Rational and moral action: a critical survey of rational choice theory

de Jonge, J.P.R.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
CHAPTER XVII

ARE JUSTIFYING REASONS MOTIVATING?

1. Introduction

The basic claim of rational choice theory is that intentional action is to be explained by reference to the beliefs and desires of the agent. Hume, as many philosophers today, argued that the role of knowledge (beliefs) is limited to the discernment of means to ends. 271 To be motivated to act requires a psychological state (a pro-attitude), which only desires can provide. Whereas beliefs are motivationally inert, desires provide the passion that activates the agent. 272 A non-

271 Sugden has recently published an article in which he makes two statements: first, Hume was not endorsing an instrumental form of practical reasoning; second, Hume opposes a propositional decision theory. (Sugden, 2006) I refer to a theory of motivation that is said to originate in Hume's view that only desires can motivate an agent to act.

272 In order to provide an alternative, I shall define belief not just as a passive domain of (scientific) knowledge, but as an active domain of
Humean alternative would be one in which a genuinely cognitive state is doing the motivation. Pure cognitivism supposes that a complete motivating state can consist of nothing but (objective) reasons. It assumes that, where there is motivation, there will be desire. But it understands the desire as the state of being motivated rather than as some part of what motivates. What motivate an agent are his or her beliefs about (desirable) states of affairs.

What about the justification of actions? From the Humean point of view reason can tell us how to satisfy our ends, but it cannot tell us whether these ends are themselves 'rational'. The opposing view holds that one can, to some extent, objectively establish the value of the end that is to be obtained or realized through the action. It is not the desire for an end, but its desirability that justifies an action. The crucial question to be answered in this chapter is: will a justifying reason also be a motivating one?

In short: there is disagreement about the motivational potential of reasons and the justificatory potential of desires. (Dancy, 1995) I shall propose to combine a desire-based and a reason-based approach, as part of a general approach, in which it is accepted that beliefs and desires can independently give rise to action. Within this general approach, it will appear that desires cannot serve as a justification of an action. Desires usually lack the evaluative force that is thought necessary to make them justifying reasons, namely some kind of evaluation of the desired object as good -e.g., pleasant, interesting, decent etc. There are, of course, circumstances in which the desirability of the action is a clear-cut case even when the agent just did what he desired

common knowledge. Beliefs about the world, which include beliefs such as that it is right to follow conventions, or to be honest etc. can provide agents with reasons for actions. The assumption is that knowledge provides good reasons.
to do. But when I refer to desires, I have in mind motives, based upon purely emotional states, that are only slightly amenable to critical reflection (see also the distinction between higher-order and lower-order desires in chapter XVIII).

2. Internal and external reasons

It is convenient to start with the question: ‘what is the basis for our ability to choose?’ If we are autonomous and choose our own purposes, then the resources that make our choosing possible have to come from within. Only two resources suggest themselves: reason and want. Here authors part company.

Humeans, believe that reason is imposed, and is something external. 273 Therefore, the identity, the being of human beings, consists in their wanting things. For, how is autonomy secured? How can we be sure that anyone’s interests, ends and purposes are really that person’s own? The obvious way to make this clear is to show that they all stem from one’s desires, since one’s desires are one’s own, whereas reason is external. Wants precede reason, for if thought shaped experience, then individuals are dependent on the sources of these thoughts, that is to say, dependent also on the norms and presuppositions of society they inhabit. This conflicts with this variant's conception of individual autonomy. There is no denial that reason and desire are inseparable aspects of choice. This is another question. There can be no choice without both a desire for an

273 As Hume formulated it: it is evident that all passions, volitions and actions, are not susceptible of any agreement or disagreement; and therefore it can never be an object of reason. Bentham stated that: nature has placed mankind under the governance of two sovereign masters: pain and pleasure. Since these masters exercise their power through the medium of desire, reason cannot motivate us. And also Mill argued that, because autonomy is fundamental, it cannot consist in reason, to which it would then be subservient, but must have its roots in what we want.
end and reasoning about how to achieve it. The role of reason is confined to the means to achieve ends. This is still the dominant point of view in rational choice theory.

The discussion about the capacity of rational arguments, principles or norms to contribute to the explanation and justification of intentional action can be clarified by applying the distinction between internal and external reasons. An authoritative article about this distinction is Williams (1981). I will take this article as my point of departure.

An internal reason has the form ‘A has a reason to Ø’. Williams insists that the claim that a person has a reason must always be based on that person’s perspective of the situation. From a motivational point of view it is a matter-of-course that a reason is only a reason when the person in question is motivated to act on that reason.

An external reason has the form ‘there is a reason for A to Ø’. This point of view does not take the perspective of the agent as its point of departure, but the perspective of the person making the judgment. If I believe that I would have reason to Ø in circumstances C, and that the agent’s situation is not different from mine, in relevant respects, then I will be inclined to think that this reason counts in favor of Ø-ing in his case as well.

According to Williams internal reason statements can be discovered in deliberative reasoning. An agent can come to see that he has a reason to do something that he did not see he had reason to do at all. In this way the deliberative process can add new options for which there are internal reasons, just as it can also add new internal reasons for actions already undertaken. Williams thus acknowledges that a person’s subjective motivational set can undergo various changes. "In his unaided deliberative reasoning, or encouraged by the persuasions of others, he may come to have some more concrete sense of what would be involved [were he to pursue a specific end], and lose his desire for it, just as, positively, the imagination can create new possibilities and new desires". (Williams, 1981, 105) Not only can desires create beliefs, beliefs can also excite (new) desires. But, according to
Williams, deliberation is always from existing motivations. "The new motivation could be in some way rationally arrived at, granted the earlier motivations". (Williams, ib., 109) Beliefs can take the agent from one desire to another (or from one connative state to another), but it cannot give rise to a desire.

The claim of an external reason statement can be formulated in the following way. "When A comes to believe that there is a reason for him to Ø, he will be motivated to Ø, even though, before, he neither had a motive to Ø, nor - in the extreme case - any motive related to Ø-ing in one of the ways considered in the account of deliberation. (...) What is it that one comes to believe there is a reason to Ø, if it is not the proposition that if he deliberated rationally, and knew all the relevant facts, he would be motivated to act appropriately". (Williams, ib., 109) However, when one says that A has a reason to Ø, isn't one committed to say that A acts irrational when A does not Ø? The answer is negative; though there is a (objective) reason to Ø, an individual may nevertheless not be sensible to this reason. 274 One could argue that the whole point of the external reason claim is that a reason can be a valid motive to act, independently of the fact that it really motivates the agent. The fact that external reasons can be true

274 The argument is that since (external) reasons are rational considerations, they cannot be rejected without raising the suspicion of irrational behavior. When it comes to moral reasons, the claim that they are rational considerations would imply that acting immorally is a matter of being irrational. This seems like a knockdown argument against immoral behavior. (see Doyle, 2000, about egoism-conversion) But it will be difficult to defend such a claim. A reasonable alternative explanation would be that some agents don't have a disposition to act according to reason because their subjective motivational set lacks certain sensibilities. This argument may be used to point at the difference between the moral and the amoral agent. An amoral agent having the same belief as the moral agent will not be motivated to act in the same way. An important consideration in this context is that reasons that support the adoption of a moral commitment may not be able to bring about "a change in the fundamental orientation in an agent's life that moral commitment requires". (Doyle, ib., 22) Moral reasons can motivate an agent, but the motivation is contingent.
independently of the motivations of an agent, would, moreover, just be fine for actions that qualify as moral actions. For the reasons that would justify an action, would be valid for every agent and this is precisely what is assumed in the case of moral reasons. 275

The internalism-externalism distinction reflects the disagreement between Humean’s and non-Humean’s. From a Humean point of view there is no sense in which our desires or passions are rational or irrational. "All reasoning is concerned either with abstract relations of ideas or with relations of objects, especially causal relations. (..) Abstract relations of ideas are the subject of logic and mathematics; (..) they yield no conclusions about action". (Korsgaard, 1996, 331) Hume also argued "reason is the faculty that judges of truth and falsehood, and it can judge our ideas to be true or false because they represent other things. But a passion is an original existence or modification of existence, not a copy of anything. It cannot be true or false, and therefore it cannot in itself be reasonable or unreasonable". (Korsgaard, ib., 218; see also Platts, 1988)

Normative requirements only provide agents with reasons to act if they are appropriately related to the agent ‘s antecedent desires. And since desires are non-rational states, there is no distinctive rational principle to enter into the explanation of motivation. Then, morality, for instance, cannot be a matter of knowledge, but is rather a matter of sentiment. The reason that moral considerations must be rooted in people 's sentiments is precisely that nothing else can motivate them, leaving no room for justifying reasons as distinct reasons. Therefore, actions are justified by the reasons that motivate them. When facing conflicting desires, rational agents will do what they most want to do.

275 A reason to act in a particular way can be a valid reason even if it does not motivate the egoist because there is nothing in it for him/her. Helping an old lady crossing a road is a matter of ordinary decency, and so we should do it even when we would not benefit from it. Sometimes there is a reason to help whether or not someone recognizes the fact. (Grice, 1991, 65)
The idea that men are impelled by ‘desire’ which can be assisted or impeded by reason but which is not itself rationally produced, is unacceptable to quite a number of moral philosophers. The problem they have with the Humean theory of motivation is either that it does not “(..) allow reason to have any critical impact on human behaviour at all.” (Hampton, 1998, 148) Or, that “(..) it is irrational to have certain sorts of beliefs, but lacking the corresponding desires and motivations.” (Wallace, 1990, 382) They argue instead: “We have reason to try to achieve some aim, when and because, it is relevantly worth achieving. Since these are reasons for being motivated, we would have these reasons even if, when we were aware of them, that awareness did not motivate us. But if we are rational, it will.” (Parfit, 1997, 130)

The non-Humean approach can be characterized as holding that people act for a reason, and that reasons are considerations in virtue of which actions are good in some respect and to some degree. One has a reason for action only if its performance is likely to produce, or to contribute to producing, something good. In this conception of action the beliefs of an agent connect agency and value. This approach allows that a gap may exist between what is good and what attracts the agent. And therefore the reason that justifies the action may not always motivate the agent. But when we deliberate about what we desire most, we deliberate about what it would be best for us to want. "This feature is essential. It explains the sense in which desires are under control, rather than being states of minds visited upon us. A desire is reasonable when we have it because of a belief in the value of its object, and it would disappear were we to abandon that belief. (..) Our desires are, in this regard, like beliefs. We cannot want what we see no reason to want (..)” (Raz, 1999, 53)

From this point of view there are no desires -besides urges and inclinations- that are not reason-based. The crucial distinction is between thinking of desires as being an essential part of a complete motivating state, or of thinking of them as identical with the state of being motivated. The state of being motivated needs an explanation and this must be
given in terms of the nature of what is desired. A desire is an endorsement of a reason that is independent of it. If a desire is reason-based then the desire does not itself add to those reasons. "Desires are held for reasons which they can transmit but to which they cannot add". (Dancy, 2000, 39)

3. The subjective motivational set

According to the Humean point of view a reason must be connected to the agent’s subjective motivational set to be motivating. We can think of this motivational set as a set of primary reasons, combinations of pro-attitudes and beliefs.276 The question that is raised by the non-Humean view is whether beliefs (or, more generally, cognitive attitudes) can give rise to a pro-attitude. A denial is often based on arguments such as: ‘since we experience that people do not always desire what we believe is right, this proves that reasons do not motivate’.

Korsgaard thinks that much discussion is based on a false impression of what the internalism requirement means. It does not require that rational considerations always succeed in motivating us. There are plenty of things that interfere with the motivational influence of a given rational consideration, such as emotional distress, weakness of will or irrationality. The scope of the subjective motivational set is, moreover, not the same for all people. Its content may vary between different agents, the impetus of elements may be latent and sometimes elements may simply be ignored. All these things may explain the differential reaction to reasons across agents and by the same agent at different moments in time. (Korsgaard, 1996, 326ff)

Formally, the items that can be found in an agent’s motivational set depend solely on the kinds of reasoning that are possible. The question is how broad or how narrow the

276 A pro-attitude combines with a belief state to produce a primary reason. The primary reason for an action is its cause. Pro-attitudes can be of very different kind. (Davidson, 1980; see chapter II)
subjective motivational set is defined. The scope of the subjective motivational set is determined by the range of pro-attitudes that are accepted as principles of action, and by the admittance of cognitive propositional attitudes next to connative propositional attitudes. Therefore, the answer depends on, firstly, whether only means-end reasoning is allowed, or also reasoning about ends and consequently whether it only contains plain desires or also commitments, principles, conventions etc. (as in Davidson’s description, see Davidson, 1980). Secondly, and, more importantly: do normative principles only exert influence on actions for which the agent already has a desire to act (as in the strict Humean version), or can these reasons motivate us, directly or indirectly. When this would be allowed the agent’s perspective and the perspective of an impartial observer correspond. This is the question that ultimately divides ‘internalists’ and ‘externalists’. In a desire-based approach a reason could never oppose a passion if it did not originate from a desire that is already part of the subjective motivational set. But in the externalists’ approach it is assumed without reservation that reasons play an important role in the emergence of new motives.

Williams takes the range of possible elements in a person’s subjective motivational set to be broader than the class of plain desires. It includes such things as ‘dispositions of evaluations, patterns of emotional reaction, personal loyalties and commitments’. (Williams, 1981, 105) Williams acknowledges that through a process of sound deliberation new elements to one’s subjective motivational set can be added and old ones can be eliminated. Even reasons can be adopted, but only on the basis of the agent’s existing subjective motivational set.

Scanlon thinks that the disagreement between internalists and externalists is limited in the end. On both sides it is conceded that: a) reasons very often have subjective conditions; b) failing to see the force of a reason need not involve irrationality; although c) it may, as in the case of cruelty and insensitivity, involve some other failing or
deficiency. (Scanlon, 1998, 372) We could resolve the remaining disagreement in either of two ways, he suggests. We could limit the universality of reasons by tying reasons to motives. An agent only has a reason when he is being moved by it. Or we could relax the requirement that in order for an agent to have a reason the argument for it must be linked to his subjective motivational set. Scanlon opts for the second solution. A person can have a reason to act although he is not motivated by it. And when a person does not see he has a reason to act, it is because of deficiencies on his side. But Scanlon adds that there is no decisive argument in favor of one of these two alternatives, nor does he think it would make a great deal of difference. As long as a), b) and c) are accepted, everything really important is in place.

I think Scanlon overlooks the really important difference between internalists and externalists, viz. the difference between two conceptions of desire. The crucial distinction is between thinking of desires as being themselves part of a motivating state, and thinking of them as identical with the state of being motivated. Nagel was the first to espouse the idea that ‘desire’ is a state of being motivated, as "motivatedness", rather than as what motivates. I discuss Nagel’s approach in the next section.

4. Motivated desires and motivating reasons

In the opinion of Nagel, "the assumption that a desire underlies every intentional act depends on confusion between two kinds of desires, motivated and unmotivated ones. Many desires are arrived at by decision and after deliberation. (..) Hunger is brought about by lack of food, not motivated by it. The desire to buy food is motivated by hunger. Motivational explanation is as much in order for that

[277] There must be no misunderstanding; in a formal way each action model needs an animating principle. The question whether reasons can be motivating depends on their relation with such an animating principle (desires).
desire as for the action itself". (Nagel, 1978, 155) When one examines the logical reason why desires must always be among the necessary conditions of reasons for action, one sees that it may often be motivated by precisely what motivates the action. "And if it is motivated by that reason, it cannot be among the conditions for the presence of those reasons ". (Nagel, ib., 156)

Nagel’s notion of a motivated desire then seems to meet the objection against the Humean theory that it gives no instrument for agents to critically reflect on their own behavior and at the same time to give in to the demand of Humean theory that agents ultimately have to have a desire to further an end. Be it that we have here a use of 'desire' that indicates merely a motivational direction and nothing more.

Pettit and Dancy, subsequently defended the idea that a motivating state can originate from cognitive considerations. Some of the desires that figure in motivating reasons are such that their presence is entailed by the presence of certain beliefs (desiderative beliefs). In such cases it would be natural to say, Pettit argued, that the desires are constituted by the beliefs. "We must regard the desire in that case as inheriting the cognitive or discursive status of the desiderative belief". (Pettit, 1988, 531) A desiderative belief is just a motivated desire.

Dancy even argues that there are no desires - besides urges and inclinations - that are not reason-based. Dancy understands 'desire' as a state of being motivated - as motivatedness - rather than as what motivates. "Starting from a state that is purely cognitive, it is conceivable that motivation should then occur without addition to or other change in that psychological state of the agent to which the motivation (the

---

278 We could compare his distinction with the contrast between a ‘spontaneous’ or ‘felt’ desire and a ‘reason-based’ desire (Raz, 1975), or with the distinction between manifest and true desires (Harsanyi, 1982) and relate it with Parfit’s distinction between desire-based and value-based motivation. (Parfit, 1997) In Nagel’s terminology a motivated desire is a desire that follows from a reason. The desire would vanish should the reason disappear. The desire is just an animating force.
motivatedness) is a response". (Dancy, 2000, 92) 279 The explanation of intentional action can always be achieved by laying out the considerations in the light of which the agent saw the action as desirable, sensible or required. Dancy understands motivating reasons in terms of normative ones, as reasons of 'what is believed'. Normative reasons are potential justifying reasons.280

Nagels' insistence that the fact that the presence of a desire does not entail that it is a necessary condition for the presence of the reason, fostered the fear among Humean theorists that the notion of a motivated desire might lead to the idea that (prudential) reasons are counterexamples to their thesis that all reasons for action depend on the agent's desires. And, this would be a refutation of the doctrine that deliberation always departs from existing desires, as Williams contended. Smith thus responded, in defence of the Humean theory of motivation, that "the idea that there may be a state that motivates a desire, but which is not itself a desire, is simply implausible." (Smith, 1987, 59). Instead of a motivated desire he argued in favor of a motivating reason.

Smith's approach is based on a strict distinction between beliefs and desires. Beliefs aim at truth, and beliefs should be changed to fit with the world. Desires aim at realization, and the world should be changed to fit our desires. From this distinction it follows that a cognitive propositional attitude (a belief) can give at most a partial specification of a reason for action. It needs in addition something non-cognitive, a state of the will or a volitional event (desire). Essential in his version of the Humean theory is also a distinction between two concepts of desire. We may associate it with a feeling or an emotion. This is the phenomenological definition of a

279 In my opinion the definition of 'desire' as 'motivatedness' is as close to Nagel's description of a motivated desire as it can be. Nevertheless, Dancy criticizes Nagel for not being a pure cognitivist. (Dancy, 2000, 93/94). I must admit that the point escapes me.

280 The 'motivating reason' in this context differs from the motivating reason that Smith introduces as a substitute for Nagel's motivated desire.
desire. It is modelled on the idea of sensations and it refers to psychological states. But in Smith’s view desires differ from sensations in that they also have a propositional content. We should define desires as dispositions. (Ib., 47)

Now he has laid the foundations for the formulation of a motivating reason. I.e., there are cognitive states an agent can be in only if he is in some non-cognitive state. Then he states that the Humean theory of motivation is entailed in the following premises: (Ib., 55)

1. having a motivating reason is, inter alia, having a goal.
2. having a goal is being in a state with which the world must fit.
3. being in a state with which the world must fit is desiring.

A motivating reason is thus the reverse of Nagel's motivated desire.

In reply to Pettit’s suggestion, Smith responded that a desiderative belief can be construed as a belief on the agent’s part about his own desires: as the belief that he desires ‘to Ø’. (Smith, 1988) But Pettit reminds him that this does not settle the dispute between Humeans and non-Humeans. "The issue that divides Humeans and anti-Humeans is whether motivating reasons always involve the presence of non-cognitive states, states which reason on its own is incapable of producing". (Pettit, 1988, 531) He suggest that Humeans should argue "that the only sort of belief which would get close to entailing the presence of desires is a belief which bears on the existence of precisely those appetitive states". (Ib., 532/3) Smith, accepting this argument and referring to Lewis’s dispositional theory of value, argues that the Humean theory of motivation can be defended by "taking a 'favourable psychological attitude' to be a species of desiring, thus displaying an internal, indeed analytical, connection between believing that Ø-ing is right and desiring to Ø". (Smith, 1989, 91)

The dispositional theory of value says that the best way to see what a person prefers is to look at what he chooses

281 Notice that Smith suddenly introduces a concept of desire - namely a psychological concept- that he had rejected before.
when he is presented with a choice in imagination. This is the way to see whether a person values something. "The thoughtful person does not value what he desires but rather he values what he desires to desire". (Lewis, 1989, 115)

Does the dispositional theory of value appeal to a cognitive state? Reason plays a part in the process that Lewis describes: it is involved in bringing the agent into full imaginative acquaintance with the object. But once she has achieved this position, valuing is simply a sort of mental state directed towards the object that is valued. Whether or not she finds herself valuing the object has nothing to do with reason. (see Broome, 1999b; see also, Vellema, 1992)

The criticism of Nagel’s notion of a motivated desire has furnished us with an answer to the question whether the belief-desire account of action permits agents to be moved, so to speak, ‘from both sides’. That appears not to be acceptable in the Humean tradition, as expounded by Smith. From his point of view a motivating reason is intelligible a motivated desire is not. Wallace has termed this the “desire-out, desire-in” principle. "The process of thought which gives rise to a desire can always be traced back to a further desire, one that fixes the basic evaluative principle from which the rational explanation of motivation begins". (Wallace, 1990, 370) The chain of explanation must eventually terminate in an unmotivated desire. Nagel’s attempt to integrate justifying and motivating reasons is not acceptable in the eyes of those who represent the Humean tradition.

---

282 It sounds as if Lewis is talking about a second-order desire, in which case he would be defending a motivated desire instead of a motivating reason. But he doesn't. Lewis represents the Humean point of view.

283 In subsequent articles Smith seems to have made a surprising U-turn by defending cognitivism and declaring that the claim that a decision is justified is the expression of a belief rather than a desire. Beliefs can cause and rationalize desires without the aid of a further desire. (Smith, 2001, 100) Nevertheless, I still consider his account of the Humean point of view as accurate.
5. Counter-preferential choice

Theorists who defend the idea of a counter-preferential choice also argue that reasons can underlie decisions. A natural way to discuss counter-preferential choice is by focusing on dynamic or sequential choice. People rarely make decisions on the spot, or decisions that are unrelated to former and future decisions. Many goals can only be realized when people undertake a sequence of activities and actions. An agent who makes plans tries to settle in advance what to do later; there is a time interval between formulating an intention and executing it. But an agent also wants, whatever his prior plans, to retain rational control over what he does when the time comes. He could suffer when bound to carry out the plan that he had adopted on the basis of past concerns. Shifts and changes in preferences are a normal and necessary part of human life. It only makes sense to deliberate about the future if one is open to the conditions at the relevant time in the future and their effects. At the other hand, sometimes it pays to stick to a prior formed preference. I will give two examples of potential conflicts between prior and later formed preferences: one relates to an individual agent, the other pictures a case of cooperation. In both cases it is argued that people have reason to stick to a prior formed intention even though they don’t desire it.

The first example is a thought experiment designed by Kavka, called the Toxin Puzzle. (Kavka, 1983) The story is that you are approached by an eccentric billionaire who offers you the following deal: he gives you a vial of toxin that, if you drink it, will make you painfully ill for one day, but will not threaten your life or have any lasting effects. He will pay you one million dollars tomorrow morning, if at midnight tonight you agree to drink the toxin tomorrow afternoon. He emphasizes that you need not drink the toxin to receive the money; and in fact the money will already be in your bank account hours before the time of drinking. You thus think you can sign the contract, believing
that you could avoid drinking the toxin and just pocket the million. But then you realize that if you think in those terms when midnight rolls around, you cannot be intending to drink the toxin tomorrow. So you decide to intend today to drink the toxin and then change your mind after midnight. But if that is your plan, then it is obvious that you do not intend to drink the toxin (at most you intend to intend to drink it). 284

Two points underlie this puzzle. The first is that while you have no reason to drink the toxin, you have every reason to intend to drink it. Now when reasons for intending and reasons for acting diverge, as they do here, confusion often reigns. For we are inclined to evaluate the rationality of an intention both in terms of its consequences and in terms of the rationality of the intended action. As a result, when we have good reason to intend but not to act, we have an inner conflict. This example also reveals that one cannot intend whatever one wants to intend: “our intentions are constrained by our reasons for action”. (Kavka, 1983, 36) The second point is that one foresees that, after tomorrow morning, but before the times comes to drink, one will recognize that one has no reason to drink and, indeed, compelling reasons not to drink for the money is already collected. One also foresees that one will be sufficiently rational tomorrow to revise one’s intentions accordingly.

On one view, what is up to an agent is what to do from now on. What are now under control are his alternatives from now on. On Monday morning the agent cannot rationally form the intention to drink. Circumstances have changed and there is no reason to execute an intention that was formed on the basis of alternatives available at a prior time. The instrumental rationality of an action depends on the agent’s evaluative ranking at the time of action of options then available to him. But a sophisticated agent will look ahead to what will in the future seem best to her. We could conceive of

284 You cannot pretend to intend to drink the toxin, for it is assumed that the billionaire has the rare capacity that makes you transparent to him.
such a sophisticated agent as someone who acts on plans. A plan specifies how an agent is to act at different choice nodes. Such an agent tries to solve sequential choice problems by the use of backward induction. In order to collect the money she has to form a resolute intention to drink the toxin. But this sophisticated agent violates one of the conditions of sequential choice, namely the separability condition, which says that at every decision node only those plans are acceptable that would be acceptable were it the initial choice point of a new dynamic decision problem.

The sophisticated agent is therefore criticized for (a) avoiding new information (at the choice nodes) and (b) for depriving herself of future freedom of choice. This may be a reason to consider an alternative approach: the so-called resolute choice. The characteristic feature of this approach is the assumption that the preferences of the dynamically rational agent are expressed by his plan of action. After having embarked upon the plan, she adjusts her preferences steadily so that at each stage of the implementation process she prefers to follow the plan rather than deviate from it. The resolute agent has committed herself to follow the plan. Due to this self-imposed obligation she transforms her preferences at subsequent choice nodes and adjusts them to the chosen plan of action. Therefore, the criticism (a) and (b) do no longer apply to her. (see also Gauthier, 1997)

Both Bratman and Gauthier defend the view of the resolute chooser whose intentions and plans at t(2) fit into a larger course of action that begins at t(1) and for whom the assessment of that larger course of action is crucial. Bratman acknowledges that this violates the assumption that agents are temporally located (the separability condition) but going through with one’s prior intentions may satisfy a non-regret condition. When the agent does not make a resolute choice to drink the toxin tomorrow, he will not be able to intend to drink the toxin tonight and, thus, forego the million. This non-regret condition can make follow-through rational even in the face of a present ranking to the contrary. Why should anticipated satisfaction of the no-regret condition matter to an
instrumentally rational planning agent? The idea is that anticipation of future regret or non-regret can be relevant to the stability of a prior intention of a planning agent; it can be relevant to the question of when it is reasonable to reconsider and abandon a prior intention, and when not. The force of the non-regret condition is grounded in one’s actual engagement of how one will see matters at a plan’s end. It is this concern that supports a distinctive kind of intention stability in certain no-anticipated-information cases, by formulating a resolute choice to execute the intention. (Bratman, 1998, 75)

In his reaction to the Toxin puzzle Gauthier argues that a person should expect to have an adequate reason to execute an intention if he should expect that, were he to execute it, he should do better than had he not formed it. Only an unreflective person who considers her reason for and against the drinking of the toxin only when the time comes will have good reason not to drink it. Surely, one’s reasons for and against drinking the toxin do not change. What does change is the context in which these reasons are weighed. Rationally forming an intention requires looking ahead to its execution. The reflective person in forming her intention considers her reasons for and against drinking the toxin and assesses her course of action -intention and execution- as a whole. "Once she undertakes a course of action as a whole she must rationally continue to assess her particular actions as part of that whole, unless she comes to have reason to abandon her course of action". (Gauthier, 1998, 55)

The second example I want to consider is a case of cooperation or reciprocal behavior. (I borrow the example from Mintoff (1997))

---

285 Previously he had written that a commitment to a plan might rationally override preference-based reasons that in themselves would support an action incompatible with the plan. In other words, rational commitment to a plan may require a counter-preferential choice. "Resolute planning requires resistance to the direct appeal of preferences -resistance, we might say to temptation (...)." (Gauthier, 1996, 242/3)
Two farmers, Jones and Smith, have adjacent farms. Next week Jones’s crops will be ready for harvesting, the week after Smith’s will be ready. There is the chance of mutual benefit, if Smith would help Jones and next week Jones helps Smith. But when Smith has helped Jones, how does he know Jones will reciprocate? After the harvest Jones is retiring, is selling his farm and moving to another place. Therefore, Smith would like to know whether Jones really now intends to reciprocate and whether Jones might reconsider his intention after Smith has helped him.

Mintoff considers the argument of Gauthier that "it is rational to act on a previously formed intention if and only if the expected outcome of acting on that intention is at least as good as the expected outcome of not having formed any intention in the first place". (Mintoff, 1997, 624) Now suppose that after Smith has helped Jones, bad weather sets in and the costs of Jones of reciprocating escalate. Jones would need to reconsider whether to carry out his intention, for the expected outcome of non-cooperation could exceed that of cooperation. In that case it would not be rational for Jones to form an unrestricted intention to help Smith after Smith has helped him. But Mintoff believes that in the case of Jones and Smith it is best for Jones to be disposed not to reconsider it. (Ib., 630) Thus it would be rational for Jones to help Smith if she helps first, for the expected outcome of reciprocating –namely mutual cooperation– is at least as good as the alternative, namely mutual non-cooperation.

Is it reasonable to execute an intention when the reason for that intention does not exist anymore? This is dilemma that Kavka raised in the Toxin Puzzle. The primary reaction is: no, unless. When the reasons for the emergence of the intention cease to exist, then the execution of the intention does not serve any rational purpose. An exception is when the intention has created an obligation/expectation. One has, for instance, accepted help and created the expectation to reciprocate. Though it seems that it is not rational to return a favor after one has received it oneself, it would neither be rational to prevent the emergence of spontaneous forms of
mutual provision of services by exhibiting ‘free-rider behavior’. The secondary reaction is: yes, if. Bratman refers to a commitment to one’s prior intentions and Gauthier thinks that an agent has reason to execute a prior formed intention if he should expect that, were he to execute it, he should do better than had he not done it. For instance, because when he did not execute the former formulated intention, he would not have been able to formulate a sincere intention and thereby he would forgo the million dollars.

In all those cases in which it was thought rational to execute the intention, it appeared that some resolute choice or some balanced reason was called for. A resolute choice is not preference-based but plan-based. The execution of the intention is not desired but argued for. Reasons not only motivate, they can even suppress preferential desires.

6. Conclusion

In this chapter I have discussed the question whether, if one has a reason to Ø, one ought (rationally) to have a motive to Ø. The rationale for this question is whether moral reasons have (contingent) motivational power.

Two issues appeared to be crucial. The first was the distinction between internal and external reasons; the second concerned the difference of opinion between Humeans and non-Humeans about the relation between reasons and desires.

The internal-external division turned around the problem that internal reasons are by definition motivating but not a priori justifying, while external reason are justifying, but not motivating. The standard view is that reasons only motivate when they are connected to the agent’s motivational set. Crucial in this context is the scope of the subjective motivational set. The essential difference between theories of practical reasoning is whether cognitive states are allowed next to connative states within the motivational set. When the scope of the motivational set is narrowly defined, and contains psychological states only, as in the Humean
tradition, the motivational force of reasons is nil indeed. When the scope of the motivational set admits moral principles, conventions, collective goals and so on (see Williams and Davidson), then we can conceive of the subjective motivational set as containing primary reasons in any combination of connative and cognitive attitudes, even allowing cognitive propositional attitudes. This would also give us the opportunity to introduce concepts like ‘coercive rationality’ or ‘resolute choice’ to solve the problem of the counter-preferential choice.

What divides the Humean and the non-Humean approaches is, whether, given a wide scope of the subjective motivational set, the emergence of a reason for action that would justify the action must always be supported by a prior desire according to the 'desire out-desire in' principle. The non-cognitivism of the Humean approach has occasionally aroused the criticism that it leaves moral life too easy. That it distorts the sense of authority that (normative) reasons have over us with regard to the desirability of the action. The reply is that, when we deliberate, we make judgments about the prima facie desirability of our options and, on this basis, reach a conclusion whether a particular option is desirable, all things considered. Someone who has a prudent (or moral) belief like, "it is right that I Ø" expresses a belief, a belief about the rightness of his Ø-ing. From the point of view of the dispositional theory of value to be a value means that, which we are disposed, under ideal circumstances, to desire to desire (Smith, 1989). But non-Humeans think that it is just impossible to have a belief about what would be good and lack the corresponding desire. The belief and desire are conjoined. Beliefs can cause and rationalize desires without the aid of a further desire.

I suggest we adopt a more general approach of practical reason in which it is accepted that agents are motivated in different sorts of ways by different considerations in different circumstances. We should accept that desires sometimes are essential parts of the motivating state (unmotivated desires) and that desires sometimes only denote a state of being
motivated (motivated desires). We should perceive the agent as a multi-faceted being, with conflicting inclinations and dispositions. This would add some psychology to moral theories and would create some common ground between cognitive and connative attitudes. We should admit that neither are desires blind nor are reasons inert. 286 We usually cannot have a desire except for a reason, neither have a reason except for some desire. (Raz, 1986, 29/30; see also, Griffin, 1986; 1996) 287 This would solve some paradoxes and puzzles. The paradox that we should have to say: “It is true that I believe in p; but that is just a psychological fact about me; about the truth of p itself, I remain uncommitted.” (Nagel, 1997, 32) And the ‘internalist’s puzzle’ that there are reasons which are not ‘reasons for us’ simply dissolves. All reasons motivate, but only good reasons justify.

286 Pettit and Smith have moved in this direction. In "Backgrounding Desire" they argued "there is room and need to see desire as a sort of state which closely parallels beliefs". (Pettit and Smith, 1990, 591)
We could envisage purely motivated and unmotivated desires as extremes on a scale ranging from pure cognitive to pure connative reasons.

287 As Ford expresses it: ‘motivation represents the integrated patterning of a subset of cognitive and arousal processes.’ (Ford, 1992; in which he presents results of empirical research).
An example of the infusion of some psychology in moral theory is Simpson’s suggestion to connect beliefs and motives by a psychological interpretation of moral dispositions. There are moral beliefs, he argues, that are not motivationally inert, but possess the disposition of ‘emotional concern’. It is arguable that reasons are also entailed by emotional beliefs. I can have a reason to be concerned. Whenever I pity others, the belief that something is hurtful to them gives me a reason to offer relief even if I am not always motivated to provide it. (Simpson, 1999, 204) The contingent awakening of his moral sensibility may explain why an agent is sometimes determined to act in an other-regarding way because he thinks he ought to, all things considered.