we do. I am able to pull off some posters from a billboard in the university hall, but my leaving them there isn't something I do, it isn't an instance of my agency. What has to be added is that my ability to pull them off must make letting them hang *explainable* by my intentional attitudes. After all, agency demands that the intentional attitudes explain what is the case, and just the ability to change something does not amount to an explanation why the agent doesn't change it. For an instance of agency it is therefore not sufficient that something is in our power, we also have to have reason to exercise or not to exercise our power.

This is where the story fits well with the account of the intentional strategy given earlier. Employing this strategy means that we treat agents as behaving like ideally reflective beings who act as if they were thinking about everything they do. But the fact that I don't go

to the university hall now and pull down the posters is not something that is explainable by the fact that I could think it over and then would come to the conclusion that it isn't worthwhile. The reason is that doing something so odd isn't in the focus of my possible reasoning at all. It is something I wouldn't think of, so to speak. It wouldn't occur to me to do such a thing. Therein lies the difference between my not tearing down the posters and not waking up my ill friend.

This difference is crucial for understanding agency. On one side there are aspects of the world we accord to in our behaviour, although we only rarely reflect on them in actual thinking; and this can take the form of active intervention or of abstaining from interference. And on the other side there are aspects of the world we usually don't care about at all, like the posters hanging around in the hall, which we simply leave to themselves. The basis for this difference, though, can readily be traced back to the origins of the intentional stance and hence of agency ascription, which is to say, to the social practice of mutual deliberation. Social interaction after all is not only in need of coordination but also in need of *responsibility ascriptions*. Responsibility isn't something that people simply have or have not, basically it is something they take over or get assigned for certain domains, for which they have to take care. (The way to ensure this in social life is to assign to them quite literally the duty to respond, to tell why something happened and another thing didn't etc., which is usually connected with positive and negative sanctions by the group.)

This practice of responsibility allocation is an essential part of social deliberation, since it ensures mutual predictability. But it also sets boundaries to what the individual person has to deliberate about and hence which course of the world can be explained by her propensity to act as if she were constantly thinking, and which can't. Only facts that fall into the agent's domain of responsibility are therefore explainable with recourse to her intentional attitudes, but within the limits of this domain it does not matter whether the facts are due to the agent's active meddling in with her limbs or not.⁶

6. The difference between bodily doing and letting happen is important, though, for the question of where the limits of the individual domain of responsibility are, since active interference is a common way of expanding one's domain

What I want to advertise, then, is a sort of ascriptivism, though it is very different from H.L.A. Hart's original version (cf. Hart 1948/49). Agents are people who have learned to live in spheres of responsibility, i.e. they have learned that with respect to certain aspects of the course of the world they are expected to act as if they would take them into account in a full deliberation about what to do. A lot of bodily movements necessarily belong to these spheres, e.g. most movements of the limbs, yet others don't, for example the beating of the heart. Moreover, since the bodily movements are apt to change things in the world, agents gain power and control over parts of the world and hence can take over responsibility for much more than the movements of their bodies. They can e.g. be responsible for what is on the toast or where the car is heading. Action ascriptions express these responsibilities, but actions, as being identified with the bodily movements, can at best play a minor part in the story. And if we only were gods or pure souls, we wouldn't need them at all, because then as Plotinus once told us, the world would simply follow our pure intellect, effortlessly and without forcing us to waste any thoughts on it.^{7,8}

of responsibility into new areas, which sets it apart from letting things happen within the given domain.

7. Plotinus, *Enn.* IV, 8, 8, 42.

8. I have benefited considerably from critical discussions at the conference, in particular with Hanjo Glock, Jens Kulenkampff, Jonathan Lowe and Fred Stoutland. John Hyman and especially Rüdiger Bittner have done a lot to improve and correct the final version of my paper, for which I am very grateful.

REFERENCES

- Achinstein, Peter, "Causation, Transparency, and Emphasis", *Canadian Journal of Philosophy* 5 (1975).
- Bratman, Michael, *Intention, Plans, and Practical Reason*, Cambridge (Mass.) 1987.
- Davidson, Donald, "The Logical Form of Action Sentences", in: N. Rescher (ed.), *The Logic of Decision and Action*, Pittsburgh 1967.
- Davidson, Donald, "On Saying That", *Synthese* 19 (1968).
- Dennett, Daniel, "Intentional Systems", *The Journal of Philosophy* 68 (1970).
- Goldman, Alvin, *A Theory of Human Action*, Princeton 1970.
- Hart, H.L.A. Hart, "The Ascription of Responsibility and Rights", *Proceedings of the Aristotelian Society* 49 (1948/49).
- Melden, A.I., *Free Action*, London 1961.
- Plotinus, *Plotini Opera, editio minor, Vol. III*, P. Henry, H.-R. Schwyzer (eds.), Oxford 1982.
- Stoecker, Ralf, *Was sind Ereignisse?*, Berlin, New York 1992
- Stoecker, Ralf, "Actions, Reasons, and their Relationship", in: R. Stoecker (ed.), *Reflecting Davidson*, Berlin 1993.
- Stoecker, Ralf, "Tun und Lassen – Überlegungen zur Ontologie menschlichen Handelns", *Erkenntnis* 48 (1998).