Level 2

Intention

Élisabeth Pacherie, Centre National de la Recherche Scientifique, Paris, France

CONTENTS

The concept of intention is used to characterize both actions and states of mind. To a first approximation, we classify actions as intentional if they are both purposeful and voluntary and we say that someone intends or has the intention to do something if he is settled on so acting.

INTRODUCTION

0144.001 The notion of intention lies at the intersection of the two domains of action theory and philosophy of mind. On the one hand, the concept of intentional action is at the heart of action theory. Providing an analysis of the distinction between intentional and non-intentional action is one of its main aims, with important consequences for work on moral and legal responsibility. On the other hand, the notion of intention is also a central element in the web of concepts used to characterize the mind and the various kinds of states and attitudes that belong to the mental realm.

0144.002 It is not surprising that intentions figure so prominently in both fields of inquiry. A capacity for intentional behavior is the mark of intelligent agency. Intentional behavior is behavior the correct explanation of which cannot be purely mechanistic but requires us to make reference to the agent's attitudes, beliefs and desires, and reasons for acting in a certain way. There is little doubt, therefore, that there is a close connection between intentions and intentional action. Yet, what exactly an intention is, what qualifies as an intentional action, and what is the connection between the two, are all matters of considerable philosophical debate.

ARE THERE SUCH THINGS AS INTENTIONS?

0144.003 On one simple and prima-facie plausible view, the connection between intentional actions and intentions is to be construed as follows: someone does something intentionally if and only if he or she has the intention to do it and that intention causes him or her to do it. Yet this simple view can be criticized in a number of ways. One objection that can be raised is that it involves the unnecessary postulation of intentions as distinctive states or events. The philosophers who raise this objection (Anscombe, 1963; Davidson, 1980, essay 1) consider that for an action to be intentional is for it to be explainable by reasons the agent has for acting that way. In other words, an action is intentional if it is explainable in terms of beliefs and desires of the agent, typically a desire for, or more generally, a 'pro' attitude towards, a certain outcome and a belief that acting in such or such a way would promote attainment of this outcome. On this view, to say that someone acts intentionally or with an intention is simply to say that some appropriate relation obtains between the agent's beliefs and desires and the agent's actions. Talk of intentions as distinct states or events is therefore superfluous.

0144.004 This relational analysis of intentions has in turn been criticized. As several philosophers, including Davidson himself, have pointed out (Davidson, 1980, essay 5; Bratman, 1987; Harman, 1986; Velleman, 1989), this analysis is inapplicable to intentions concerning the future, intentions that, although we may now have them, are not yet acted upon, indeed might never result in action. Acknowledging the existence of future-directed intentions forces one to admit that intentions can be states separate from the intended actions or from the reasons that prompted the action. But, as Davidson notes, once this is admitted there seems to be no reason not to allow that intentions of the same kind are also present in all or at least most cases of intentional actions. Once intentions are acknowledged as separate mental states and not just relational constructs, a new issue arises. Can these states be given a reductive analysis – are they

assimilable to complexes of desires and beliefs, to special kinds of beliefs or judgments – or do they form a *sui generis* and irreducible class of mental states?

BELIEFS, DESIRES AND INTENTIONS

0144.005 Although they acknowledge the existence of states of intending, a number of philosophers want to resist the idea that intentions constitute some new, irreducible kind of attitude. They see beliefs and desires as the two fundamental kinds of mental state and think they can provide an analysis of intentions that remains within the bounds of the belief–desire framework. These elaborations of the belief–desire model have taken several forms.

0144.006 One approach sees intentions as reducible to combinations of beliefs and desires. It is commonly admitted that intentions have both cognitive and motivational components. Belief–desire reductionism about intentions suggests that the cognitive component of an intention can be identified with a belief and its motivational component with a desire. In particular, it stresses the link between an intention to perform an action and a belief that one will, and claims not only that intending to do something entails believing that one will do it but that the intention consists in this belief. Yet, not all our beliefs about our future actions constitute intentions: some may simply be predictions. A proponent of this view must therefore introduce further constraints that beliefs should meet to qualify as intentions. Here, the motivational component of intentions becomes highly relevant. Thus, Davis (1984) suggests that an agent's intention to *A* consists in believing that the agent will *A* because he or she desires to *A*, together with believing that this desire will motivate him or her to *A*. Alternatively, Velleman (1989) identifies intentions with self-fulfilling beliefs that are motivated by a desire for their fulfilment and that represent themselves as such.

0144.007 Another approach, advocated by Davidson (1980, essay 5), rejects this belief–desire reductionism. In particular, Davidson does not accept the claim that intending to *A* entails believing that one will *A*. Yet, his approach maintains a close correspondence between intentions, beliefs and desires. According to Davidson, intentions should be viewed as a special kind of evaluative judgment different from the evaluative judgments that correspond to desires to act. Desires to act are what Davidson calls 'prima-facie' judgments, judgments that actions of a certain kind are desirable insofar as they have a certain attribute. As such, prima-facie judgments are not directly associated with actions, for it is not reasonable to perform an action on the sole ground that it has some desirable feature. By contrast, an intention to act is what Davidson calls an 'all-out' judgment, an unconditional judgment, made in the light of one's beliefs, that a certain action is desirable. In making an all-out judgment as opposed to a prima-facie judgment, we settle on a course of action. Intentions are thus associated with actions in a way that mere desires are not. This analysis of intentions introduces a new element, an all-out judgment in the belief–desire framework, yet it avoids having to postulate a completely *sui generis* kind of mental entity. Intentions, together with other pro attitudes, are deemed to belong to the general class of evaluative judgments.

0144.008 Each of these proposals has encountered specific objections. For instance, it is not clear how, on Davidson's view, an agent would decide between two courses of action that are equally desirable all things considered but are not mutually compatible. It is also disputed whether intending to *A* need always entail believing that one will, not to mention whether it consists in this belief. A more general criticism of such analyses is that they are too backward-looking. They focus on how intentions are arrived at, and therefore fail to appreciate important aspects of the role of intentions.

THE FUNCTIONS OF INTENTIONS

0144.009 Opponents of reductive approaches to intentions tend to emphasize a number of functions plausibly attributed to intentions. They argue that intentions have their own complex and distinctive functional role that warrants considering them as forming an irreducible kind of psychological state, on a par with beliefs and desires. Among these functions are some that are most apparent when one considers future-directed intentions and the way they function in practical reasoning. Philosophers pursuing this line of inquiry (Bratman, 1987; Harman, 1986; Mele, 1992) propose to articulate what is at stake in being settled on or committed to some course of action as a result of having formed or acquired an intention. Thus, Bratman stresses three aspects of the roles played by intentions in the period between their initial formation and their eventual execution. Firstly, intentions have a characteristic stability or inertia: once we have formed an intention to *A*, we will not normally continue to deliberate whether to *A* or not; in the absence of relevant new information, the intention will resist reconsideration. Secondly, during this period between the formation of an intention and action,

we will often reason from such an intention to further intentions, for example, reasoning from intended ends to intended means or preliminary steps. Thirdly, this intention will itself often be an element in a larger plan: it will constrain the other intentions we may form and will help us achieve both intrapersonal and interpersonal coordination.

0144.010 The role played by intentions in practical reasoning is especially salient when we consider future-directed intentions. The executive aspect comes to the fore when we concentrate on immediate intentions. Here the philosophical focus is on the role of the intention in the production of the corresponding action (Brand, 1984; Mele, 1992). This role is both cognitive and motivational. Many philosophers agree that intentions are motivating causes of actions. Furthermore, this role as motivators is not only to trigger or initiate the intended action but to sustain it until completion. If, say, while on my way to my office, I ceased to intend to go my office, this would bring my action to a halt. Intentions also involve the guidance and monitoring of action. The cognitive component of an immediate intention to *A* incorporates a plan for *A*-ing, a representation or set of representations specifying the goal of the action and how it is to be arrived at. Moreover, intentions may be assigned a monitoring function, a capacity to detect progress towards the goal or to detect and correct deviations from the course of action as laid out in the guiding representation.

0144.011 An exclusive focus on the practical reasoning aspect of future-directed intentions may give the impression that only agents endowed with sophisticated cognitive abilities may be capable of intentional actions. Yet, some philosophers think that it would be unduly restrictive to limit the sphere of intentional actions to premeditated or planned actions. They are ready to accept that some actions that we perform relatively spontaneously and without prior deliberation, or some actions performed by creatures devoid of the sophisticated capacities necessary for forward planning, still qualify as intentional actions. Concentrating on the executive aspect of immediate intentions may be a way to make sense of this idea. We will return to this question, after considering one further idea related to the executive aspect of intentions.

ARE INTENTIONS SELF-REFERENTIAL?

0144.012 Reflection on the executive functions of intentions has led some philosophers (Harman, 1986; Searle, 1983) to claim that intentions involve a kind of causal self-referentiality. Put very simply, the self-referentiality thesis is the thesis that an intention to *A* is the intention that this very intention be the cause of the agent's *A*-ing. Two main considerations may be adduced in favor of this thesis.

Firstly, if one accepts a causal analysis of intentional actions according to which an action's being intentional depends on its being appropriately caused by a corresponding intention, one must find a way to exclude cases of deviant causal chains. An example, adapted from Harman, may illustrate the nature of the problem. Mabel may intend to kill Ted by running him over in her car. Ted happens to be walking by when she backs out of her driveway and she runs him down, killing him without even seeing him. Here, although Mabel intended to kill Ted by running him over and killed him in this way, her action is not intentional: she does not do what she intends. The suggested moral is that for an act of *A*-ing to be intentional, it is not enough that it be caused by an intention to *A* or even an intention to *A* in a certain way, but the intention must be an intention that this very intention lead one to *A* in a certain way. 0144.013

This conclusion has been challenged by Mele (1992), who argues that what accounts for Mabel's action being unintentional may be a function of the plan component of her non-self-referential intention. For instance, he would say that Mabel's action is not intentional because her intention did not incorporate a plan to run down Ted while backing out of her driveway. Mele insists that it not enough for an action to be intentional that it fits some intended plan; it must be guided by the plan. It is unclear, however, what is involved in the notion of guidance and whether it can be defined so as to exclude causal deviance without invoking the idea of causal self-reference of intentions. 0144.014

A second kind of consideration supporting the self-referentiality thesis refers to the conditions of satisfaction of intentions. Searle (1992) argues that the conditions of satisfaction of an intention include that the intention itself play a causal role in bringing about the rest of its conditions of satisfaction. Moreover, Searle claims that this causal self-referentiality condition is internal to the content of the intention. Thus, the content of my intention to raise my arm should include that I raise my arm as a result of this very intention. 0144.015

This construal of the self-referentiality thesis has also been challenged. It has been objected that, even if self-referentiality is part of the conditions of satisfaction of intentions, there is no obvious reason 0144.016

why this condition of satisfaction should be reflected in the content rather than attached to the psychological mode itself or considered part of the background conditions. It has also been objected that Searle's account of self-referentiality is psychologically implausible in requiring of the intenders unrealistically strong cognitive and conceptual capacities. Recently, attempts have been made to dispel the charge of conceptual oversophistication by showing how self-referentiality can be construed as an implicit element of the content of intentions, reflecting architectural and circumstantial facts about how the intention connects with action (Pacherie, 2000; Roth, 2000).

THE CONNECTION BETWEEN INTENTIONS AND INTENTIONAL ACTIONS

0144.017 The discussion of self-referentiality above raised the question whether it is sufficient for an action of *A*-ing to be intentional that the agent have the intention to *A* or whether it is furthermore required that the intention be self-referential. We can also ask the converse question, namely whether it is always necessary for an action of *A*-ing to be intentional that the agent intends to *A*. In other words, are our folk-psychological notions of intention and intentional action sufficiently univocal that philosophers may hope to provide a unified analysis of their connection? We will now examine three kinds of cases that seem, in different ways, to shed light on the connection between intentional actions and intentions.

Expected side-effects and responsibility

0144.018 Both Harman (1986) and Bratman (1987) have argued that certain actions we perform are intentional despite not being intended because they are expected side effects of some (larger) action we intend to do. For instance, Bratman claims that if someone intends to run a marathon and believes that he will thereby wear down his running shoes, his action of wearing down his running shoes is intentional despite not being intended. Similarly, Harman suggests that in firing his gun to try to kill a soldier, a sniper who knows that he will thereby alert the enemy to his presence does not intend to alert the enemy, yet does so intentionally. If this analysis is accepted, we are forced to conclude that there is a gap between intentions and intentional actions. One way to try to avoid this conclusion, suggested by Mele and Moser (1994), involves drawing a threefold distinction among intentional actions, unintentional actions and non-intentional actions. Since the sniper does not unknowingly or accidentally alert the enemy, the action is not unintentional, but since, on the other hand, alerting the enemy was not a goal or aim of the agent, nor is it intentional. Non-intentional actions would then constitute a middle ground between intentional and unintentional actions.

However, as both Bratman and Harman note, the 0144.019
tendency to consider an expected side effect as an intentional action depends on the significance of the action and on the reasons one might have not to do it. We can contrast alerting the enemy with another expected side effect of firing the gun, such as thereby heating the barrel. Although we have no inclination to say that the action of heating the barrel was intentional, we may want to say that alerting the enemy was intentional insofar as the sniper had reasons not to want to alert the enemy and yet chose to do so. Here the topic of intentional action becomes entwined with issues of responsibility and moral assessment. Maintaining that an action was intentional may justify holding the agent responsible for it. At the same time, maintaining a distinction between intentional actions that are intended and intentional actions that are not may also make an important difference for moral or legal purposes.

Goals, Plans, and Skills

We have considered cases where one might be said 0144.020
to act intentionally in *A*-ing despite not having *A*-ing as a goal, and thus not having the intention to *A*. Conversely, there are cases where one does *A*, one has *A*-ing as a goal and in this sense an intention to *A*, and yet it is unclear whether the *A*-ing qualifies as an intentional action.

Take the following example adapted from Mele. 0144.021
Tom, who has never handled a gun before but is offered a large prize for hitting the bull's-eye on a distant target, aims carefully, fires and hits the bull's-eye. Does he do it intentionally? Here, intuitions diverge. Some philosophers consider that Tom intentionally hits the target, others deny it. Those (Velleman, 1989; Mele and Moser, 1994) who deny that the action is intentional typically insist that in this and similar cases, successfully *A*-ing is not sufficiently within the control of the agent because he lacks a reliable plan for or sufficient skill at *A*-ing and because there is therefore an important element of luck involved in the action resulting in success of failure. Yet, in Tom's case,

even philosophers who would deny that the action was intentional accept that the action was done for a reason – Tom wanted the prize money and thought that to win it he had to hit the bull's-eye – and that in this sense he acted with the intention to hit the bull's-eye. Velleman sees such cases as proof of a fundamental ambiguity of the term 'intention', which is used to denote both plan states and goal states. It would seem that those who think that Tom's and similar actions qualify as intentional actions have intentions as goal states in mind, and those who deny this think of intentions as plans or representations that guide and control the action.

0144.022 Yet Velleman's distinction may still not fully capture the complexity of our intuitions. Take the following more extreme example, borrowed from Mele and Moser. Lisa selects a sequence of six numbers to win a fair Florida instant lottery, and, as it happens, she wins the lottery. Here, it seems that we would not want to say that Lisa wins intentionally, and that we would also be reluctant to say that she acted with an intention to win, despite the fact that she chose the numbers out of a desire to win the lottery. Our unwillingness to consider the action as intentional is not motivated by Lisa's lack of skill or the frailty of her plans. Assuming that the lottery is fair and that there is no way to win by cheating, Lisa has done all that was in her power to win the lottery. The difference between the Lisa and Tom cases lies in how much the successes of their actions, considered as types, can in principle depend on the agents' skills or planning abilities.

0144.023 It seems, therefore, that to account for our intuitions a threefold distinction is needed. At one extreme are actions over which the agent has a sufficient degree of control, and which are therefore uncontroversially intentional. In the middle are actions over which the agent could in principle have a sufficient degree of control (as witnessed by the fact that some other agents do exhibit such control – for instance expert marksmen), and which can be deemed intentional in a weaker sense (Tom's case). At the other extreme are actions over which we could not in principle have sufficient control, because there is some important element of luck on which their outcome depends essentially (Lisa's case).

Sudden, Impulsive, and Subsidiary Actions

0144.024 Some philosophers (Searle, 1983; Wilson, 1989) argue that sudden and impulsive actions can be intentional despite being performed without the agent's forming, either consciously or unconsciously, a prior intention to do them. Thus, if a pile of books on my desk starts to topple, I might suddenly reach for the pile to keep it from collapsing; or, frightened by an earth tremor while in California I might impulsively throw myself under a table; acting in both cases intentionally. Similarly, Searle claims that even in cases where we have a prior intention to do something, there are often many subsidiary actions that are not represented in the prior intention and yet are intentionally performed. I might have a prior intention to type a certain sentence, without the intention representing the key pressings that must be performed for the sentence to be typed. It may be thought that sudden and impulsive actions are too fast for there to be time to form an intention. It may also be pointed out that such actions do not meet the strong consistency requirements that are distinctive of actions done in conformity with a prior intention. Typically, they are performed with no regard for their potential side effects or for their general coherence with the agent's other beliefs and goals.

On Searle's account, these actions are intentional 0144.025
despite the absence of a prior intention because they involve an 'intention-in-action'. In his terminology, a prior intention is an intention that is formed prior to the action and that represents and causes it as a whole. Thus, both future-directed and immediate intentions are prior intentions. An intention-in-action is the mental component of the action itself; it presents, causes, and is contemporaneous with, bodily movements. The intention-in-action together with the bodily movements it causes constitute the action. According to Searle, all intentional actions have intentions-in-action but not all have prior intentions. And even when both a prior intention and an intention-in-action are present, the intention-in-action will typically be much more determinate than the prior intention.

The claim that some intentional actions need not 0144.026
involve prior intentions can be challenged in two different ways. Some philosophers agree that in the case of sudden or impulsive actions, intentions are absent, but, because of this absence, deny that such actions really qualify as intentional. Bratman (1987) suggests that they are purposeful and voluntary but still not intentional (nor unintentional). Conversely, one might agree that such actions are intentional, but claim that they involve an intention. Thus, Mele (1992) distinguishes between forming an intention and (passively) acquiring one, and suggests that although in the case of sudden and impulsive actions an agent may lack sufficient time

to form an intention, he or she may still have time to acquire one. Similarly, in the case of actions involving many subsidiary actions, one may contend that a prior intention normally includes a detailed plan for acting and thus that intentions-in-action are superfluous.

0144.027 As Mele and Moser (1994) suggest, highly relevant to this debate is the difficult question of how much of what is going on representationally in the preparation and execution of an action should be included in the content of intentions. Work in the neuroscience of action suggests that there are multiple levels of representation and control of action, from high-level specifications of goals down to low-level specifications of neuromuscular activity. It remains an open question, partly empirical and partly conceptual, which of these representations should be considered part of the content of intentions or, if one countenances them, intentions-in-action, and what role criteria such as accessibility to consciousness or representational format (conceptual or not) should play.

CONCLUSION

0144.028 Philosophical analyses of the notion of intention reveal a number of aspects – practical reasoning, planning, goal-directedness, control of action – which it may either essentially include or be closely associated with. Presumably, what we take as prototypical cases of intentions and intentional actions are cases where all dimensions are present. In less clear cases, philosophical intuitions differ. The debates over the nature of the connection between intentions and intentional actions bespeak the lack of general agreement on the respective roles and relative importance of these aspects. Different approaches privilege different aspects or weigh them differently when characterizing intentions as states of mind and when characterizing intentional actions. These disagreements may in turn be explained in part by interest in the different issues – rationality, ethics, freedom, mental kinds, motor organization – with which the notion of intention is connected. A number of recent studies explore intentions from an intersubjective point of view. Some theorists are interested in collective actions and the types of shared or cooperative intentions that make them possible. Others try to characterize the abilities that underlie our capacity to interpret observed actions as intentional and to attribute intentions to the agents performing them, thus connecting work on intention with the topic of mindreading. These new perspectives may well pay special attention to features of intentions less salient in other approaches.

References

Anscombe GEM (1963) *Intention*, 2nd edn. Ithaca, NY: Cornell University Press.

Brand M (1984) *Intending and Acting*. Cambridge, MA: MIT Press.

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

Davidson D (1980) *Essays on Actions and Events*. Oxford: Clarendon Press.

Davis WA (1984) A Causal Theory of Intending. *American Quarterly* **21**: 43–54.

Harman G (1986) *Change in View*. Cambridge, MA: MIT Press.

Mele AR (1992) *Springs of Action*. Oxford: Oxford University Press.

Mele AR and Moser PK (1994) Intentional action. *Noûs* **28**: 39–68.

Pacherie E (2000) The content of intentions. *Mind and Language* **15**: 400–432.

Roth AS (2000) The self-referentiality of intentions. *Philosophical Studies* **97**: 11–52.

Searle JR (1983) *Intentionality*. Cambridge, UK: Cambridge University Press.

Velleman JD (1989) *Practical Reflection*. Princeton, NJ: Princeton University Press.

Wilson G (1989) *The Intentionality of Human Action*. New York, NY: Peter Lang.

Further Reading

Audi R (1993) *Action, Intention, and Reason*. Ithaca, NY: Cornell University Press.

Castañeda H-N (1975) *Thinking and Doing*. Dordrecht: Reidel.

Davis L (1979) *Theory of Action*. Englewoods Cliffs, NJ: Prentice-Hall.

Donagan A (1987) *Choice: The Essential Element in Human Action*. London: Routledge.

Gallese V and Goldman A (1998) Mirror neurons and the simulation theory of mind reading. *Trends in Cognitive Science* **2**: 493–501.

Ginet C (1990) *On Action*. Cambridge, UK: Cambridge University Press.

Grice H (1971) Intention and uncertainty. *Proceedings of the British Academy* **57**: 263–279.

Jeannerod M (1997) *The Cognitive Neuroscience of Action*. Oxford: Blackwell.

O'Shaughnessy B (1980) *The Will*. Cambridge, UK: Cambridge University Press.

Taylor C (1964) *The Explanation of Behaviour*. London: Routledge and Kegan Paul.

Keywords: (Check)

action; belief; desire; practical reasoning; self-referentiality

ECS Article 144 Author queries

`Article Intentions Self-Referential?', 2nd paragraph; It seems that part of the point of this story is that Mabel is backing out of her drive *because* she intends to run Ted over (i.e., that's why she's setting off in her car). If so, could this be made clearer in the story? (If not, then the act of *A*-ing may not have been caused by an intention to A, which makes the last sentence puzzling.)

`Are Intentions Self-Referential?', 4th paragraph, 2nd sentence: Please supply reference for (Searle, 1992).

Can you supply a small glossary defining complex terms for the layperson?