Action in Context

Edited by
Anton Leist

Walter de Gruyter · Berlin · New York

∞ Printed on acid-free paper which falls within the guidelines of the ANSI to ensure permanence and durability.

Library of Congress Cataloging-in-Publication Data

A CIP catalogue record for this book is available from the Library of Congress.

ISBN 978-3-11-018893-6

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.d-nb.de.

Printed in Germany
Cover design: Martin Zech, Bremen.
Printing and binding: Hubert & Co GmbH & Co KG, Göttingen.

Contents

II. Action, Persons and Life

III. Action and Epistemology

Introduction: Through Contexts to Actions

Anton Leist

1. Action in Context: Why not Holism in the Theory of Action?

In one of his last articles, D. Davidson pointed out the remarkable fact that in the history of (Western) philosophy there have been only two periods of pronounced interest in what actions *are*—as opposed to the usual, moral interest in how we *ought to* act—, namely in Aristotle's time and the present day (Davidson 2005, 277). This observation could be complemented by another: it is perhaps no coincidence that during Aristotle's period of thought, as in current philosophy, traditional enthusiasm for metaphysics and epistemology was and is counterbalanced by a critical distance towards these disciplines, so that in the resulting philosophical pluralism, in addition to the pervading interest in all aspects of knowledge and the mind, it was and is possible to take an extensive look at actions. In actual fact, a wealth of literature and discussion covering an independent discipline known as 'action theory' has been open to 'analytic philosophers' for decades, but the fact that this discipline is a speciality alongside many others already underlines the theoretically ambivalent and uncertain role of action theory within the context of philosophy taken as a whole. The traditional quibbles concerning language, mind and consciousness continue to loom more heavily than such small-scale quandaries surrounding why and how human beings act.

If it is true that action theory has been pushed into a specific box by other disciplines, it is also true that it has not been overwhelmingly enthusiastic to enlarge upon its specific topic. For Davidson and the representative parts of analytic action theory he has influenced, "what actions are" is the central issue which upon second glance is typically not as harmless as it first appears. What actions are, or what constitutes an action, is usually viewed as *the* central question of action theory, usually phrased with the help of Wittgenstein's formulation: "What is left over if I subtract the fact that my arm goes up from the fact that I raise my arm?" (PU, I. §621) Many authors have accepted this question as the central

What Good is a Will?

J. David Velleman

1. Introduction

As a philosopher of action, I might be expected to believe that the will is a good thing. Actually, I believe that the will is a great thing—awesome, in fact. But I'm not thereby committed to its being something good.

When I say that the will is awesome, I mean literally that it is a proper object of awe, a response that restrains us from abusing the will and moves us rather to use it respectfully, in a way that does it justice. To say that the will is a good thing, however, would imply that having a will is better than not having one, or that using it is better than not using it—neither of which I am prepared to assert as a general rule.

Speaking metaphorically, I would say that the will is like a magic wand. In fairy tales, the character who looks upon a magic wand as an unalloyed good is destined to be sadder but wiser in the end. Being a magician isn't better than being an ordinary human, just different; and a magician must value his powers by respecting them and therefore using them appropriately, even sparingly, not by using them as much as possible.

Kant expresses something like this view in the first section of the *Groundwork:*

> In the natural constitution of an organic being—that is, of one contrived for the purpose of life—let us take it as a principle that in it no organ is to be found for any end unless it is also the most appropriate to the end and the best fitted to it. Suppose now that for a being possessed of reason and a will the real purpose of nature were his ***preservation,*** his ***welfare,*** or in a word his ***happiness.*** In that case nature would have hit on a very bad arrangement by choosing reason in the creature to carry out this purpose. For all the actions he has to perform with this end in view, and the whole rule of his behaviour, would have been mapped out for him far more accurately by instinct; and the end in question could have been maintained far more surely by instinct that it can be by reason. (Kant 1956, 62–63, Acad. 395)

Later in the *Groundwork*, Kant is going to declare that "the will is nothing but practical reason" (80; Acad. 412). So when he belittles the utility of reason in "a being possessed of reason and a will", Kant is actually belittling the utility of the will itself. If Nature had merely been looking out for our well-being, according to Kant, she would have placed our behavior under the rule of instinct, without bothering to endow us with a will at all.

An opposing view of the will can be found in Michael Bratman's classic book *Intention, Plans, and Practical Reason*. Bratman doesn't speak of the will as such, but his conception of intention, as an effective commitment to act, can be interpreted as a conception of that which issues from the will and mediates its influence on behavior. The will can therefore be conceived as the faculty of intention, in Bratman's sense of the term. Bratman argues convincingly that intention is distinct from desire and belief, and hence that the capacity for intentions is supplementary to the capacity for desire-belief motivation. In Bratman's view, this supplementary capacity "has a pragmatic rationale, one grounded in its long-run contribution to our getting what we (rationally) want".[1] Bratman thus appears to think, in opposition to Kant, that the will is designed for our preservation, welfare, and happiness.

In particular, Bratman believes that the capacity for intentions enhances the effectiveness of desire-belief motivation, by enabling it to secure greater desire-satisfaction in the long run. Many non-human animals have the capacity for desire-belief motivation, which enables them to navigate their environment intelligently in the pursuit of survival and comfort; but they cannot frame commitments of the sort that would store the results of advance problem-solving and coordinate their actions over time. The latter capacity belongs exclusively to humans, and in Bratman's view it makes them even more intelligent than other animals in the pursuit of desirable ends.

I don't want to deny that the will often provides the instrumental advantages that Bratman claims. What I deny is that the will is best understood as designed to provide those advantages. The way to figure out how the will works is not to figure out how it serves the purpose of making us more effective in pursuing desirable ends. Indeed, the way to figure out how the will works may not be to assume that it has any

1 See Bratman 1987, 35. See also p. 34: "[T]his role of intentions in providing a background framework for the weighing of desire-belief reasons is itself grounded in pragmatic considerations concerning the satisfaction of (rational) desire."

purpose at all. For in my view, the will may well be an accident, of the sort that Richard Lewontin and Stephen J. Gould have called a spandrel—that is, a feature formed not by design but by the accidental confluence of other designed features (see Gould/Lewontin 1979; Gould 1997).

2. Bratman's Theory

Bratman calls his theory of intention a "planning theory" (Bratman 1987, 9), because it takes future-directed plans as the paradigm for all intentions, including intentions directed at present or immediately forthcoming actions (4). Future-directed plans serve two purposes, according to Bratman (2 f.). First, they facilitate scheduling of deliberative effort, by enabling us to deliberate in advance, when time and materials for deliberation are plentiful, and then to store the results until the occasion for action arrives, when resources for deliberation may be scarce. Second, plans facilitate coordination of actions, by committing us to future actions whose performance we and others can count on.

Plans can serve these purposes because they embody commitments that are, in Bratman's terminology, both volitional and reasoning-centered. In their volitional aspect, plans determine our behavior: unless we change our minds, we will do what we plan, when the time arrives. In their reasoning-centered aspect, plans set the agenda for, and constrain the scope of, further deliberation (15–20).

Bratman analyzes the latter, reasoning-centered aspect of plans in terms of norms and corresponding dispositions of practical reason (29–49). First, a plan is rationally required to resist reconsideration, so as to constitute a stable making-up of our minds. Second, a plan is rationally required to be means-end coherent—that is, to be filled in with instrumental details in time for its execution. Third, our plans are rationally required to be consistent with one another and with our beliefs, so that their joint fulfillment is possible given the facts as we see them. In accordance with the first requirement, a plan to spend one's vacation in Maine, for example, tends to close the question where to spend one's vacation, and it tends to keep the lid on that question unless circumstances change significantly. In accordance with the second requirement, the same plan tends to prompt deliberation about means of getting to Maine, issuing in a plan about how to get there. In accordance with the third requirement, the plan excludes from consid-

eration any plan of spending one's vacation elsewhere, or of doing anything else that would be incompatible with being in Maine at that time.

In their volitional and reasoning-centered aspects, plans occupy a different functional role from desires. A desire to spend one's vacation in Maine might influence one's behavior but wouldn't determine it, given the possibility of countervailing desires that may be stronger; and so the desire would lack the volitional nature of a plan. A desire to spend one's vacation in Maine would also lack the reasoning-centered features of a plan, since it could rationally be abandoned at any time, could rationally be retained in the presence of a conflicting desire—say, to spend the vacation in Florida—and wouldn't oblige one to plan a means of getting to Maine. Bratman sums up these differences by saying that a desire fails to settle our course of action (18 f.). Because a plan would settle things, as a desire does not, it is entitled to be regarded as a distinct mental state, which Bratman calls intention.

So characterized, intentions differ even from those desires which set goals for our deliberations and actions. As Bratman points out, we can rationally adopt goals that are not jointly attainable, as a way to improve our chances of at least partial success, but we cannot rationally adopt incompatible intentions (chap. 8). For example, we can rationally aim at being awarded each of two fellowships, even if the award of either would automatically cancel our application for the other, but we cannot rationally plan or intend to win each of these fellowships, given their mutual incompatibility. Intentions have the function of settling the outcomes that they represent, and we cannot settle outcomes that are incompatible; we can aim at incompatible outcomes because aiming doesn't settle anything.

Bratman believes that every future-directed intention requires a present-directed intention to convey its motivational force:

> Future-directed intention involves volitional commitment because of its relation to present-directed intention: if my future-directed intention manages to survive until the time of action, and I see that this time has arrived and nothing interferes, it will control my action then. Present-directed intentions have a special relation to action, and future-directed intentions are the sort of state that will have this special relation if they survive until the time of action is seen to have arrived. (108)

Bratman identifies no other useful role for present-directed intention, and he might therefore be expected to conclude that it exists only to play this role and hence only when some formerly future-directed intention has matured to the point of requiring present action. Yet Bratman believes that whenever an agent does something intentionally, he has an intention to do something (though not necessarily the same thing; see 119 ff.).[2] Bratman is therefore committed to the existence of present-directed intentions that didn't develop out of future-directed intentions, motivating actions that are intentional but were not antecedently planned.

Bratman's view on the prevalence of present-directed intentions can be extrapolated from remarks that he makes about the association between intentional action and intention. In considering potential counterexamples to this association, Bratman mentions only "automatic and unreflective" actions, as illustrated by this example: "Suppose you unexpectedly throw a ball to me and I spontaneously reach up and catch it." (126) Bratman suggests that such spontaneous actions, which do not spring from an intention, fail to disprove the association because they are not quite intentional, either. What's revealing about these remarks is that they show how narrow is the class of actions that do not spring from an intention, in Bratman's view. Even if actions are unplanned, they spring from an intention so long as they are not "automatic and unreflective".

3. A Problem for Bratman's Theory

Now, I agree with Bratman that even in circumstances for which we have no antecedent plans, we often make up our minds before we act: we settle on a course of action before taking it, rather than taking it automatically and unreflectively. And as we have seen, settling what we are to do is characteristic of intention. Thus, Bratman is right in thinking that many unplanned actions are informed by a mental state of the sort that he calls intention.

Yet the frequency with which we form immediate intentions *de novo* raises a potentially troubling question for Bratman's theory. The question is why we bother to form intentions when there is no longer any opportunity for them to serve the functions for which, according to Bratman, the mental state of intention is designed. When the plate of cookies is held out to us, why do we make up our minds to take one? Why doesn't our hand just shoot out and grab, as it does when we spontaneously and automatically react to a sudden throw?

2 Bratman concedes the possibility of minor exceptions to this rule.

We might have planned to take a cookie if we had known in advance that we would be offered one. In that case, an intention would have placed our action under the control of deliberations for which there would be no time when the plate of cookies was already under our nose; and it might have given us grounds for skipping dessert earlier in the day, so as to coordinate our diet. But if we didn't know in advance, and consequently didn't plan, then the offer of cookies signals that we have already missed the opportunity to form a well-considered intention and to benefit from its coordinating role. We have to make up our minds on the spot, to take a cookie or not to take one. The question, to repeat, is why we bother to make up our minds. The intention to take a cookie doesn't incorporate the results of any prior deliberation, or set the stage for any further planning, or provide a basis for any coordination. There appears to be nothing for the intention to do.[3]

The same question can be raised not only about the intentions involved in unplanned actions but also about various cases of planning. Why, for example, don't I wait until I'm standing in the voting booth to settle how I am going to vote?

Of course, I have good reason for starting to think about my vote in advance: I want to be sure of making sufficient progress in my deliberations before time runs out. But having started my electoral deliberations in good time, why don't I leave them open until the moment arrives? Why conclude them prematurely? No prior preparations are needed for marking my ballot in one way rather than another, and there is no other behavior, either mine or anyone else's, that needs to be coordinated with my vote. The only foreseeable effect of planning my vote is that I will become resistant to reconsidering it, and yet there seems to be no harm, and potentially some good, in leaving it open to reconsideration, or in continuing actively to consider it in my spare time. Making up my mind in advance therefore seems to carry potential costs but no conceivable benefits. So why don't I—indeed, why doesn't everyone—show up in the 'undecided' column of pre-election polls?

3 One might think that forming an intention at least serves the purpose of terminating whatever brief deliberation we have time to undertake on such an occasion. But a spontaneous action would serve this purpose equally well; and we sometimes make up our minds without deliberating at all. What does intention add to action?

4. A Solution: Anscombe's Theory

I think that this case points to a function of intention that Bratman has omitted from his account. So long as I am still deliberating, I don't know how I am going to vote, and such uncertainty about my own future behavior is an uncomfortable state of mind, especially when the behavior in question is significant for my conception of myself. I am quite happy to leave my dinner selection undecided until the last moment, when the waiter's pencil will be poised to record it. But my vote in an election is not just a momentary matter of taste: it defines my stance on questions of importance. Not to know whom I favor in an election is to suffer an undesirable kind of ignorance about myself.

The idea that intention provides self-knowledge is a major theme of an alternative theory of intention—namely, that presented in Elizabeth Anscombe's book *Intention* (1963). Anscombe's primary claim on this topic is that an intention on which one acts embodies "knowledge without observation" of what one is doing. This claim has generally met with puzzlement and skepticism, and so I'd like to spend a moment defending it.[4]

Anscombe compares knowledge of one's intentional action with knowledge of the position of one's limbs, describing both as non-observational knowledge. She expands on the latter case as follows:

> [A] man usually knows the position of his limbs without observation. It is without observation, because nothing *shews* him the position of his limbs; it is not as if he were going by a tingle in his knee, which is the sign that it is bent and not straight. Where we can speak of separately describable sensations, having which is in some sense our criterion for saying something, then we can speak of observing that thing; but that is not generally so when we know the position of our limbs. Yet, without prompting, we *can say* it. I say however that we *know* it and not merely *can say* it, because there is a possibility of being right or wrong: there is a point in speaking of knowledge only where a contrast exists between 'he *knows*' and 'he (merely) *thinks* he knows'. (13 f.)
>
> It is not ordinarily possible to find anything that shows one that one's leg is bent. It may indeed be that it is because one has sensations that one knows this; but that does not mean that one knows it by identifying the sensations one has. (49)

4 For recent defenses of Anscombe's claim, see Hursthouse 2000; Falvey 2000. My interpretation of Anscombe has benefited especially from Falvey's paper.

> If a man says that his leg is bent when it is lying straight out, it would be incorrect to say that he had misjudged an inner kinaesthetic appearance as an appearance of his leg bent, when in fact what was appearing to him was his leg stretched out. (50)

These passages are best understood, in my view, as imprecise descriptions of the fact that, when it comes to the position of one's limbs, one is subject to perceptual appearances that involve no sensory qualities. That one's leg is bent is an appearance that doesn't feel like anything—like anything other than one's leg's being bent, that is. The appearance is empty of any sensations, such as pressure, tingling, and the like. Appearances without sensations are simply perceptual beliefs—proprioceptive beliefs, in the present case.[5]

Anscombe's claim is not that one is infallible about the position of one's limbs: she not only acknowledges but emphasizes the possibility of being wrong. Her claim is rather that when one's proprioceptive beliefs are true, they usually constitute knowledge, because they are connected in the right way with the facts that make them true. Anscombe is thus a reliabilist about proprioceptive knowledge. Her view is that one needs no sensory evidence on which to judge the position of one's limbs because the position of one's limbs generates proprioceptive beliefs via a mechanism that is generally reliable: a leg's being bent reliably causes one to believe that it is bent. I think that recognizing Anscombe's reliabilism is the key to understanding her view on non-observational knowledge of intentional action.

What corresponds to a proprioceptive belief, in the case of intentional action, is the intention itself. Descriptions of one's intentional actions are, in Anscombe's phrase, "known by being the content of intention" (53). The notion that knowledge can be embodied in an intention requires explanation and defense, which Anscombe provides as follows.

We need to distinguish between two kinds of indicative statements about the future: expressions of belief, such as "I am going to be sick", and expressions of intention, such as "I am going to take a walk" (1). If someone responds to the statement "I am going to be sick" by asking

5 This interpretation implies that the second passage quoted above contains a slight error. Anscombe should not have said "it is because one has sensations [pl.] that one knows this". What she should have said is "it is because one has sensation [sing.] that one knows this". Proprioceptive appearances are received via sensory pathways, blockage of which, as in anesthesia, leaves one unable to judge the position of one's limbs. But the sensations conveyed by these same pathways are not intrinsic to proprioception.

"Why would you do a thing like that?", he has misinterpreted the speech act, by failing to recognize it as an expression of belief rather than intention. Conversely, if someone responds to "I am going to take a walk" with "How can you tell?", he has failed to recognize it as an expression of intention rather than belief. The difference between these statements cannot lie in the former's being informative and hence potentially knowledge-conveying, since the latter is informative and hence potentially knowledge-conveying as well. As Anscombe puts it, "the indicative (descriptive, informatory) character is not the distinctive mark of 'predictions' *as opposed to* 'expressions of intention', as we might at first sight have been tempted to think" (3).

The "indicative (descriptive, informatory) character" of a statement expressing an intention indicates that the statement simultaneously expresses knowledge on the part of the speaker. When one says "I am going to take a walk", one lets the hearer know that one is going to take a walk. One's assertion is meant to provide the justification in virtue of which the hearer then knows that one is going to take a walk, and it is meant to provide that justification by virtue of expressing one's own knowledge to the same effect. Hence an expression of intention must at the same time be an expression of knowledge—of something known, in other words, by being the content of intention.

In Anscombe's view, the underlying difference between "I am going to take a walk" and "I am going to be sick", given that both express knowledge on the part of the speaker, is that the latter expresses speculative knowledge, whereas the former expresses knowledge that is practical, in the sense that it causes the facts that make it true (87). "I am going to be sick" expresses a belief that is caused by evidence of the speaker's becoming sick, whereas "I am going to take a walk" expresses an intention that causes the speaker to take a walk. Both the belief and the intention may amount to knowledge on the part of the speaker, if they are true and appropriately connected to the facts that make them true. The difference lies in the causal order of the connection. A belief amounts to knowledge if it is appropriately caused by (evidence of the) facts that make it true; an intention amounts to knowledge if it appropriately causes facts that make it true.[6]

6 Here I am using the purposely vague term "appropriately" to encompass whatever additional conditions are necessary to rule out various things that might defeat an attribution of knowledge. I am not trying to develop a precise epistemology for the knowledge embodied in intention, since the details of such an epistemology

5. A Problem for Anscombe and Its Solution

The most puzzling part of Anscombe's view is her claim that what is known by being the content of intention can include, not just how one is moving one's limbs, or what one is thereby attempting to accomplish, but the fact that one is actually accomplishing it:

> I think that it is the difficulty of this question that has led some people to say that what one knows as intentional action is only the intention, or possibly also the bodily movement; and that the rest is known by observation to be the *result*, which was also willed in the intention. But that is a mad account ....
>
> Another false avenue of escape is to say that I really 'do' in the intentional sense whatever I think I am doing. E.g. if I think I am moving my toe, but it is not actually moving, then I am 'moving my toe' in a certain sense, and as for what *happens*, of course I haven't any control over that except in an accidental sense. The essential thing is just what has gone on in me, and if what happens coincides with what I 'do' in the sphere of intentions, that is just a grace of fate. ...
>
> But this is nonsense too. (51 f.)

Thus, Anscombe believes that if one is trying to shoot a bull's-eye, intends to shoot one, and will consequently end up having shot one intentionally, then one already knows without observation that one is shooting a bull's-eye, not just that one is intending or trying to do so, or moving one's limbs with that aim. But how can the content of one's intention embody knowledge of whether the bull's-eye is going to be hit?

An answer to this question is suggested by the following passage:

> 'Why are you pumping?'—'To replenish the water supply.' If this was the answer, then we can say 'He *is* replenishing the water supply'; unless indeed, he is not. This will appear a tautologous pronouncement; but there *is* more to it. For if after his saying 'To replenish the water-supply' we can say 'He is replenishing the water-supply', then this would, in ordinary circumstances, of itself be enough to characterise *that* as an intentional action. (The qualification is necessary because an intended effect just occasionally comes about by accident.) (38 f.)

The parenthetical remark at the end of this quotation indicates that, in Anscombe's view, one can bring about an intended result without doing so intentionally if the result comes about accidentally. Hitting the bull's-eye by luck doesn't count as hitting it intentionally.

are not relevant to my purposes. All that's relevant is that the order of causation between facts and knowledge of those facts is reversed from that which is characteristic of speculative knowledge.

Now compare the two passages just quoted, and notice that the concept of accident figures in the first passage as well. In that passage, Anscombe ridicules the view that what I do, and know that I'm doing, when I intentionally move my toe does not include my toe's actually moving. What's ridiculous about that view, according to Anscombe, is the implication that my toe's moving is accidental, 'a grace of fate'. Anscombe suggests that my toe's moving is part of my intentional action, and hence part of what I know without observation, because it is a reliable result of my intending to move my toe.

The lesson to be drawn from these passages is that the reliability of the connection between intention and movement ought to be necessary both to the movement's being intentional and to the intention's being knowledge, according to Anscombe's reliabilist epistemology. Speculative knowledge, which is caused by the facts, must be caused by them via some reliable mechanism, or it doesn't qualify as knowledge, after all. If knowledge can also be connected to the facts by virtue of causing them, this connection must be reliable as well. Unless an intention with the content "I'm going to move my toe" reliably causes my toe to move, it won't amount to practical knowledge. If my toe moves accidentally, I will neither have moved it intentionally nor known that I was moving it.

Anscombe's view, then, is that what is done intentionally and what is known by being the content of intention will tend to coincide, because they are constituted by the same reliable connection and undermined by the same failures of reliability. If your intention to hit the bull's-eye doesn't amount to knowledge of the fact that you're hitting it, the reason is probably that you can't reliably hit the bull's-eye; and in that case, your hitting it won't qualify as intentional, either. If your hitting the bull's-eye does qualify as intentional, then you must have a reliable way of hitting the bull's-eye, in which case your intention to hit it is reliably connected to that result and will probably amount to knowledge.[7] What you do

7 What is done intentionally is not perfectly coextensive with what is known without observation. Whether something is done intentionally depends on the reasons for and against doing it—especially against. An unreliable but lucky marksman may not be credited with hitting the bull's-eye intentionally, but he may be blamed for hitting the President intentionally, simply by virtue of trying to hit the President and despite his lack of a reliable method for doing so. (This example is due to Gilbert Harman. Thanks to Gideon Rosen for reminding me of it.) In neither case does he know without observation that he is hitting his target. If his target is the President, then, what he does intentionally exceeds what he knows without observation. As I understand Anscombe, she believes that

intentionally, you tend to do knowingly, because intentional action and knowledge are generally two ends of the same chain of reliable causation.

6. Intention and Belief

What I want to preserve in Anscombe's theory is the claim that intention can embody knowledge of the resulting action, knowledge that is practical in the sense of causing what it represents.[8] Anscombe's view that intention embodies knowledge doesn't lead her to assert that intention requires belief; rather, she seems to think that knowledge can subsist either in a true belief or in a fulfilled intention. But in viewing intention as potentially embodying knowledge, she clearly conceives of it as similar to belief in being a cognitive commitment to the truth of its propositional object.

Anscombe is therefore at odds with philosophers who deny that intention entails belief, because they do so on grounds that would rule out its being any sort of cognitive commitment that could embody knowledge. Consider, for example, Bratman's argument against the thesis that intention entails belief:

> First, there seem to be cases in which there is intention in the face of agnosticism about whether one will even try when the time comes. I might intend now to stop at the bookstore on the way home while knowing of my tendency toward absentmindedness—especially once I get on my bike and go into 'automatic pilot'. If I were to reflect on the matter I would be agnostic about my stopping there, for I know I may well forget. It is not that I believe I will not stop; I just do not believe I will.
>
> Second, there seem to be cases in which there is intention in the face of agnosticism about whether one will succeed when one tries. Perhaps I intend to carry out a rescue operation, one that requires a series of difficult steps. I am confident that at each stage I will try my best. But if I were to reflect on the matter, I would have my doubts about success. (Bratman 1987, 37 f.)

Here Bratman argues that intending to stop at the bookstore, or to carry out the rescue operation, is compatible with remaining agnostic about the truth of the corresponding descriptions, "I am going to stop at the bookstore" or "I am going to carry out the rescue". He seems to regard the intentions as conative commitments to *making* these future-tense descriptions true but not as cognitive commitments to their *being* true. But then if the agent does stop at the bookstore or carry out the rescue on the basis of his intention, he cannot claim to have known that he was going to, since he can hardly claim to have known things with respect to which he was cognitively uncommitted. Bratman must therefore deny Anscombe's claim that intention embodies knowledge of the resulting action.

This denial leaves much of Bratman's functional account unmotivated. Why, for example, should an agent be rationally obliged to arrange means of carrying out an intention, if he is agnostic about whether he will in fact carry it out? Suppose that I form an intention to fly to Chicago next Tuesday, well knowing that I often forget to take trips that I have planned. (I am even more forgetful than Bratman.) Buying a ticket for my flight to Chicago will turn out to have been a waste of money if I forget to take the trip. If going to Chicago were merely a goal, in competition with my other goals, then the rationality of buying a ticket would depend on the expected benefits of attaining this goal, by buying the ticket, and the expected costs of foregoing other goals, by investing in the ticket and then possibly failing through forgetfulness to attain even this one. Part of the reasoning-centered function of intention, in Bratman's view, is to cut through such cost-benefit calculations by generating a categorical requirement to identify means of doing what is intended. But why should I be categorically required to invest in means whose benefits I am not yet prepared to believe in? If I am still entertaining the possibility that a ticket will go to waste, why shouldn't I weigh its expected benefits against those of alternative investments?

Bratman also believes that a plan of flying to Chicago on Tuesday rationally constrains my subsequent practical reasoning by ruling out options inconsistent with my being in Chicago—for example, the option of accepting your invitation to dinner on the same day. But if, for all I know, I might forget my flight and still be here on Tuesday evening, then it would certainly be advantageous to have a dinner appointment to replace whatever activities I would then be missing in Chicago. If I form two intentions—to fly to Chicago and to have dinner with you here—then I will have a better chance of doing something that I intend, and a better chance of enjoying myself on Tuesday, just as I can raise my

intentional action is invariably known without observation. On this point, among others, I believe that Anscombe is mistaken.

8 I do not want to preserve either Anscombe's reliabilism or her view that intentional action can be analyzed in terms of intention. On the problem of analyzing intentional action, see the preceding note. For an evidentialist rather than reliabilist conception of knowledge without observation, see Velleman 1989, Part I.

chances of winning a fellowship by applying for two of them, despite the known impossibility of winning both.

As I have explained, Bratman argues that rationality forbids such inconsistencies between our intentions, though not between our goals. But why should my intentions be subject to a requirement of consistency if I can remain cognitively uncommitted to their truth? If I am agnostic as to whether I'll be in Chicago on Tuesday evening, why should my plans for Tuesday evening have to be consistent with my being there?

Of course, the requirement of consistency among intentions is essential to their role in coordinating behavior, since behavior cannot be coordinated with inconsistent states of affairs. Yet the coordinating role of intentions would itself come into doubt if intentions did not involve a cognitive commitment. When an intention coordinates behavior, the agent and his associates proceed on the assumption of its being executed—which would be an odd way to proceed if the agent himself were agnostic on the question. If I am agnostic as to whether I will be in Chicago on Tuesday, why should anyone plan or act on the assumption of my being there? And why should anyone hesitate to plan or act in ways inconsistent with that eventuality?

Bratman might say that intentions must be consistent because they are agglomerative, in the sense that rationality favors combining an intention to *A* and an intention to *B* into a single intention to *A* and *B*. Surely, I cannot rationally intend to be in Chicago and have dinner with you here at the same time. This argument would only shift the problem, however, by raising questions about the requirement of agglomerativity. Why are attitudes agglomerative? Beliefs are agglomerative because they aim to fit the world, of which there is just one, in whose complete characterization the contents of all true beliefs are conjoined. The rational pressure to conjoin beliefs is a pressure to fuse them into a single characterization of the single world that all of them aim to fit. An agglomerativity requirement would be equally appropriate for intentions if they had to be jointly executable in a single world. As before, however, the question is why my intentions must be jointly executable, if I can be agnostic as to whether they will be executed. If I can have a plan without believing that it will be carried out in this world, why should I confine myself to planning for a world in which it is carried out?

Note finally that Bratman must deny, what is axiomatic for Anscombe, that the natural expression of an intention to do something is the assertion that one will do it. For if Anscombe were right about the natural expression of intention, then Bratman's view would imply that someone could sincerely say "I am going to act" without believing it, since the assertion might still be sincere as an expression of intention. What's more, if Anscombe were right, then the statement "I am going to act but I don't believe it" would express an intention while reporting the lack of a corresponding belief—a combination of mental states that should not be paradoxical, according to Bratman. Yet such a statement is clearly an instance of Moore's paradox, because someone who says "I am going to act" without believing it is guilty of insincerity by ordinary conversational standards. Unless Bratman wants to revise those standards, he will have to say that "I am going to act" is not the natural expression of an intention. Perhaps, then, he will say that the natural expression of an intention is "I intend to act".

Bratman might also point out that "I intend to" can be a way of avoiding an outright affirmative answer to the question "Are you going to act?". But avoiding an outright affirmative answer does not necessarily indicate a lack of belief. After all, one can also avoid an outright affirmative answer by saying "I believe so", which positively requires that one believe that which one is trying to avoid asserting. Hence expressions used to avoid an outright assertion cannot in general be interpreted as indicating a lack of belief.

Now, if one answers the question "Are you going to act?" with "I believe so", then failing to act will not make one guilty of having told a lie, but it will entail that one has misled a listener who was inclined to adopt or defer to one's point-of-view. Surely, the same will be true if one answers the question "Are you going to act?" with "I intend to". If one then fails to act, one can retrospectively defend this reply by saying "Well, I only said that I intended to"; but this defense will be just as narrowly legalistic as "I only said that I believed so". In either case, one will have raised an expectation of action in a listener who was inclined to adopt or defer to one's point of view. If intention didn't entail belief, however, then the statement "I intend to act" wouldn't license even a credulous listener to expect action. "I intend to" would be similar to "I hope so"—that is, the report of an attitude that warrants no expectation on the part of the listener because it requires none on the part of the speaker.

7. What is Intention For?

In my view, these incongruities in Bratman's account can ultimately be traced to the fact that for many of the purposes that Bratman assigns to intention, knowledge of one's forthcoming actions is what actually does the work. Knowledge of one's forthcoming actions is what provides the basis for coordination, and it is what constrains the options available to further practical reasoning. Intention would be ill-suited to serve these purposes if it were compatible with agnosticism about one's forthcoming actions, as Bratman believes. Hence intention would better serve the purposes for which it is designed, according to Bratman, if it were capable of embodying self-knowledge in the manner envisioned by Anscombe.

I have now argued that the epistemic role of intention supports the pragmatic functions elucidated by Bratman. Yet I previously considered cases in which no pragmatic purpose is served by the self-knowledge embodied in an intention. Making up one's mind to take a cookie that has been offered, or making up one's mind how to vote in the next election, puts one in the position of knowing what one is going to do, but not of being able to coordinate or plan any related activities. These cases therefore raise the question whether the epistemic function of intention is entirely subservient to its pragmatic functions.

On the one hand, these cases might exemplify a re-purposing of intention—a use of the attitude for a purpose that, in the normal case, is ancillary to its ultimate purpose but, in the present cases, is the only purpose at stake. That is, intention might have been designed to embody self-knowledge for the sake of facilitating coordination, but it might then be used, on occasion, for the sake of self-knowledge alone, when coordination isn't necessary. On the other hand, the pragmatic uses of intention might be the ones that exemplify re-purposing. That is, having been designed to embody self-knowledge for its own sake, intention might turn out to yield coordination as a fortuitous byproduct.

If these alternative hypotheses seem too fanciful in attributing the human will to intelligent design, never mind: I am not going to argue for either of them. I suspect that the human will is an accident—an absurd and wonderful accident. But as I shall argue, it is an accident that more closely approximates the second rather than the first of the envisioned designs. Attributing the will to accident may seem just as fanciful as attributing it to intelligent design. Any hypotheses about the origins of the will must be closer to creation myths than to scientific theories. Why entertain creation myths at all?

My justification for entertaining creation myths about the will is that such myths are implicit in the functionalist moral psychology that Bratman practices. Bratman seeks to characterize the attitude of intention in terms of the function that the attitude typically serves and the rational standards for serving that function. In this effort to identify the function of intention, he is guided by the assumption that it is a function, not merely in the value-neutral, philosophical sense that it is a causal role, but in the evolutionary sense that it confers some benefit on intention-forming creatures. He seeks to characterize intention, not just by observing what an intention does, but by figuring out what an intention might do that needs doing. He thereby adopts what Daniel Dennett has called the design stance, assuming for methodological purposes that the attitude of intention will turn out to be as an intelligent designer would have fashioned it.

My own creation myth about the will, which follows shortly, is meant as an antidote to this methodological assumption.[9] My purpose in telling this myth is not so much to gain acceptance for it as to show that the standard myth is optional. Our theorizing about the will should not necessarily be guided by the assumption that this practical faculty is designed to confer practical benefits. It often does confer practical benefits, of course, but those benefits may be accidental and should therefore not be assumed to have governed its design.

8. The Will as a Spandrel

On many other occasions I have defended a particular explanation of the causal relation by which an intention is reliably connected to the facts that constitute its fulfillment. Let me take a moment to rephrase that explanation in the terms that I have adopted for the present occasion.[10]

An agent has a standing desire to understand what he does, and the best way for him to satisfy this desire is not to do anything until he is prepared to understand it, and then to do that which he is prepared to

9 For direct arguments against the assumption, see Fodor 2000, chap. 5.

10 An important difference in terminology is that, on the present occasion, I have avoided claiming that intention entails belief in what is intended. Rather, I have said that intention entails a cognitive commitment to the truth of what is intended—a formulation that leaves open whether that cognitive commitment should be called a belief.

understand.[11] The act-description embedded in an intention, such as "I am going to help myself to a cookie", conveys information about what the agent is up to, information that implicitly explains why he is behaving as he is—or, more precisely, would explain why he was behaving as he would be if he now proceeded to reach for something on the platter before him. This act-description thus prepares the agent to understand the behavior of reaching toward the platter and, conversely, to be puzzled by alternative behaviors. ("If I'm helping myself to a cookie, then why am I scratching my shin instead of reaching toward the platter?") The agent's desire to understand what he's doing inhibits him from acting until he has committed himself to the truth of some such description, whereupon it reinforces his antecedent motives for acting in accordance with that description, with the result that he does what he's prepared to understand and consequently understands what he's doing.

The agent can commit himself to the truth of any one of the act-descriptions that he has some antecedent motives for satisfying, because that commitment will reliably enlist reinforcement for the motives favoring satisfaction of the description. The agent's commitment to the truth of the description, being a reliable cause of the facts that will make it true, then constitutes knowledge of those facts—practical knowledge, in Anscombe's terminology. The commitment therefore plays the functional role of an intention.

This conception of intention has the consequence that to have a will—that is, a capacity for intentions, so conceived—is to have freedom, in the form of an open future. Because my cognitive commitments to the truth of propositions about my forthcoming behavior reliably cause those propositions to come true, I can commit myself either to the proposition that I will help myself to a cookie or to the proposition that I won't, and my commitment will amount to knowledge in either case. With respect to my own intentional actions, then, I can invent my knowledge of the future rather than discover it, and my future is therefore open in a sense in which it is closed with respect to other events. My will therefore gives me genuine options, in the form of alternative truths of which I could invent either one.[12]

11 The explanation offered in this paragraph involves empirical assumptions that, in my experience, strike psychologists as obvious while striking philosophers as bizarre. I summarize some of the psychological evidence in Velleman 2000a.

12 For further elaboration on this theme, see Velleman 2000b.

My present point about this account of intention is that it posits nothing more than the predictable consequences of two motivational states whose utility in the design of a creature is far more general than that of the human will. The first of these states is the creature's desire or drive to understand what's going on in its environment. The second of these states is the realization on the part of the creature that it is one of the inhabitants of that environment, a realization that revolutionizes the creature's conceptual scheme, by giving it an objective conception of itself. If a creature is driven to understand what goes on in its world, and if it also realizes that its own behavior is among those goings-on, then it will be driven to understand its own behavior, and it will have acquired the capacity for intentions.

Of course, the creature may not immediately realize that it can understand its behavior by doing what it is prepared to understand, but that realization must naturally come, as the creature frames ideas of things to do, does them, and sees that it then understands what it is doing. Discovering its capacity for intentions is simply a matter of discovering how the egocentrically conceived world of doing things is connected to the objectively conceived world of things understood—which is rather like a dog's discovering that what it wags is the tail visible to its rear. Once the creature finds its own tail, in the form of its behavioral contributions to the intelligible world, it is ready to start forming intentions.

In short, the will emerges as a byproduct of curiosity plus self-awareness, which are fundamental endowments of human nature, designed for purposes far more general than scheduling deliberation and facilitating coordination. Once Nature had made self-aware inquirers, she didn't have to give them a faculty to serve these specific purposes, because they already had one.

I have already suggested one moral of this creation myth. The moral is that if we adopt the design stance in theorizing about the will, we should remember how medieval architects designed spandrels: by accident. We shouldn't assume that the will is perfectly suited to the purposes that it serves, because it may not have been designed for those purposes, or indeed designed for any purposes at all.

A further moral that one might draw from this creation myth is that the will and its freedom are less central to human nature than we previously thought. I decline to draw this moral for two reasons. First, I think that a feature acquired by accident can become central to the nature of a creature by shaping its subsequent evolution, as I imagine that the

will may have shaped ours. Second, I think that centrality is partly an evaluative matter, and that we are entitled to value our accidental endowments more than our designed features, if we like.

I realize, however, that this question is one on which we are likely to bear out William James's remark that philosophical disagreements are at bottom differences of temperament.[13] The reason why I am willing to view our free will as an accident is that its being an accident wouldn't undermine my sense of its value, mine being a temperament that welcomes absurdity. Those whose temperament favors order in the human condition may find that the idea of the will's being an accident dampens their enthusiasm for it.

My creation myth does have one additional moral, owing to the fact that human life goes better when informed by an understanding of human nature. We do better at managing our lives if we understand the kind of creature whose life we are managing. In the case of the will, I think, we stand to benefit in the management of our lives from an understanding of how it is related to our motivational psychology, over which it exerts a kind of supervision and control. Such an understanding can indicate how smooth or stormy we should expect the relations between these faculties to be.

On this question, Bratman's rationalist conception of the will counsels optimism. Our capacity for intentions is an enhancement of our capacity for desire-belief motivation, according to Bratman, designed to make us more effective in getting what we want. Of course, the will must sometimes restrain desire-belief motivation, but it does so for the sake of greater desire-satisfaction in the long run. In the long run, then, we can expect harmony between willing and wanting.

My absurdist conception of the will is less optimistic. I agree with Bratman that our capacity for intention enables us to realize gains in desire-satisfaction, but I see no reason to assume that it is perfectly or even moderately well suited to this role. Our nature as motivated animals is sometimes enhanced by our nature as self-aware inquirers, but these two sides of our nature can also be at odds—fundamentally at odds, in a way that isn't resolved in the long run. Managing a human life therefore entails striking an essentially uncomfortable balance between wanting and willing, or between motivation and intention.

13 James makes this remark in the first lecture of *Pragmatism* (James 1995, 2).

9. Conclusion

In closing, I'll offer two illustrations of this balance and its discomforts. The first is an instance that favors willing, the second an instance that favors wanting.

My conception of the will suggests that we should not be surprised by the conflict between desire-satisfaction and the requirements of morality. If the latter requirements inhere, as Kant believed, in the very structure of the will, then we should not expect them to be desire-syntonic, because the will is not just an enhancement to desire-belief motivation. The will makes its own demands, which can be fortuitously instrumental to those of desire-belief motivation, but only sometimes and only imperfectly. To seek desire-based reasons for being moral is therefore a mistake. What we should seek instead are ways of harmonizing these two sides of our nature, by cultivating desires that are at least minimally syntonic with the will.

My conception of the will also suggests that, although it can enhance desire-belief motivation, we should also expect it to get in the way, and not just when it malfunctions. Simply by being what it is, the will disrupts many activities that go best when left to the guidance of our motivational nature. I am not saying, with Kant, that instinct would do better than practical reason at securing our very preservation, welfare, and happiness; but I am saying that instinct does better in guiding many activities.

The activities that I have in mind have been characterized by the psychologist Mihaly Csikszentmihaly as providing an experience that he calls "flow" (Csikszentmihaly 1990). He conducted research in which subjects were prompted to record their activities, and their feelings about them, at regular intervals during the day. He then identified a category of "optimal experiences" that occur in the course of highly challenging activities in which the subject exercises appropriate skills. Czikszentmihaly writes:

> When all of a person's relevant skills are needed to cope with the challenges of a situation, that person's attention is completely absorbed by the activity. There is no excess psychic energy left over to process any information but what the activity offers. All the attention is concentrated on the relevant stimuli.
>
> As a result, one of the most universal and distinctive features of optimal experience takes place: people become so involved in what they are doing

> that the activity becomes spontaneous, almost automatic; they stop being aware of themselves as separate from the actions they are performing. (53)

As Czikszentmihaly goes on to explain, this loss of self-consciousness "does not involve a loss of self, and certainly not a loss of consciousness, but rather, only a loss of consciousness *of* the self" (64). Not surprisingly, it also involves a suspension of deliberation:

> In normal life, we keep interrupting what we do with doubts and questions. 'Why am I doing this? Should I perhaps be doing something else?' Repeatedly we question the necessity of our actions, and evaluate critically the reasons for carrying them out. But in flow there is no need to reflect, because the action carries us forward as if by magic. (54)

Activities that provide this experience can of course be initiated deliberately, but then they follow a course determined by sequences of stimulus and skilled response. They require intelligence, but they require it to be expressed directly in behavior rather than in prior decisions; they require us to think with our bodies, without pausing to make up our minds. For this reason, the will cannot enhance these activities; it can only disrupt them.

I think that the phenomenon observed by Czikszentmihaly provides indirect support for my conception of the will. To begin with, it illustrates the intimate connection between willing and self-awareness, both of which must be suspended if we are to 'find flow' (see Csikszentmihaly 1997). What's more, the difficulty of finding flow illustrates the self-assertive nature of the will. In order to enter the state of optimal experience that Czikszentmihaly describes, we need to find an engrossing activity that will draw us out of ourselves and thereby silence our deliberations and planning. The elusiveness of such experiences should remind us that will is not just a capacity for control but a drive toward control, which must sometimes be beguiled into relinquishing its hold on our behavior. Having a will entails having a tendency to be willful—a tendency that cannot be explained by conceptions of the will as a passive instrument of practical reasoning. Unlike that passive instrument, the will is a mixed blessing.

Bibliography

Anscombe, G. E. M. (1963): *Intention*, Ithaca: Cornell University Press.

Bratman, Michael (1987): *Intention, Plans, and Practical Reason*, Cambridge, MA: Harvard University Press.

Csikszentmihaly, Mihaly (1990): *Flow: The Psychology of Optimal Experience*, New York: Harper and Row.

— (1997): *Finding Flow: the Psychology of Engagement with Everyday Life*, New York: Basic Books.

Falvey, Kevin (2000): "Knowledge in Intention", in: *Philosophical Studies* 99, 21–44.

Fodor, Jerry (2000): *The Mind Doesn't Work That Way: The Scope and Limits of Computational Psychology*, Cambridge, MA: MIT Press.

Gould, Stephen J./Lewontin, Richard C. (1979): "The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme", in: *Proceedings of the Royal Society of London* 20, 581–598.

Gould, Stephen J. (1997): "The Exaptive Excellence of Spandrels as a Term and Prototype", in: *Proceedings of the National Academy of Science* 94, 10750–10755.

Hursthouse, Rosalind (2000): "Intention", in: Teichmann, Roger (ed.): *Logic, Cause and Action*, Cambridge: Cambridge University Press, 83–105.

James, William (1995): *Pragmatism*, New York: Dover Publications.

Velleman, David J. (1989): *Practical Reflection*, Princeton: Princeton University Press.

— (2000a): "From Self-Psychology to Moral Philosophy", in: *Philosophical Perspectives* 14, 349–377.

— (2000b): "Epistemic Freedom", in: Velleman (2000), 32–55.

— (2000c): *The Possibility of Practical Reason*, Oxford: Oxford University Press.

Kant, Immanuel (1956): *Groundwork of the Metaphysics of Morals*, translated by H. J. Paton, New York: Harper & Row.