

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE FILOSOFIA E CIÊNCIA HUMANAS
PROGRAMA DE PÓS-GRADUAÇÃO EM FILOSOFIA**

RAFAEL GRAEBIN VOGELMANN

WILLING AND WANTING - A VOLITIONALIST ACCOUNT OF MOTIVATION

Porto Alegre
2021

RAFAEL GRABIN VOGELMANN

WILLING AND WANTING - A VOLITIONALIST ACCOUNT OF MOTIVATION

Tese apresentada ao Programa de Pós-Graduação em Filosofia da Universidade Federal do Rio Grande do Sul (UFRGS) como requisito parcial para a obtenção do Título de Doutor em Filosofia.

Orientador: Prof. Dr. André Nilo Klaudat

Porto Alegre
2021

CIP - Catalogação na Publicação

Vogelmann, Rafael Graebin
Willing and Wanting - A Volitionalist Account of
Motivation / Rafael Graebin Vogelmann. -- 2021.
155 f.
Orientador: André Nilo Klaudat.

Tese (Doutorado) -- Universidade Federal do Rio
Grande do Sul, Instituto de Filosofia e Ciências
Humanas, Programa de Pós-Graduação em Filosofia, Porto
Alegre, BR-RS, 2021.

1. Vontade. 2. Motivação. 3. Desejos. 4. Metas. 5.
Razões. I. Klaudat, André Nilo, orient. II. Título.

RAFAEL GRAEBIN VOGELMANN

WILLING AND WANTING - A VOLITIONALIST ACCOUNT OF MOTIVATION

Tese apresentada ao Programa de Pós-Graduação em Filosofia da Universidade Federal do Rio Grande do Sul (UFRGS) como requisito parcial para a obtenção do Título de Doutor em Filosofia.

Orientador: Prof. Dr. André Nilo Klaudat

Defendida em 24 de março de 2021

BANCA EXAMINADORA:

Prof. Dra. Barbara Herman
University of California, Los Angeles – UCLA

Prof. Dr. David Horst
Universidade Federal do Rio Grande do Sul – UFRGS

Prof. Dr. Leonardo de Mello Ribeiro
Universidade Federal de Minas Gerais – UFMG

Profª. Dra. Sílvia Altman
Universidade Federal do Rio Grande do Sul – UFRGS

AGRADECIMENTOS

A Laura, pelo amor, pela companhia e pelo carinho.

Aos meus pais, Isabel e Ervino, pelo contínuo suporte e incentivo.

A meu orientador, André Nilo Klaudat, pelo conhecimento compartilhado e por me desafiar constantemente.

APOIO DE FINANCIAMENTO CNPQ

O presente trabalho foi realizado com apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

This study was financed in part by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

RESUMO

O objetivo desse trabalho é motivar e defender a tese de que a vontade é a fonte de nossas ações. De acordo com o modelo volicionalista que defenderei, nós somos dotados de vontade, uma capacidade de tomar decisões. Quando agimos intencionalmente e por uma razão, a atividade da vontade é parte da etiologia da ação. Isto é, parte do que explica a ação é o fato de que o agente exercitou sua vontade de maneira a decidir agir à luz de certa consideração. Ademais, de acordo com esse modelo, a atividade da vontade não pode ser reduzida às operações de desejos ou juízos normativos. O agente, por meio de sua vontade, desempenha um papel irreduzível na produção de suas ações. Minha defesa desse modelo emerge gradualmente da crítica de modelos alternativos. Em primeiro lugar, eu rejeito a ideia de que somos movidos por desejos compreendidos como forças motivacionais. Eu argumento que essa ideia é incompatível com a existência de casos de incentivos múltiplos (isto é, casos nos quais o agente tem mais de um incentivo para agir mas nos quais seu motivo corresponde a apenas um desses incentivos). Para dar conta de tais casos, nós temos que atribuir a agentes a capacidade de determinar ativamente os objetivos visados por suas ações. Essa capacidade é a vontade do agente. Em segundo lugar, eu sustento que casos de incentivos múltiplos mostram que a vontade não pode ser compreendida como a capacidade de identificar razões para ação e pesá-las de modo a chegar a um veredicto normativo sobre o que devemos fazer. Antes, devemos conceber a vontade como razão prática, entendida como a capacidade de se engajar em episódios de raciocínio que concluem não em juízos normativos mas em intenções. Após argumentar em favor da concepção da vontade como razão prática, me volto para o modelo padrão da ação, segundo o qual nossas ações são causadas por pares desejo-crença. Sustento que uma vez que recusamos a noção de forças motivacionais, desejos (no sentido amplo que defensores do modelo padrão usam o termo) apenas podem ser compreendidos como disposições para decidir agir à luz de certas considerações e que, conseqüentemente, o modelo padrão colapsa no modelo volicionalista. Isso encerra minha defesa da tese de que nós não somos movidos por nossos desejos, mas antes determinamos nosso próprio comportamento por meio do exercício da nossa vontade. Por fim, argumento que devemos compreender a vontade não como a capacidade de decidir à luz de nossas crenças, mas antes como a capacidade de decidir à luz de fatos – uma capacidade que não é perfeitamente exercitada quando decidimos agir à luz de uma crença (mesmo que verdadeira).

Palavras-chave: Vontade; Motivação; Desejos, Metas; Razões.

ABSTRACT

The goal of this work is to motivate and defend the view that the will is the source of our actions. According to the volitionalist model I will defend, we are endowed with a will, a capacity to make decisions. When we act intentionally and for a reason, the activity of the will is part of the etiology of the action. That is, part of what explains an action is the fact that the agent has exercised her will so as to decide to act in light of a particular consideration. Furthermore, according to this model, the activity of the will cannot be reduced to the operation of desires or normative judgments. The agent, through her will, plays an irreducible role in the production of her actions. My defense of this model emerges gradually from the criticism of alternative models. First, I reject the idea that we are moved by desires conceived of as motivational forces. I argue that this idea is incompatible with the existence of multiple-incentives cases (i.e., cases in which the agent has more than one incentive to act but in which her motive corresponds to only one of these incentives). In order to account for such cases, we have to ascribe to agents the capacity to actively determine the goals at which their actions aim. This capacity is the agent's will. Second, I argue that multiple-incentives cases show that the will cannot be understood as the capacity to identify reasons to action and to weigh them in order to reach normative verdicts about what we should do. Rather, we should conceive of the will as practical reason, understood as the capacity to engage in pieces of reasoning that conclude not in normative judgments but in intentions. Having argued for the conception of the will as practical reason, I turn to the standard model of action, according to which our actions are caused by belief-desire pairs. I argue that once we abandon the notion of motivational forces, desires (in the broad sense in which supporters of the standard model use the term) can only be understood as dispositions to decide to act in light of certain considerations and, consequently, that the standard model collapses on the volitionalist model. That concludes my defense of the view that we are not moved by desires, but rather determine our own behavior through the exercise of our will. Lastly, I argue that the will should be understood not as a capacity to decide in light of our beliefs, but rather as a capacity to decide in light of facts – a capacity that is not perfectly exercised when we decide to act in light of a belief (even if it is true).

Key words: Will; Motivation; Desires; Goals; Reasons.

Contents

Introduction	7
1. The volitionalist model.....	7
2. Overview of the Argument.....	13
1. Why do we need the notion of Will	18
1. Introduction.....	18
2. Motives and Incentives.....	20
3. Multiple-incentives cases.....	22
4. Multiple-incentives cases, Desires and the Goals of our Actions.....	27
5. Davidson Causalism.....	28
6. Motivational Strength.....	33
7. Making room for the agent.....	37
8. Why do we need the notion of Will?.....	44
2. Willing, Weighing, Planning	48
1. Introduction.....	48
2. Reasons-to-Motivation Model.....	52
3. The Reasons-to-Judgment Model.....	56
4. Disregarding Reasons.....	60
5. Reasons-to-Decision Model.....	64
6. The Will as the Capacity for Planning.....	67
7. Choosing a means.....	73
3. The Will as a Capacity for Practical Reasoning	76
1. Introduction.....	76
2. The Will as the Capacity to Choose for Reasons.....	77
3. Desires and the Will.....	92
4. Akrasia and Motivational Forces.....	99
5. Taking Stock.....	105
4. Desires, Beliefs and the Explanation of Action	107
1. Introduction.....	107
2. Reasons for Action.....	109
3. The standard model as a Reductive View.....	114
4. Desires and the Explanation of Practical Reasoning.....	121
5. Conclusion.....	127
5. Acting in Light of a Fact and Acting in Light of a Belief	129
1. Introduction.....	129
2. Motivating Reasons and Acting in Light of Beliefs.....	132
3. A Counterexample.....	135
4. The Disjunctive Model.....	136
5. Disjunctivism about Considerations.....	141
Conclusion	147
References	150

Introduction

1. The Volitionalist Model

Philosophers have long wondered about the psychological basis of motivation. During the modern period, a central debate concerned the relative roles of reason and the passions in the production of action. The contemporary heirs of this debate prefer to ask whether normative judgments, concerning our reasons for action, can move us to act by themselves or if they require the aid of desires. Some argue that only desires can move us. Some argue that only the recognition of reasons can move us. And yet others argued that both judgments and desires are sources of motivation and can conflict. The question this dissertation deals with also concerns the sources of motivation. The main idea I shall defend is that the will is the source of our actions, at least when we act for a reason, and that its activity cannot be reduced to the motivational effects of desires or normative judgments.

The view can be summarized as follows: we are endowed with a will, a capacity to make decisions. When we act intentionally and for a reason, the activity of the will is part of the etiology of the action. That is, part of what explains an action is the fact that the agent has exercised her will so as to decide to act in light of a particular consideration. The activity of the will that leads to action consists in practical reasoning. Indeed, the will just is practical reason. But practical reason is not a capacity to engage in pieces of reasoning that conclude in normative judgments about our reasons for action. Rather, it is a capacity to engage in pieces of reasoning that conclude in an intention. Following Hieronymi (2011), we can understand practical reasoning as the activity of settling the question of whether to perform a particular action. The considerations that settle the question for the agent are the reasons in light of which the agent decides to act. As such, the will is a capacity to form intentions in light of certain considerations the agent treats as reasons to act. Intentions, rather than being caused by other mental states (such as desires), are the product of the activity of the will. Intentions are best understood as plans of action, that specify a goal and a strategy to achieve that goal. To form an intention is to settle on a plan of action. Once a plan is in place, it will lead to action when the time comes (as long as it is not forgotten or revised). According to this view, then, what explains an action is not simply a mental state, such as a belief-desire pair or a normative judgment, but rather the complex fact that (i) the agent, through the exercise of her

will, decided to act in light of a consideration, (ii) thereby formed a corresponding intention and (iii) eventually executed that intention.

These claims form the core of what I will call the *volitionalist model* of our motivational psychology. This presentation of the model is bound to raise a number of questions. To some, this view may seem perfectly trivial, to others, highly implausible or even a version of a long-refuted theory. Some clarification is in order.

1.1. The Modern Theory of the Will

First, it is important to distinguish the view I am going to defend from what Hyman calls the modern theory of the will (Hyman, 2015, p.1). What is distinctive of the latter view is the claim that we have to postulate the will in order to distinguish voluntary from involuntary actions. A voluntary action is one that originates in the will. In particular, what makes an action voluntary is the fact that it is caused by a conscious choosing or willing, a *volition*, which is an act or an operation of the will.

This view is vulnerable to a very compelling objection. As Ryle (1949, p.67) has argued, it leads to a dilemma: are the volitions, the acts of the will, themselves voluntary or involuntary? If a volition is voluntary, then it must issue from a prior volition and that from yet another and so on – we face a regress. If it is involuntary, then how can the actions that follow from it be voluntary? If an involuntary thought makes me blush, and I cannot stop myself from blushing once the thought has occurred to me, then my blushing is involuntary as well. It would seem that in much the same way, if an involuntary volition makes the act, and I cannot stop myself from acting given the volition, then the action that issues from the volition is involuntary as well.

Ryle's dilemma is a powerful objection to the theory of volitions considered as a theory of voluntariness. Because the modern theory postulates the will in order to explain voluntariness, it is vulnerable to the objection. Although the view I am going to defend shares the idea that the actions we perform for a reason are always the product of the activity of the will, it differs from the modern theory of the will in that it does not postulate the will in order to distinguish between voluntary and involuntary actions. I hold that whenever we act for a reason, the activity of the will is part of the etiology of the action and its role cannot be reduced to the role of mental states as desires or normative judgments. This view is perfectly compatible with the claim that an action that is the product of the activity of the will can be

involuntary (if, for instance, it was performed due to duress or coercion). It also admits that, to the extent that actions that are not done for reasons can be voluntary, actions that do not originate in the will can be voluntary. If actions that are not done for reasons cannot be voluntary, the explanation of that fact is not to be found in the fact that they do not originate in the will, but in the connection between voluntariness and acting for a reason.

Rather than postulating the will in order to distinguish between voluntary and involuntary actions, the view I am going to defend postulates the will as a way of accounting for a special ability we have. We are capable of setting goals for our actions, not only in the sense of deciding to do something in light of the fact that it will contribute to something that we care about, but in the sense that we can decide to do something with a view to a particular goal, even when it would make perfect sense for us to perform the same action with a view to another goal. We can, for instance, decide to act with a view to helping someone, and only that, even when we know that the action we will perform is also a means to another goal we cherish. It is in order to account for this ability that we have to postulate the will – or so I shall argue.

One could insist that Ryle's dilemma still poses a threat. If actions originate from the will and acts of the will are themselves actions, then we have a potential regress in our hands. But this objection rests on a misconception of the volitionalist model I will defend. According to this model, whenever we act for a reason, our action is the product of the activity of the will. The activity of the will, however, is not itself an intentional action done for a reason. That does not mean that it is something that merely occurs to the agent either. Rather, it is an active process, a process of practical reasoning, of which the agent is the subject. Even if it is not an intentional action done for a reason (at least not always), practical reasoning is something the agent does.

1.2. Volitions

It is also important to distinguish the view I defend from other forms of contemporary volitionalism. It is usual to distinguish between reductionist theories of action and anti-reductionist theories of action. According to the latter, the role the agent plays in the production of her action cannot be reduced to the role mental states or events play in that process. According to the former, the role the agent plays in the production of her action can be reduced to the role mental states or events play in that process. According to one

reductionist view, for instance, for an agent to decide to act in light of a certain consideration is simply for her action to be caused in the appropriate way by a belief-desire pair of hers. In that way, the contribution of the agent to the production of the action (the decision) is reduced to the operation of certain mental-states. In contrast, anti-reductionist views hold that this reduction cannot succeed.

A somewhat popular anti-reductionist view holds that acts of the will (volitions) are basic mental actions that can cause the agent's body to move but are themselves not caused by other mental states, such as desires or normative judgments. Rather, they are the result of the agent's exercise of her power or capacity to will. Because this capacity is not reduced to other mental states, the agent has an irreducible role to play in the production of actions. Views along these lines have been defended by Ginet (1997) and Lowe (2008, p.148). According to these views, volitions are the immediate cause of the action or, more precisely, the agent's movement. In order to cause her arm to move the agent wills her arm to move, that is, forms a volition that her arm is to move. This volition is a mental action that amounts to her trying to move her arm. Indeed, Ginet holds that when the agent succeeds in moving, the volition is perceived by the agent as a feeling that she made the movement occur (Ginet, 1997, p.89). For that to be the case, the volition must be something that accompanies or immediately precedes the movement.

This is not the view I am going to defend. According to my view, the will is a capacity for practical reasoning. To exercise the will is to engage in practical reasoning. And the product of practical reasoning is an intention, a plan. Of course, a plan is not an action, so what results from the activity of the will is not a mental action. Whether executing the plan by performing certain bodily movements will require some basic mental actions is a question about which my view remains silent. One could argue that the very activity of engaging in practical reasoning can be seen as a mental action. And even though, according to my view, an episode of practical reasoning cannot cause a bodily movement directly, it can result in the adoption of a plan which can eventually lead to action. In that sense, the activity of the will can cause actions. But even if we admit that, acts of the will, as I conceive of it, fall short of the volitions postulated by Ginet and Lowe. According to them, a volition can amount to the act of trying to do something: trying to move one's arm is constituted by the agent's willing her arm to move. But clearly, settling on a plan in light of a consideration cannot amount to trying to move. Furthermore, Ginet holds that a volition can have the phenomenological

quality of seeming to the agent as if she made the movement of her body occur. But the activity of the will as I conceive of it, cannot be experienced by the agent as the feeling that she is causing her movements, simply because a complete exercise of the will may take place long before any action or even fail to lead to action (as when we form a plan but later forget about it completely).

Therefore, the will, as I conceive of it, is not a power to produce volitions, understood as basic mental actions, but rather a capacity for practical reasoning. Nevertheless, the view I am going to defend qualifies as an anti-reductionist theory of action. It holds that the agent has an irreducible role to play in the production of action. As I said, according to my view, what explains an action is the complex fact that the agent engaged in practical reasoning, thereby formed an intention and eventually executed it. The role the agent plays is that of the reasoner. And the activity of reasoning or deciding cannot be reduced to the operation of other mental states, such as desires. Therefore, even if agents do not have an irreducible role to play in initiating movement and action, they play an irreducible role upstream in the causal chain that leads to action.

1.3. The Will and Desires

A third clarification concerns the relation between the will and desires or pro-attitudes in general. According to the volitionalist model I am going to defend, actions are explained by a complex fact that includes the fact that the agent exercised her will in a particular way. Does that mean that the activity of the will cannot be explained by appeal to further mental states, such as desires and beliefs? No. The volitionalist model is compatible with the view that belief-desire pairs can explain, even cause, the activity of the will that figures in the explanation of the action. But is not the activity of the will reduced then to a mere epiphenomenon, an unnecessary step between desire and action? And is it not false then that the will is the source of our actions? If our desires cause the activity of the will, is it not true that the source of our actions is in our desires, not in the will? I do not think so.

The word “desire” can be understood in more than one way. What we usually mean by “desires” is what I will call, in chapter 3, “substantive desires”. A substantive desire is a mental state such that it makes sense to ask whether or not an agent desired to act as she did. In this sense, one can, for instance, attend a meeting even though one has no desire to do so. One can decide to do something one has no substantive desire to do. In that case, the

source of the action is in the will and the activity of the will that results in the decision is not explained nor caused by a substantive desire. Clearly, this is not the sense of “desire” that raises problems for the volitionalist model.

Philosophers usually use “desire” in a different sense. In this sense, it is impossible to intentionally do something one does not want or desire to do. In this sense, if one decides to go to the meeting, then one has a desire to go to the meeting. It is in this broad sense of “desire” that it makes sense to claim that the activity of the will can always be explained, even caused, by desires. The problem is to explain what desires, in this broad sense, are. According to the volitionalist model, desires in this sense are simply dispositions to decide in light of certain considerations. To decide to act in light of a consideration is to engage in practical reasoning – is to settle the question of whether to act. Now, that is exactly what the activity of the will consists in. Therefore, desires, in this broad sense, are nothing but dispositions to engage in the activity of the will in a particular way. They can only manifest in the agent engaging in practical reasoning in a particular way. Even if desires cause the activity of the will, therefore, the episode of practical reasoning in which the exercise of the will consists is not a dispensable step in the causal chain that leads to action. Desires, in the broad sense, can only be understood by reference to that very activity.

Admitting that desires cause the activity of the will, which results in intentions, which, in their turn, lead to action, is perfectly compatible, therefore, with an anti-reductionist theory of action. Even if the desires cause the agent’s practical reasoning to take a certain course, it is still the agent that is doing the reasoning. Given that the episode of practical reasoning is an indispensable link in the causal chain leading to action, the agent still has an irreducible role to play in the production of action.

Is it not true, however, that once we admit that desires can cause the activity of the will we have to admit that the source of our actions is in our desires and not on the will? No. Because desires, in the broad sense, can only be understood as dispositions to decide in light of certain considerations, to ascribe a desire to an agent is simply to register her disposition to reason practically in a particular fashion. To be moved by a desire is to manifest that disposition, and when the agent manifest that disposition, she is moved by the activity of her will. I defend this view in chapter 4.

2. Overview of the Argument

According to the volitionalist model, then, when we act for a reason, the activity of the will plays an irreducible and central role in the etiology of the action. My defense of this model will emerge gradually from the criticism of alternative models.

According to a popular view, our motivational processes take the following form: certain mental states (such as beliefs, belief-desire pairs or belief pairs) or certain mental operations (such as the weighing of *pro tanto* reasons) produce in us motivations to perform certain actions. These motivations are conceived as forces. They have a particular intensity and a certain direction. Motivations whose directions coincide can combine their strengths, thus producing a stronger composite motivation. The agent is moved by the strongest motivational force at play in the struggle for the determination of her behavior. This simple schema can be developed in a number of ways. One can hold that the agent's motivations always correspond to or are determined by her belief-desire pairs. According to the crudest version of this view, the agent is completely passive with respect to the determination of her behavior – she is effectively reduced to the condition of an observer of the power struggle that takes place within her. Some philosophers supplement this view with the idea that the agent has an active role to play in the creation and suppression of non-derived desires and, in that way, can contribute to the determination of her own behavior. Others suggest that the agent can be identified with a particular desire, namely, the desire to act according to her reasons, and can take part in the determination of her behavior by forming normative judgments that direct the motivational force of that desire. Yet another option is to hold that motivational forces are not produced by belief-desire pairs, but rather by normative beliefs, or pairs of normative beliefs and factual beliefs or by the mental operation of weighing *pro tanto* reasons. In this case, the agent has an active and direct role to play in the production of her motivations and, consequently, in the production of her actions.

I will argue, in chapters 1 and 2, that this model, independently of how it is developed, is false. We are not moved by motivational forces in dispute, regardless of what the source of these motivation happens to be and regardless of the extent of the control we have over them.

My starting point, in chapter 1, will be the discussion of the question about how are the intentions with which we act in a particular case determined. A statement of the form

“agent S did action A with the intention of G-ing” informs us about the goal at which action A was aimed. The question is how are the goals at which our actions aim determined. It is commonly held that the goal at which an action aims (and, therefore, the intention with which the agent acts) is specified by the belief-desire pairs in light of which the action seems appealing to the agent. In fact, I will argue in chapter 1 that this view is implied by the idea that our behavior is determined by motivational forces which correspond to our belief-desire pairs. I will refer to this idea as *hydraulic model* of our motivational psychology, because it is a way of expressing the view that our behavior is determined by the power struggle between our impulses. I then argue that the view that the goals of an action are fixed by the belief-desire pairs that render the action appealing to the agent is false. My argument is simple: there are *multiple-incentives cases* (that is, cases in which an agent has more than one incentive to act but in which her motive for acting corresponds to only one of these incentives) and the view that the goals of our actions correspond to our belief-desire pairs is incompatible with the existence of these cases. This conclusion has two important consequences. First, given that the view in question is entailed by the hydraulic model, it follows that this model is false. Second, if the goals at which our actions aim are not passively determined by the belief-desire pairs we happen to have, then these goals must somehow be actively determined by the agent. I refer to the capacity agents have to actively determine the goals at which their actions aim as their “will”. The activity of the will fixes the goals of our actions by producing a particular intention. If my arguments are correct, they show that we need the notion of the will in order account for the existence of multiple-incentives cases.

I should emphasize that throughout chapter 1 I discuss the position of a number of philosophers assuming that they subscribe to the hydraulic model. For instance, I argue that Davidson’s causalism fails to accommodate multiple-incentives cases as long as we hold on to the hydraulic model. Davidson, however, most likely does not subscribe to the hydraulic model. One can hold that we are moved by desires but reject the idea that they move us as motivational forces. In chapter 1, my target is the latter idea. However, I do return to the former idea in chapter 4 – where I argue that once we have abandoned the idea of motivational forces, the view that we are moved by desires collapses on the volitionalist model.

Now, if the goals at which our actions aim are determined by intentions that are the product of the will, how should we conceive of the will and the intentions it produces?

One possibility is to identify the will with practical reason. The content of this conception of the will depends, of course, on how we understand the notion of practical reason. One intuitive option is to think of practical reason as the capacity to identify certain considerations as reasons for acting or refraining from acting and to weigh these considerations in order to arrive at a verdict about what one should do. According to this suggestion, intentions are normative judgments about what we should do, or about what we have most reason to do. I refer to this capacity to identify and weigh reasons as “judicative reason”. Chapter 2 discusses the proposal of identifying the will with judicative reason. There are two ways to understand this proposal. According to what I call the *reasons-to-motivation model*, the *pro tanto* reasons that are acknowledged by the agent produce a motivational force in the direction of the action they favor. The strength of these motivations corresponds to the weight the agent ascribes to the reason. This guarantees that the agent will always be more motivated to perform the action she believes she has most reason to perform (except in cases of *akrasia*). I argue that this model is simply a variation of the hydraulic model and, as such, faces exactly the same objections. According to what I call the *reasons-to-judgment model*, we should abandon the notion of motivational forces. This model holds that we are not moved by motivational forces. Rather, we are beings that move from the consideration of *pro tanto* reasons to intentions, conceived as normative judgments, and then, if everything goes well, execute these intentions. I believe that rejecting the notion of motivational forces is a move in the right direction. Nevertheless, I argue that as long as we take intentions to be normative judgments, we still cannot account for multiple-incentives cases. By the end of chapter 2, I argue that in order to account for such cases we need to conceive of intentions as plans of action. I conclude that we should think of the will as a capacity to adopt plans in light of certain considerations.

In chapter 3, I defend a positive conception of the will. The conclusion that intentions are plans of action may suggest that the formation of an intention cannot be the result of a process of practical reasoning. Only judgments, one could argue, can be the conclusion of a piece of reasoning. I dispute this claim. If we conceive of practical reason as the capacity to reach normative judgments, then we have to distinguish between the will and practical reason. That becomes clear when we consider cases of *akrasia* and decisions made in circumstances of normative uncertainty. These cases suggest that the will is an executive capacity whose job is to convert the normative judgments into intentions. But this conception of the will obscures the fact that we form intentions *for reasons*. We decide to act and form

intentions in light of certain considerations. These considerations are the reasons in light of which we decide. Thus, given a broader understanding of reasoning (according to which any process by which we come to form, revise, or sustain an attitude for a reason counts as reasoning), intentions can be seen as conclusions of pieces of reasoning. What cases of akrasia and decisions under normative uncertainty show is that we should distinguish between theoretical reasoning regarding practical matters (a kind of reasoning which results in normative judgments) and practical reasoning (a kind of reasoning which results in intentions). The will, I argue, should be identified with our capacity for practical reasoning in this sense. After defending this point, I explore the relation between the will and our desires. In particular, I try to put to rest the concern that by abandoning the notion of motivational forces we obscure the way in which desires interfere with our decision process. I argue that it is an illusion to suppose that talk of motivational forces provides any advantage over the volitionalist model when it comes to explaining that interference.

In chapter 4, I return to the claim that we are moved by desires. In chapters 1 and 2, I rejected the idea that we are moved by desires conceived of as motivational forces. But that is not the same as showing that we are not moved by desires. One could argue that even if there is no such thing as motivational forces, we are moved by desires because belief-desire pairs cause our actions. I refer to the view that rejects the notion of motivational forces while holding that whenever we act for a reason our action is caused by a belief-desire pair as the *standard model*. In chapter 4, I argue that once we abandon the idea of motivational forces, the standard model collapses on the volitionalist model. The standard model does provide a genuine alternative to the volitionalist model if we take it to be a reductive account of what it is to decide to act in light of a consideration. However, once we reject the notion of motivational forces, this reductionist project fails because we have to understand desires as dispositions to decide to act in light of certain considerations. If we adopt a non-reductive reading of the standard model, according to which the activity of deciding to act in light of a consideration is not reduced to the operation of belief-desire pairs, then it presupposes that we are capable of engaging in practical reasoning and that the process in which practical reasoning consists plays an irreducible role in the production of action. But this is exactly what the volitionalist model holds. At this point the standard model is no longer an alternative to but a version of the volitionalist model. One could still insist that the standard model differs from the volitionalist model in that it holds that we are moved by desires. But, given that the

desires that figure in the standard model are simply dispositions to decide to act in light of a certain consideration, the claim that we are moved by desires is reduced to the claim that whenever we decide to act in light of a consideration, we manifest a disposition to decide to act in light of that consideration. This is perfectly trivial and in no way conflicts with the volitionalist model.

Finally, in chapter 5, I turn to a problem regarding how exactly to characterize the will. I defended the view that the will is a capacity to decide in light of certain considerations. But there is more than one way in which to understand this claim. One option is to understand it as the claim that the will is a capacity to decide to act in light of certain beliefs. This option is problematic. We conceive of ourselves as beings capable of deciding to act in light of normative reasons. And normative reasons are facts, not beliefs. Therefore, we should be able to decide to act in light of facts. One could suggest that to decide in light of a fact is simply to decide in light of a true belief. This suggestion is supported by an argument from error cases. In cases in which we decide in light of a belief that turns out to be false, we cannot be said to have decided to act in light of fact. In these cases, we decide to act in light of a belief. Given that from the subjective perspective of the agent there is no difference between error cases and non-error cases, we should conclude, according to this argument, that the agent is doing the same thing in all cases. That is, we always decide to act in light of a belief. To decide in light of a fact is simply to decide in light of a true belief. I think this conclusion is false and I offer a counterexample to it. The question that remains is how to defuse the argument from error cases. I argue that the only way to do that is to accept a disjunctive conception of what it is to decide in light of a consideration. According to this view, deciding in light of a belief (be it true or false) and deciding in light of a fact are different (although possibly subjectively indistinguishable) ways of deciding to act in light of a consideration. It is only when we decide to act in light of a fact that we perfectly manifest our capacity for practical reasoning. The will, therefore, has to be conceived as a capacity to decide to act in light of *facts* – a capacity that is not perfectly exercised when we decide to act in light of a belief (even if it is true). This view allows us to defend the claim that when we act in light of a normative reason, the normative reason itself (and not some psychological state) is the reason that move us. This is an important result, one that Jonathan Dancy tried but ultimately failed to establish in his *Practical Reality* (2000).

1. Why do we need the notion of Will

1. Introduction

A statement of the form “agent S did action A with the intention of G-ing” informs us about the goal at which action A was aimed. It has the same content as “agent S did action A in order to G”. For instance, to say that “she ran with the intention of catching the bus” is just to say that she ran in order to catch the bus. Likewise, “he added sage to the stew with the intention of improving its taste” is the same as “he added sage to the stew in order to improve its taste”. It is commonly held that the goal at which an action is aimed and, therefore, the intention with which the agent acts (i.e., the content of the G-slot in the statements above) is specified by the belief-desire pairs in light of which the action seems appealing to the agent. Thus, “he added the sage to the stew in order to improve its taste” is correct if the agent wanted to improve the taste of the stew and believed that adding sage to the stew would do just that.

This view is strongly associated with Davidson¹ but it is shared by a number of philosophers. As a matter of fact, I shall argue that it is entailed by a widely shared view about motivation. This view is composed by three theses: (a) we are directly moved by belief-desire pairs: a pair composed by a pro-attitude towards G and the belief that doing A is conducive to G (what I will call a pro-A pair) motivates us to perform A; (b) belief-desire pairs differ in strength² and when an agent has several pro-A pairs these combine their strengths to produce a stronger motivation to perform A (I refer to this as the thesis of compositionality)³ and (c) when faced with appealing but incompatible alternative actions we perform the action that we are more strongly motivated to perform. I refer to this set of theses as the hydraulic model⁴ of

1 Davidson claims that we can explain an action by indicating “what it was about the action that appealed” to the agent, that we do so by presenting the “primary reason why the agent performed the action” which is nothing but a belief-desire pair (Davidson, 1980, p.3-4) and that to “know a primary reason why someone acted as he did is to know an intention with which the action was done” (Davidson, 1980, p.7).

2 The strength of a belief-desire pair or of the motivation it produces is probably a function of the strength of the pro-attitude towards G, the agent’s estimation of how likely it is that performing the action A will bring about G and the degree of confidence of the agent in that estimation. The details are irrelevant to the arguments that follow.

3 It is not part of this thesis that compositionality is linear, so that, for instance, if two desires of equal strength combine then the resulting motivation is twice as strong as each of them. All that follows from it is that if desires D_1 and D_2 are equally strong, then the motivation produced by set $\{D_1, D_3\}$ is to some degree stronger than the one produced by set $\{D_2\}$. These sets of desires are what Mele calls the “motivational base” of the motivation to act (see Mele, 1992, p.58-60).

4 The expression comes from McDowell (2002: 213) and Wallace (2006: 55)

our motivational psychology because it is naturally understood as a way of expressing the view that our behavior is determined by the tug of war between impulses or forces within us.

The hydraulic model is explicitly upheld by philosophers such as Velleman (I discuss his position in section 7) and many others are implicitly committed to it. For instance, this view is what is behind Schroeder's claim that since "your desires are what motivate you to act" you can act according to your moral reasons only to the extent you have a "collection of desires whose strengths match the weights of [your] independently existing reasons" (Schroeder, 2007: 169). A case can also be made for the claim that Hume upheld this view, but it is important to notice that the hydraulic model is compatible with an anti-Humean theory of motivation. Dancy (2000, p.85-7) describes (but does not subscribe to) a view he calls *pure cognitivism*. According to this view beliefs produce motivation directly. It allows for the claim that where there is motivation, there is desire but only because it conceive of desires as the very state of being motivated, and not as a part of what motivates us. The motivation produced by beliefs, however, differ in strength and when there is a conflict of motivation we are moved to action by the stronger motivation. Pure cognitivism is, therefore, an anti-Humean thesis but a version of the hydraulic model nevertheless.

My primary goal in this chapter is to reject the view that the goals our actions aim at are specified by the belief-desire pairs that render the action appealing to the agent. My argument is simple: there are multiple-incentives cases, i.e., cases in which the agent has more than one incentive to act but in which her motive corresponds to only one of these incentives, and the view that the goals of our actions are specified by our belief-desire pairs is incompatible with the existence of these cases. This conclusion has two important consequences. Given that, as I shall argue, that view is entailed by the hydraulic model, it follows that we should reject the hydraulic model and with it the idea that we are directly moved by belief-desire pairs. The other consequence is this: if the goals our actions aim at are not passively determined by the belief-desire pairs that render them appealing, then they must be somehow actively determined by the agent. I refer to the capacity to actively determine the goals one's actions aim at as the agent's "will". If my arguments are successful, therefore, they show that we need the notion of will in order to account for multiple-incentives cases.

In section 2, I discuss the notions of motive and incentive. In section 3, I introduce the notion of multiple-incentives cases. In section 4, I argue that the hydraulic model entails

the view that the goals of our actions are determined by the belief-desire pairs that render them appealing to the agent and that this view is incompatible with the existence of multiple-incentives cases. In section 5, I consider whether Davidson's causalism provides a way in which to hold on to the idea that we are moved by our belief-desire pairs while rejecting the hydraulic model. Sections 6 and 7 are concerned with attempts by the supporters of the hydraulic model to reject the claim that this view entails that the goals our actions aim at are determined by our belief-desire pairs. Finally, in section 8, I argue that introducing the notion of the will allow us to account for multiple-incentives cases.

2. Motives and Incentives

Attributions of motives usually take the form "person *P* did action *A* because *M*", as in "she ran because her bus was about to leave" or "he added sage to the stew because it would improve its taste". Statements of this form ascribe motives when the fact that fills in the *M*-slot renders the action intelligible as an action aimed at a particular goal. If I say something like "she moved her leg because I hit her knee" or "he yawned because the person next to him yawned" I am simply asserting the cause of the action, not ascribing a motive to the agent. Usually, therefore, the following entailment holds:

MOTIVE-GOAL LINK: If *M* is the motive for which agent *P* did *A*, then *A* aims at a goal *G* and *M* is part of what explains why doing *A* is an effective or necessary means to *G*.⁵

5 When I say that the entailment "usually holds" I mean it. There is a number of exceptions to the motive-goal link. If "she ran because she had an appointment" is a correct motive ascription in a particular case, it follows that the action of running was aimed at a goal, namely, getting at her appointment in time, but the fact that the agent had an appointment does not contribute to the explanation of why running is a necessary or effective mean to getting at her appointment in time. Rather, the fact that she had an appointment helps to explain why she chose to perform an action with a view to that particular goal. Sometimes, therefore, an agent's motive explains her acting with a view to a particular goal instead of rendering intelligible the performance of a particular action in light of the goal it aimed at. There are also cases of error in which the agent acts under the belief that something is the case when it is not. Thus, for instance, if in the example above the bus was not about to leave, we would say that "she run because she thought the bus was about to leave". Of course, the fact that she believed the bus was about to leave (when in fact it was not) does nothing to explain why running was a necessary means to catching the bus (even if it explains why the agent thought so and, therefore, engaged in that action). It is commonly thought that the possibility of this kind of error shows that our motives are always, even in cases in which we are not deceived about our circumstances, provided by mental facts instead of facts about the circumstances of action. I do not think that is the case – a form of disjunctivism about motives can be defended. According to this view, in cases in which we are actually responding to facts about our situation those facts provide our motive for action, whereas in cases in which we are not responding to facts because we are deceived our motives are provided by our mental states. I cannot defend this point here, however. The arguments in this chapter should work even if one holds that our motives are always provided by mental states (see note 8). For now, I wish simply to remark that even though there are important exceptions to the motive-goal link as formulated, this conditional holds in relation to an important class of motive attributions, namely, those motive attributions in which the motive

Thus, for instance, if “she ran because her bus was about to leave” is a correct motive ascription, then the action of running aims at a goal and the fact that the bus was about to leave helps to explain why running was conducive to that goal. We can easily infer that the action of running was aimed at catching the bus⁶ and the fact that the bus was about to leave explains why running was necessary to the achievement of that goal. In a similar way we can immediately see that the action of adding sage to the stew aims at improving its taste and the fact that adding sage to the stew would improve its taste explains, vacuously, why doing just that contributes to the achievement of that goal. This connection between motive ascription and the goal at which the action aims is made even clearer if we pay attention to the fact that both “she ran because the bus was about to leave” and “she ran in order to catch the bus (which was about to leave)” are equally satisfactory and roughly equivalent answers to the question “why did she run?”, even though the latter is not an ascription of motive but simply states the goal at which the action was directed.

The correction of a motive ascription also entails that the agent has (i) a pro-attitude towards the goal G her action A aims at (she either desires G , wants G , prizes G , is inclined towards G , etc.) and (ii) believes that performing A is either a means to G or constitutes G .⁷ Thus, for instance, if “she ran because the bus was about to leave” is a correct motive ascription, then the agent in this example wants to catch the bus and believes running to be an effective means to achieve that goal.⁸

Let me now introduce the notion of an *incentive*:

corresponds to an incentive to act that the agent recognizes. In this chapter I will focus on this class of motive attributions.

- 6 These inferences, of course, depend on a series of suppositions about the situation in which the agent finds herself, about available alternative courses of action, about the agent herself and, perhaps, about the human form of life in general. It clearly is not a deductive inference. Most likely it is a kind of inference to the best explanation.
- 7 This claim is shared by a number of philosophers, such as Nagel (1978: 29-30), McDowell (1978: 15), Smith (1994: 116) and Dancy (2004: 85).
- 8 What follows from this is that the presence of a corresponding belief-desire pair is a condition of correction for a motive ascription. It is very common to hold, however, that our motives are always constituted by the corresponding belief-desire pairs (or better, by the fact that we have those belief-desire pairs). I will call this the Foreground View. In contrast, the Background View holds that the appropriate belief-desire pairs are a background condition for the correction of a motive ascription but do not constitute the motive itself. The Foreground View is problematic, especially because it entails that the motives for which we act never correspond to the facts that provide us with reasons to act, which are not facts about our mental states. An objection along these lines has been pressed by Dancy (see 2001, p. 103-106). The argument of this chapter, however, should work even if one holds the Foreground View. When pertinent I will add notes to clarify how the argument should be understood by someone who holds the Foreground View.

INCENTIVES: A fact *I* is an incentive for agent *P* to perform action *A* if *A* is an available course of action for *P*, *P* has a pro-attitude towards *G* and *I* is part of what explains why doing *A* is an effective or necessary means to *G*.

Thus, for instance, the fact that the bus is about to leave is an incentive for the agent in our example to run because she wants to catch the bus and the fact that the bus is about to leave explains why running is a necessary means to catch the bus. And the fact that adding sage to the stew would improve its taste is an incentive for an agent to do so if she desires⁹ to improve the taste of her stew given that the fact that adding sage to the stew would improve its taste explains, vacuously, why adding sage to the stew is an effective means of improving its taste.

An agent recognizes an incentive *I* for doing *A* when she realizes that, in light of fact *I*, doing *A* will promote (or is necessary to promote) goal *G*, which she happens to desire or want or prize, etc. Thus, an agent recognizes the fact that the bus is about to leave as an incentive for her to run when she realizes that, in light of the fact that the bus is leaving, running is a necessary means to catch the bus (which is something she wants to do). And an agent recognizes the fact that adding sage to the stew would improve its taste as an incentive to do so when she realizes that, in light of that fact, adding sage to the stew is an effective means of improving its taste (which is something she wants). An agent recognizes an incentive *I* to perform action *A* if, and only if, she has a belief-desire pair composed by a pro-attitude towards *G* and a belief that *A* is conducive to *G*.

When a correct motive ascription entails that the agent had a pro-attitude towards the goal her action aimed at and the belief that performing that action was a means to that goal, the agent's motive is an incentive she happens to recognize. That was the case in the examples considered so far. There are cases, however, in which despite recognizing several incentives to perform a particular action the agent's motive in performing the action in question corresponds to only one of those incentives.

3. Multiple-incentives cases

Consider the following example:

VOLUNTEER WORK: Mary is a really benevolent person. She cares for the well-being of others and does what she can to promote their well-being.

9 As is usual in philosophical discussion, in what follows I will use "desire" to refer to pro-attitudes in general.

One fine day she learns that a local soup kitchen is in need of volunteers. She has a few free hours that she could spend in the soup kitchen. She decides to volunteer there because it will contribute to relieving the suffering of people in need. As a matter of fact, she also wants to be admitted to the University next semester and believes, from what she hears, that volunteer work increases one's chances of being admitted. She is well aware of that, but that is not why she volunteers.

In this example Mary recognizes two incentives to volunteer at the soup kitchen: the fact that it will contribute to relieving the suffering of people in need and the fact that it will increase her chances of being admitted at the University. Her motive corresponds only to the former incentive, however. I will refer to cases such as this, in which the agent recognizes several incentives to perform an action and yet acts for a motive that corresponds to only one of those incentives, as *multiple-incentives cases*.¹⁰

Tracking the agent's motive in such cases is relevant to the moral assessment of their actions. Surely, we are all inclined to think that Mary's action is more praiseworthy than the action of another agent, Pete, that volunteers at the same soup kitchen and invest the same amount of time in that activity but does it as a way of improving his chances of being admitted at the University. The difference of course is that Mary's motive is other-regarding while Pete's motive is selfish, even though Mary also recognizes a self-interested incentive to perform the action in question.¹¹

The existence of multiple-incentives cases may be contested.¹² One may hold that if an agent recognizes more than one incentive to perform action *A*, then her motive for doing *A* must be a compound-motive that combines all the incentives she recognizes. The correct

10 If one holds the Foreground View, multiple-incentives cases must be understood as cases in which an agent performs an action *A*, has more than one pro-*A* belief-desire pair and only one of these pairs constitutes the motive for which she performed action *A*.

11 In order to avoid an obvious objection, we must distinguish between the moral desirability of an action and its moral worth. A consequentialist will not accept that the assessment of an action as morally right, wrong, desirable or undesirable depends on the motive for which the agent acts. He may, however, accept that the extent to which an action is worthy of praise or blame (that is, its moral worth) depends on the motive for which the agent acted. He may, for instance, claim that the moral worth of an actions is a matter of how it reflects on our assessment of the agent that performs it. See Arpaly (2002, p.224-5) for a discussion of this distinction.

12 I should emphasize, however, that the existence of multiple-incentives cases is widely acknowledged. See, for instance, Dancy (2000, p.161-2), Wallace (2006a, p.61), Dickenson (2007, p.3-4) and Setiya (2007, p.39-40). More importantly, Davidson, the most prominent supporter of the view that our actions are caused by desires, clearly recognized multiple-incentives cases: "[...] you may err about your reasons, particularly when you have two reasons for an action, one of which pleases you and one which does not. For example, you do want to save Charles pain; you also want him out of the way. You may be wrong about which motive made you do it" (Davidson, 1980, p.18).

motive ascription in Mary case would then be “she volunteered at the soup kitchen because it would contribute to relieving the suffering of people in need and would increase her chances of being admitted at the University”. And that would entail that her action aimed both at relieving the suffering of people in need and at furthering her admission at the University.

We do, however, assume that multiple-incentives cases are possible. First, is it simply a fact that we usually point out *the* motive for which an agent acted even in cases in which the agent had several incentives to act as she did. Just think of how often people say things like “I did not do it for the money” (granted this is very often a false statement, but one that we assume can be true on occasion). One can also act for a particular motive when one has several self-interested incentives to act. A smoker, for instance, may recognize several incentives to stop smoking (“I will be healthier. I will save money that can be spend in things I care more about. People will praise my strength of will”) but, nevertheless, quit smoking for a motive that corresponds to only one of these incentives (“What really got me to stop was the concern for my health. Of course, having more money available is a nice perk, but that is not why I quit smoking”).¹³

Second, it is often remarked that a truly virtuous person performs virtuous actions for their own sake.¹⁴ A truly benevolent person, for instance, performs a charitable action not because she will get something out of it, but for its own sake. That is not to say that in performing a charitable action the benevolent agent does not aim at a further end (namely, the promotion of the well-being of others).¹⁵ Rather it is to say that the benevolent person’s action does not aim at procuring a personal advantage for herself. Given the motive-goal link, it follows that the motive for which a truly benevolent person performs a charitable action is not provided by the fact that it promotes a personal interest of hers (or by any fact that explains why the action promotes a personal interest of hers). If that is the case, then the benevolent

13 One could object that this is not a reason to believe that there are multiple-incentives cases because the agent in our example could eventually find out that the motive for which she stopped smoking was actually a compound-motive. But that is beside the point. What matters here is that it is possible for agents to act for a particular motive even when they have several incentives. If it is conceivable that an agent could stop smoking because it would be beneficial to her health, even though she recognized other incentives to stop smoking, then we already have to admit the possibility of multiple-incentives cases. That is perfectly compatible with the view that agents may be mistaken in their self-ascription of motives.

14 Aristotle makes this claim in *Nicomachean Ethics*, II.4.

15 Bernard Williams and Rosalind Hursthouse hold that the claim that the virtuous agent chooses virtuous actions for their own sake even if they aim at further ends should be understood as the claim that the virtuous agent acts for a particular repertoire of reasons (see Williams, 1995, p.17-18 and Hursthouse, 1995, p.24-25). I agree with them.

person's motive to perform a virtuous charitable action is *pure*, in that it does not include self-interested incentives and the same should be true of other virtues. But surely a virtuous person can recognize the self-interested incentives she has to perform a virtuous action. Mary is no less benevolent for having a strong desire to be admitted at the University and believing that volunteering is conducive to that goal (nor, if she falls short of fully benevolence, could she come any closer to virtue by losing that desire or the associated belief). Neither self-denial nor ignorance of the personal benefits that may accrue from moral behavior are conditions of virtue, much less a path to it. Therefore, if virtuous agents do perform virtuous actions for their own sake, then multiple-incentives cases are possible.¹⁶

Third, we recognize cases in which it is important that we do *not* act for an incentive that is available to us. Suppose, for instance, you are in a loving relationship with a very rich person. Suppose further that you have a strong interest in financial stability and knows that maintaining that relationship is an effective way to guarantee that. That provides an incentive for you to maintain your relationship (which entails the performance of a number of actions). But no one will doubt that it is perfectly intelligible that it should matter to your partner not only that you maintain your relationship out of love or mutual care but also that you do not maintain it out of love *and* financial interest. That is clearly not a demand that you stop caring about your own financial stability nor that you lose your knowledge about your partner wealthiness. Therefore, it reflects the belief that you can act out of love only, even though you have financial incentives to act in the same way. Indeed, it may be the case that an act can only be considered a genuine act of love or friendship if considerations of personal gain play no role in moving the agent to it (as an act of devotion to a person or a cause is one to which no consideration of personal gain is mixed).

Fourth, frequently our acts establish certain relations and some of these seem to presuppose that it is possible to single out the motive for which the agent acted. It makes sense for the beneficiary of a charitable act to be grateful to his benefactor but only to the extent that the benefactor is moved by the proper other-regarding motives and not by the

16 This defense of the view that there are multiple-incentives cases does not presuppose any controversial view about morality. It is compatible with deontological as well as consequentialist views. My point is simply that we usually take virtuous agents to act with a view to a determinate goal even when they have other incentives to act. That can be true even if the value of actions is completely independent of the motives for which we act. Certain moral theories, such as Kant's, emphasize the possibility of multiple-incentives cases and hold that these cases *must* be possible if our moral assessments are to make sense. My point is humbler. I am simply pointing out that we usually assume that multiple-incentives cases are not only possible but common.

prospect of personal gain. Unqualified gratitude presupposes then that considerations of personal gain had no role in moving the benefactor. It would be preposterous, however, to claim that unqualified gratitude is not a proper reaction in a particular case simply because, say, the benefactor (being virtuous) derives satisfaction from helping others, is well aware of that and (being a normal person) is inclined towards her own satisfaction (i.e., recognizes the fact that she will derive satisfaction from acting as she does as an incentive to act). So, it should be possible to establish that the benefactor acted solely on an other-regarding motive, even though a self-interested incentive to act was also available.

One could object that as a matter of fact, whenever we do something, there is a large number of psychological factors (such as desires, concerns, emotions, patterns of thought, etc.) that are at play. If cases where only one such factor is at play in the production of the action are possible at all, they are quite rare. So, the objector would continue, as a rule we act for compound-motives corresponding to this multitude of psychological factors. But that would be a mistake. From the fact that several psychological factors were at play in the production of an action it does not follow that the agent's motive must incorporate all these factor or considerations related to them. For instance, the fact that an old lady reminds me of my grandmother may be one of the factors at play when I decide to help her. That is perfectly compatible with the claim that in acting I aim solely at helping her, and, thus, that my only goal in acting is to help her. It follows, given the motive-goal link, that the motive for which I helped her is simply that she needed help (or something similar). That is, even though there were several psychological factors at play in the production of my action, my motive is not a compound-motive. The fact that there are always several psychological factors at play in the production of action is perfectly compatible, therefore, with the existence of multiple-incentives cases.

Now, if there are multiple-incentives cases, it follows that the motives for which we act are not determined by the incentives we happen to recognize. And this, I shall now argue, entails two things: (a) that the goals at which our actions aim are not passively determined by the belief-desire pairs we happen to have and (b) that we are not moved by belief-desire pairs

4. Multiple-incentives cases, Desires and the Goals of our Actions

If an agent is moved to perform a particular action by her desire for a reward she believes she will get by so acting, then her action aims at getting her that reward. For instance, if Pete is moved to volunteer at the soup kitchen by his desire to be admitted at the University combined with the belief that doing so would increase his chances of admission, then his action is directed at the goal of being admitted at the University. In the same way, if one is moved to return a lost dog to its owners by one's desire for money combined with the belief that by returning the dog one is likely to get a reward, then one's action aims at getting that reward. And the same seems to apply to cases in which the agent is moved by several desires or pro-attitudes: if one is moved to perform a particular action by a concern for others and by the desire to obtain a certain reward, then one's action aims both at relieving the suffering of others and at getting the reward in question. In general, the following entailment seems to hold:

DESIRE-GOAL LINK: If a pro-attitude towards G combined with the belief that action A is conducive to G is part of what moved agent P to perform action A , then A aims at goal G .

Now, according to the hydraulic model, we are moved by our desires or pro-attitudes, combined with appropriate means-end beliefs. And given the thesis of compositionality, belief-desire pairs that favor the same action combine in motivating us to perform that action. In combination with the desire-goal link, the hydraulic model entails that the goals at which a particular action aims are determined by the belief-desire pairs in the agent's motivational set that favor the action in question. If we accept this consequence, however, we cannot account for multiple-incentives cases.

Consider the volunteer work case. Mary recognizes two incentives to volunteer: the fact that it will contribute to relieving the suffering of people in need and the fact that it will increase her chances of being admitted at the University. Nevertheless, if this is a genuine multiple-incentives case, the correct motive ascription is (a) "she volunteered because doing so would contribute to relieving the suffering of people in need" and not (b) "she volunteered because doing so would contribute to relieving the suffering of people in need and would also increase her chances of being admitted at the University". But why is (b) incorrect? That requires an explanation and given that Mary wants to be admitted at the University and believes that volunteering contributes to that goal, the only available explanation is this: as a

matter of fact, her action does not aim at getting her to be admitted at the University. Given the motive-goal link, it follows that she did not volunteered because it would increase her chances of being admitted at the University, i.e., it follows that (b) is incorrect.¹⁷

The problem, of course, is that if the goals an action aims at are determined by the agent's belief-desire pairs that favor the action in question, then this explanation is not available. In multiple-incentives cases the agent recognizes more than one incentive to act as she does. That entails that she has more than one belief-desire pair that favors the action she performs. If the goals at which our actions aim are determined by those pairs, it follows that the action in a multiple-incentive case aims at more than one goal, each one corresponding to one belief-desire pair that favors the action. In particular, it follows that Mary's action does actually aim at getting her to be admitted at the University (in addition to aiming at relieving the suffering of people in need).

The existence of multiple-incentives cases leads, therefore, to two important conclusions. First, that the goals at which action A aims are not fixed by the pro-A belief-desire pairs the agent happens to have (or, which is the same, that the goals action A aims are not fixed by the incentives to perform A the agent happens to recognize). Second, given that the hydraulic model, in conjunction with the desire-goal link, entails that the goals at which our actions aim are determined by our belief-desire pairs, the existence of multiple-incentives cases shows the hydraulic model to be false – in particular, it shows that we are not moved by our belief-desire pairs conceived of as motivational forces.

I return to the former conclusion in section 8. In the next sections I explore some attempts to avoid the conclusion that we are not moved by our belief-desire pairs conceived of as motivational forces.

5. Davidson Causalism

The hydraulic model is composed of three theses. This set of theses is incompatible with the existence of multiple-incentives cases, at least as long as we accept the desire-goal link. But perhaps we could hold on to claim (a) that we are moved by our belief-desire pairs by rejecting either (b) the thesis of compositionality or (c) the thesis that when

17 If one holds the Foreground View the question is why, given that Mary has two belief-desire pairs that recommend the action of volunteering, only one of these pairs constitutes her motive. The answer has to be that her action does not aim at the goal corresponding to the belief-desire pair that does not constitute her motive.

faced with incompatible alternatives we perform the action we are more strongly motivated to perform. I doubt this will work.

On the one hand, rejecting (c) while holding on to (b) does not solve the problem: if our belief-desire pairs combine to produce stronger motivations than the ones they would produce on their own and we are moved by one such motivation, even if not by the stronger one, then, given the desire-goal link, the goals at which an action A aims are determined by the agent's pro-A pairs. On the other hand, the option of rejecting (b) while holding on to (a) and (c) is problematic. If our behavior is determined by the relative strength of our belief-desire pairs but they do not combine, we would always act on our stronger belief-desire pair even if it was opposed by several pairs that are stronger than it in combination. But that is implausible: it is possible for several lesser incentives to surpass a stronger incentive. For instance, if an agent is offered a job in another city, her stronger relevant incentive may be provided by the fact that she will get a significant raise if she accepts the job but she may be motivated to reject the offer on account of several lesser incentives (such as the fact that by rejecting the job she gets to stay close to her friends and family, she can keep her current job which she takes to be more stimulating, she will avoid the stress of a longer commute etc.).

It should be granted that all this counterexample shows is that an agent's belief-desire pairs can combine to produce a stronger motivation, not that they always do. One could then propose that we account for multiple-incentive cases by claiming that in these cases the agent's pro-A pairs simply do not combine in motivating her to perform A. That cannot mean, however, that the agent is left with several isolated motivations to perform A: we can make sense of the claim that providing a further incentive makes the agent more motivated to perform an action A, but not of the claim that providing a further incentive produces in the agent a further motivation to perform action A but does not make her any more motivated to perform A. The suggestion must be, then, that in multiple-incentives cases some of the agent's belief-desire pairs simply do not motivate the agent – they are there, the agent recognizes the corresponding incentive, but they remain inoperative.

According to an interpretation of Davidson's causalism, that is exactly his view.¹⁸ Davidson's argument, the interpretation goes, consists in pointing to multiple-incentives cases

18 At his point I am assuming that Davidson is committed to hydraulic model. That is probably an incorrect interpretation of Davidson's view. The point of this section is simply to show that Davidson's solution to the problem of accounting for the possibility of multiple-incentives cases fails *as long as we are committed to the hydraulic model*. I return to Davidson's view, considering alternative interpretations, in chapters 2 and 4.

and claiming that causalism can account for them whereas non-causalism cannot: in these cases, only one of the available belief-desire pairs causes the action.¹⁹ If only one of several pro-A pairs caused action A, then it alone moved the agent. It must be the case that other pro-A belief-desire pairs the agent happens to have remained inoperative in this case. In that way the claim that we are moved by belief-desire pairs can be reconciled with the claim that multiple-incentives cases are possible, even if the desire-goal link is accepted.

One difficulty this view face is to explain why some of the agent's belief-desire pairs fail to cause the action: why they failed to play any part in bringing about the action if they were available and live, did not conflict with the belief-desire pair that actually moved the agent and, in many cases, would move the agent in the absence of that pair? Given that we only have a genuine multiple-incentives case if the agent recognizes multiple incentives, the option of claiming that the agent failed to put together the inoperative belief-desire pairs is unavailable. And simply claiming that some of the agent's belief-desire pairs remained inoperative because they did not cause the action is uninformative. That is like saying that some of the engines in a ship remained inoperative because they did not cause the ship to move – clearly the explanation has to go the other way around. Objections along these lines have been put forward in the literature and I do not intend to press them further.²⁰

Rather, I will argue that this interpretation of Davidson's view has unacceptable consequences concerning multiple-incentives cases, namely, that in these cases the agent's motivation is always out of line with the correct assessment (from her own perspective) of how desirable the action in question is.

According to Davidson, our belief-desire pairs constitute the perspective from which we assess the desirability of possible actions. According to him, we should think of these attitudes as providing the premises for an argument whose conclusion is that the action is *prima facie* desirable (Davidson, 1980, p.77). Consider, for instance, the case of someone who adds sage to the stew with the intention of improving its taste: the belief "adding sage to the stew will improve its taste" provides a corresponding premise and the desire to improve

19 See, for instance, Dancy (2000, p.161-2), Wallace (2006a, p.61) and Dickenson (2007, p.3-4).

20 See, for instance, Dancy (2000, p.161-163) and Dickenson (2007, p.13-4). Both hold that Davidson cannot provide an informative answer to the question "why did primary reason R1 caused the action and primary reason R2 did not, provided that both were available?". Dickenson suggests that Davidson could give content to that claim by introducing the notion of motivational strength and claiming that which among the agent's pro-A pairs caused her action A is determined by the relative strength of the pairs (2007, p.15-6). I discuss this suggestion in the next section.

the taste of the stew provides the evaluative premise that “it is *prima facie* desirable to improve the taste of the stew” (Davidson, 1980, p.78). From that, we can conclude that it is *prima facie* desirable to add sage to the stew. Presumably, when comparing incompatible courses of action, we can weigh the *prima facie* judgments favoring each one to determine which is more desirable.

Davidson also accepts thesis (c). He claims that the following principle seems self-evident to him: “if an agent wants to do x more than he wants to do y and he believes himself free to do either x or y, then he will intentionally do x if he does either x or y intentionally” (Davidson, 1980, p.23). In normal, non-akratic cases, therefore, the agent’s motivation to perform an action *will* track the agent’s assessment of how desirable the action is: if actions *A* and *B* are incompatible, an agent takes *A* to be more desirable than *B* and she is not incontinent, then she will perform *A* and not *B* (if she performs either), and that means, given (c), that she is more motivated to perform *A* than *B*.

Given thesis (c) and the claim that in multiple-incentives cases the agent is moved to perform action *A* by only one of her pro-*A* pairs it follows that in these cases the belief-desire pair corresponding to the agent’s motive must be strong enough to move the agent on its own, while the belief-desire pairs corresponding to the other incentives remain inoperative. Consider then Mary’s case. Suppose her concern for the well-being of others provides her with a motivation that is strong enough to get her to volunteer. Given that hers is a multiple-incentives case, the belief-desire pair composed of her desire to be admitted at the University and her belief that volunteering would increase her chances of achieving that goal remains inoperative even though she is well aware of the corresponding incentive. If nothing new comes up she will volunteer and her action will aim solely at relieving the suffering of people in need. But she finds out that her volleyball practice was moved to the same time at which she would volunteer at the soup kitchen (suppose that was the only time she could do it). She is quite passionate about volleyball and attending the practice is also a way for her to get in touch with some friends and to stay fit – given these incentives she acquires a motivation to attend the practice that is slightly stronger than her motivation to volunteer. Now let us suppose that were Mary to weigh the incentives favoring the option of volunteering and the option of attending the practice, she would conclude that the option of volunteering is more desirable. Two things could happen at this point: Mary could simply be moved by the strongest operative motivation and attend the volleyball practice or she could be prompted to

reflect about her options and weigh the available incentives. If the former, she would be led to act in a way that is sub-optimal from the point of view of the incentives she recognizes. If the latter, she would come to declare the option of volunteering more desirable than the option of attending volleyball practice. And then two things could happen: either her motivation to volunteer would fall in line with her assessment of how desirable the action is or it will not. If the latter, then she will act akratically against her best judgment. If the former, she becomes more motivated to volunteer than she was before and that means that her previously inoperative belief-desire pair becomes operative. This is by itself an odd result: ordinarily, an agent does not become more motivated to do A simply because she found out that in order to do it she must give up something she wants almost as much as she wants to do A. That, however, seems to be what happened to Mary, for she did not come to recognize new incentives she was previously unaware of, she did not come to a fuller appreciation of the good she could do or of the benefits she could accrue nor has she come to a deeper resolve to volunteer (as before, she simply takes that to the most desirable option available). More importantly, however, since she did not become aware of any new incentives, it follows that her previous, lesser motivation to volunteer, was out of line with the correct assessment, from her own idiosyncratic perspective, of the action's desirability. And that means that the possibility of her acting with a view solely to an altruistic goal rested upon a flawed or incomplete assessment of the desirability of the option of volunteering or upon a quasi-akratic misalignment between her motivation and her judgment of desirability. That becomes clearer if we suppose that the volleyball practice is again moved so that now Mary can both volunteer and attend the practice. Now she has already consciously entertained the judgment that volunteering is more desirable than attending the practice, but she will be able to volunteer with a view solely to an altruistic goal only by rendering her self-interested belief-desire pair inoperative and thus dialing down her motivation, so that as a matter of fact it becomes weaker than her motivation to attend volleyball practice – and then her motivation will be out of line with her judgment.

Surely, however, (i) the possibility of multiple-incentives cases does not rest upon flawed or incomplete assessments of desirability. One can marry out of love only even if one is fully aware of other incentives to marry and correctly weighs these incentives in assessing

the desirability of marrying.²¹ And the benevolent person can perform charitable actions for their own sake even if she is aware of the personal advantages that may result from so acting and correctly assess the balance of incentives – a joyful realization that this balance favors the option of helping others is by no means incompatible with true virtue. And (ii) it is hardly the case that the possibility of multiple-incentives cases rests upon a misalignment between motivation and judgment of desirability. Given thesis (c), Davidson must assume that in normal, non-akratic cases, motivation tracks the agent’s judgments of desirability and there is no reason to suppose that multiple-incentives cases deviate from this rule or that in these cases agents present some kind of quasi-akratic flaw in motivation. I conclude, therefore, that Davidson’s causalism cannot account for multiple-incentives cases in a satisfactory way, at least as long as we hold on to the view that we are moved by our belief-desire pairs.²²

6. Motivational Strength

Another option open to supporters of the hydraulic model is to reject the desire-goal link. If one takes this path one must provide an alternative explanation of how the goals an action aims at are determined. In this and the next section I consider and reject some proposals along these lines.

One could suggest that the goals our action aims at are a function not only of what belief-desire pairs move us to act but of how strong they are. In particular, one may think that the goals our actions aim at are not fixed by the whole set of belief-desire pairs that moved us but rather by the belief-desire pairs that have some kind of preponderance in that set. A way to cash out that suggestion (and, I believe, the most promising one) is to hold that:

SUFFICIENCY CRITERION (SC): An agent performs an action A with a view only to goal G even if she is moved by several belief-desire pairs if, and only if, (i) the agent performs A, (ii) the agent desires G and believes A to be conducive to G and (iii) this belief-desire pair is strong enough to motivate the agent to perform A in any counterfactual situation in which the circumstances of action are the same but the agent does not have any of the

21 Granted, if one’s decision to marry is based on a weighing of reasons for and against marrying, then one is probably not marrying out of love. But surely marrying out of love is not rendered impossible by a clear-eyed assessment of the balance of reasons. Sincere love is not the privilege of the ignorant or the fool.

22 It should be noted that this is not a refutation of the thesis according to which when a motive ascription of the form “agent S did A because M” is correct the belief-desire pair whose possession is a background condition for the correction of this ascription causes the action A. What I hope to have shown is that if multiple-incentives cases are possible, then this thesis does not fit well with the idea that our belief-desire pairs move us as motivational forces. In chapter 4, I discuss the question of what is left of the idea that belief-desire pairs cause our actions when we reject the notion of motivational forces.

other belief-desire pairs that favor the performance of action A in her actual circumstances.

This proposal clearly involves the rejection of the desire-goal link for it entails that one can be moved by several belief-desire pairs without it being the case that one's action aim at corresponding goals.

The appeal of the proposal is clear. Consider the volunteer work case. SC entails that in this case we are authorized to claim that "Mary volunteered because doing so would contribute to relieving the suffering of people in need" only if it is true that Mary would have acted in the same manner even if she did not want to be admitted at the University or did not believe that volunteering would increase her chances of admission. And that seems reasonable enough. Nevertheless, SC faces serious problems.

First, it cannot account for some multiple-incentives cases, namely, those in which more than one pro-A belief-desire pair is strong enough to motivate the agent to perform action A. For instance, it may be true that Mary would have acted the way she did even if she did not want to be admitted at the University and thus that, according to SC, her goal in volunteering was to contribute to relieving the suffering of these in need. But it may be also true that she would have acted in the same way if she did not care for the well-being of others – and, according to SC, that would entail that she volunteered with a view to being admitted at the University. In this case, SC would either lead to a contradiction or, in a charitable reading, entail that Mary's action aimed at a compound-goal. We could only claim that she acted with a view to the altruistic goal alone if the self-interested incentive prompting her to volunteer was not strong enough to motivate her to volunteer on its own. And that would be the case only if she had a stronger set of belief-desire pairs prompting her to perform an action incompatible with the option of volunteering. That is to say that whether or not Mary acts with a view to the altruistic goal alone depends on how strong are the belief-desire pairs that prompt her to perform alternative actions. This introduces a problem of irrelevance. Suppose that Mary wants to dedicate more hours to practicing volleyball (a sport about which she is passionate) and that this option is incompatible with the option of volunteering. According to the suggestion under consideration, assuming that the intensity of her concern for the well-being of others and her desire to be admitted at the University remains constant, she volunteers with a view to a compound-goal if the belief-desire pair prompting her to dedicate more hours to volleyball is weaker than the belief-desire pair composed of the desire to be

admitted at the University and the belief that volunteering increases her chances of admission. If the former pair is stronger, however, then her action of volunteering aims only at the altruistic goal. This means that her action could go from aiming at a compound-goal to aiming at a pure altruistic goal simply because her desire to dedicate more hours to volleyball became stronger. But that cannot be right: why should the fact that Mary became more passionate about volleyball make it the case that her act of volunteering does not aim at securing a personal advantage for her, if she still recognizes the same incentives and is moved by the same desires?

In order to avoid this problem one could restrict SC to certain goals. For instance, one could hold that we ascribe altruistic goals to an action when the conditions specified by SC are obtained because we are willing to ascribe a pure altruistic motive to altruistic actions we deem praiseworthy and, as a matter of fact, we deem altruistic actions praiseworthy when the agent was moved by an altruistic desire that would have moved her even if she had not recognized the self-interested incentives she did recognize. And, one would continue, for this reason SC holds when goal *G* is an altruistic goal but not when it is a self-interested goal. However, even if we restrict SC to altruistic goals in this manner, we should reject it for it leads to arbitrary goal attributions.

Consider Mary again. She recognizes two incentives to volunteer at the soup kitchen, one altruistic, the other self-interested, and has, therefore, the corresponding belief-desire pairs. She actually volunteers and, let us suppose, she acts in the same manner in the counterfactual circumstance in which she does not recognize the self-interested incentive. According to the restricted version of SC, therefore, she volunteers in order to contribute to relieving the suffering of people in need and her action is commendable. Now consider Megan. She recognizes the same incentives as Mary and, let us assume, the belief-desire pairs underlying these incentives are exactly as strong as Mary's belief-desire pairs. Megan also volunteers at the soup kitchen. Nevertheless, Megan does not act in the same way in the counterfactual situation in which she does not recognize the self-interested incentive to volunteer. Let us suppose that in this counterfactual situation Megan decides to spend her few free hours practicing volleyball instead of volunteering. According to the restricted version of SC, therefore, it is not the case that Megan volunteers in order only to contribute to relieving the suffering of people in need (at the very least her action aims at a compound-goal that includes a self-interested goal) and, therefore, her action is not as commendable as Mary's.

But does that counterfactual difference provide any ground to ascribe different goals to Mary's and Megan's action? What is the actual difference between them that grounds the counterfactual difference? If we are committed to the hydraulic model, it must be the case that Mary's altruistic belief-desire pair was stronger than the set of belief-desire pairs favoring the option of practicing volleyball, while for Megan the contrary is true. Given the stipulation that the altruistic desire was equally strong in Mary and Megan (and that their beliefs are the same), it must be the case that the set of belief-desire pairs favoring the option of practicing volleyball is weaker in Mary than in Megan. The fact is, let us suppose, that Megan is simply more passionate about volleyball than Mary. That means that in deciding to volunteer Megan had to overcome a stronger opposing desire. Other than that, the belief-desire pairs that moved her were exactly the same that moved Mary and exactly as strong. Why should we say then that they acted with a view to different goals? The natural thing to say is that they acted with a view to the same goals, and for the same motives, although that Megan had to overcome a slightly stronger temptation to act otherwise.²³ Our interlocutor could insist that we ascribe a pure altruistic goal only to Mary because we judge her action more praiseworthy than Megan's on account of the fact that Mary would have acted in the same way even without any self-interested incentive. But, given that the only actual difference between Mary and Megan that a supporter of the hydraulic model can identify is that Megan is more passionate about volleyball, that is the only ground on which to claim that her action is less praiseworthy than Mary's. And that is simply absurd: surely Megan's action would not become any more praiseworthy on account of her losing her enthusiasm for volleyball (assuming her concern for others remains the same) nor is it the case that morality requires her to become less passionate about the sport.

We should, therefore, reject SC. The supporter of the hydraulic model could try to provide another alternative to the desire-goal link. He could simply claim that we perform action A with a view only to G when the belief-desire pair composed of the desire for G and the belief that A is conducive to G is the strongest belief-desire pair among the agent's pro-A belief-desire pairs. This proposal also has implausible consequences. Suppose that Mary cares more about the well-being of others than she does about being admitted at the University.

23 To be clear, my point is not that counterfactual claims about how an agent would have acted if the incentives available to her were different are not relevant in justifying a motive ascription. My point is rather that we cannot reconcile the idea that we are directly moved by our belief-desire pairs with the idea that is possible to individuate the goal with a view to which an agent acts in a multiple-incentives case by appealing to such counterfactual claims.

According to this suggestion, her action of volunteering aims only at contributing to alleviate the suffering of people in need. Now suppose that we increase the strength of her desire to be admitted at the University so that she cares about being admitted at the University as much as she cares for the well-being of others. It would follow that her action aimed at a compound goal – her action is not purely altruistic but her motive is still partially altruistic. Now suppose that we increase the strength of her desire to be admitted at the University a lit bit more, so that the self-interested incentive to volunteer becomes stronger than the other-regarding one. It would follow that Mary’s action aimed solely at increasing her chances of being admitted at the University. Given the motive-goal link, it follows that her motive now is completely selfish. All traces of altruism in her action were obliterated. But that cannot be right: she is still moved by a concern for the well-being of others and for all we know it could be the case, given the framework of the hydraulic model, that she would not have volunteered where it not for that other-regarding attitude (for, perhaps, she had a self-interested incentive to perform an incompatible action that is stronger than the self-interested incentive to volunteer but not stronger than the combination of her two incentives to volunteer). Surely the intensity of our desires may be a relevant factor in determining the goals at which our actions aim, but it does not determine those goals in such a straightforward way.

The project of providing a criterion that specifies the goal with a view to which an agent acted in terms of the relative strength of her belief-desire pairs is not, therefore, very promising. If supporters of the hydraulic model hope to provide an alternative to the desire-goal link they should look elsewhere.

7. Making room for the agent

In this section I will explore other alternatives to the desire-goal link that are suggested by the views of some supporters of the hydraulic model and argue that they fail to account for multiple-incentives cases as long as we hold on to this model.

Hierarchical Complexity. Harry Frankfurt seems to commit to the hydraulic model. He identifies an agent’s “will” with “the desire (or desires) by which he is motivated in some action he performs” and he claims that this notion of the will is not coextensive with the notion of what the agent intends to do, for an agent may intend to do X but do something else because “his desire to do X proves to be weaker or less effective than some conflicting desire”

(Frankfurt, 1971, p.8). Our actions, then, are determined by the tug of war between our desires. He adds complexity to that view by introducing the notion of second-order desires, i.e., desires about our desires. Particularly, we have desires concerning our will – we may want a particular desire to be the one that effectively motivated us (Frankfurt, 1971, p.10). This kind of second-order desires are what Frankfurt calls “second-order volitions”. By coming to have a second-order volition the agent identifies himself with one of her conflicting desires (namely, the one she wants to constitute her will) and withdraws from others (Frankfurt, 1971, p.13). Frankfurt also describes this operation as one by means of which the agent endorses one of the conflicting desires as a legitimate candidate for satisfaction and rejects the other (Frankfurt, 1988, p.170).

Frankfurt does not consider multiple-incentives cases and the way he presents his use of “will” suggests he believes there are no such cases – when one is moved by more than one desire, one is moved by their combination. The notion of second-order volition, however, suggests a criterion to specify the goal at which one’s action is aimed in a multiple-incentives case that is not based on the relative strength of the incentives:

VOLITION-GOAL LINK: An agent performs an action A with a view only to goal G even if she is moved by several belief-desire pairs if, and only if, (i) the agent performs A, (ii) she has an incentive I (associated with goal G) to perform action A and (iii) she endorses that incentive by means of a second order volition.

It would then be possible to act with a view to a particular non-compound goal in a multiple-incentives case by identifying with only one of the available incentives. But that is not enough to account for multiple-incentives cases. Mary, for instance, may identify both with her altruistic concern for others and with her desire to be admitted at the University – she sees both attitudes as her own, neither is perceived by her as an alien force prompting her to act and she takes both to be legitimate candidates for satisfaction. The volition-goal link would then either lead to a contradiction or entail that she aimed at a compound-goal (and the same would be true of most multiple-incentives cases).

This problem could be avoided if we assumed that in multiple-incentives cases the agent has a more complex second-order volition – she desires to be moved only by a particular incentive. But this leads to another problem. The trouble is that, as explicitly acknowledged by Frankfurt (1988, p.172), second-order volitions may fail to have any direct

impact on the dispute between our first order desires. Unfortunately, second-order volitions may be left unsatisfied. And that will always be the case with the volition above, at least as long as we accept the hydraulic model, according to which in performing action A an agent is moved by her whole set of pro-A belief-desire pairs. It is hard to see why the fact that the agent has a systematically unsatisfied second-order desire should make any difference concerning the goals at which her actions are aimed.

Neil Sinhababu explicitly applies Frankfurt's notion of second order volition to multiple-incentives cases and he argues that a second-order volition to be moved exclusively by a particular desire could effectively prevent other desires from moving the agent (Sinhababu, 2013, p.687-8). According to him, such a volition, were it strong enough, could render all pro-A desires except the one it favors inoperative, in which case the agent would perform action A only if the preferred desire is strong enough to move the agent on its own. But that cannot be right. If a second-order volition could render inoperative a belief-desire pair that favors the action the agent hopes to perform, then it could also render inoperative belief-desire pairs that oppose that action, and that clearly is not the case. Think of someone who is addicted to a drug. This person has a very strong desire to have the drug. Suppose, however, that this person also wants very much not to be moved by that desire – the addiction has ruined her life and the thing she wants more in the world is to overcome it. Her second-order volition not to be moved by the desire for the drug is as strong as can be, but, unfortunately, we all know that is not enough to render that desire inoperative. It is not even enough to diminish its strength. And the same goes for non-compulsive desires: the dieter's desire to have one more desert is not silenced by an opposing second-order volition nor is the philanderer's desire to cheat on his wife. As Sinhababu himself acknowledges, it is "a sad fact of life that desiring to desire ϕ doesn't directly increase one's desire for ϕ " (Sinhababu, 2013, p.688). It is an equally sad fact of life that the desiring not to desire ϕ does not directly weakens the desire for ϕ .

Second-order volitions do not provide the key to account for multiple-incentives cases.

A Rational Incentive. According to Velleman, an agent can participate in her own action only if she adds something to the normal motivational influence of her desires and beliefs

(Velleman, 1992, p.465). Velleman intends to account for that possibility by ascribing to every agent capable of practical thought a desire to act according to reasons, whatever those happen to be (Velleman, 1992, p.479).²⁴ An agent with this desire can contribute to the determination of her own behavior by reflecting on and coming to a conclusion about what she has reason to do and, thus, tapping into the motivational force of that desire.²⁵ The agent can then reinforce a motive by throwing her weight behind it but “what is thrown behind those motives, in fact, is the additional motivating force of the desire to act in accordance with reasons” (Velleman, 1992, p.479).

It should be clear that Velleman is committed to a hydraulic model. We are always moved by the strongest combination of motives (Velleman, 1992, p.480; see also Velleman, 1989, p.35). The desire to act according to reasons is just one among others – it only happens to be oriented by our normative conclusions about what we have reason to do. By itself, the inclusion of this desire in the agent's motivational set does not make it any easier to specify the agent's goal in a multiple-incentives case.²⁶ It does suggest a criterion however:

RATIONAL MOTIVE-GOAL LINK: An agent performs an action A with a view only to goal G even if she is moved by several belief-desire pairs if, and only if, (i) the agent performs A and (ii) she takes the fact that A is conducive to G to provide her with a sufficient reason to do A.²⁷

This is not enough to account for multiple-incentives cases. Mary, for instance, may take both the fact that her action of volunteering promotes the end of relieving the suffering of people in need and the fact that it is conducive to her goal of being admitted at the University as sufficient reasons to volunteer. The rational motive-goal link would then either lead to a contradiction or entail that she aimed at a compound-goal (and the same would be true of many multiple-incentives cases).

24 This idea is formulated in slightly different terms in other texts. For instance, in his *Practical Reflection* (1989) he ascribes to every agent the desire to do what makes sense for them and in his *The Possibility of Practical Reason* he ascribes to every agent an inclination “to do what one accepts that one will do” (Velleman, 1996, p.722).

25 For a very similar proposal see Broome (1997, p.142).

26 Gary Watson (1975) defends a position similar to Velleman's. According to Watson we have two sources of motivation: the desires we happen to have and our judgments, guided by our values, about what is the thing to do in a particular situation (1975, p.215). These two sources of motivation may be aligned or they can conflict (in which case one may be led, by one's desires, to act in disagreement with one's practical judgment). Watson's proposal will face exactly the same problem as Velleman's.

27 In the next chapter, I discuss in detail the idea that the goals with a view to which we act are determined by our normative judgments. Here my goal is simply to show that the rational motive-goal link is not an acceptable proposal as long as we hold on to the hydraulic model.

This problem could be avoided if we assumed that multiple-incentives cases are only possible when the agent takes only the fact that her action is conducive to one particular goal as a sufficient reason to act. But insisting on that is insisting that Mary could act with a view solely to relieving the suffering of people in need only if there was some fact testifying against the option of volunteering, so that the fact that it would increase her chances of admission at the University would not provide a sufficient reason to perform A. This will lead to the same problem of irrelevance we discussed when assessing SC. Whether or not Mary acted with a view only to an altruistic goal will depend on how strong are her reasons to perform an action incompatible with the option of volunteering, say, on how strong are her reasons to dedicate some extra hours to practicing volleyball. If the latter reasons are stronger than the reason provided by the fact that volunteering increases her chances of being admitted at the University, then she volunteers with a view to the altruistic goal alone. If her reasons to dedicate more time to volleyball is weaker than the latter reason, then she acts with a view to a compound-goal. But it is not clear at all why that should be relevant in determining the goal at which Mary's action aimed, if we assume that in both scenarios her assessment of her reasons to volunteer is the same, she recognizes the same incentives and is moved by the same belief-desire pairs.

Finally, one could claim that the rational motive-goal link holds with respect to certain goals. One could, as before, hold that we ascribe pure altruistic goals to actions when the agent recognizes the fact that the action is conducive to an altruistic goal as a sufficient reason to act (regardless of what other reasons she acknowledges) because we are willing to ascribe a pure altruistic motive to altruistic actions we deem praiseworthy and, as a matter of fact, we deem altruistic actions praiseworthy when the agent recognizes the fact that it is conducive to an altruistic goal as a sufficient reason to act. This is not an implausible suggestion, but as long as we hold on to the hydraulic model it has troubling consequences. Consider Pete. He really does not care for other people – let us suppose he is a bitter, cold person. On the other hand, he is very dutiful. According to Velleman, this means he has a very strong desire to act according to reasons *de dicto*. He is also a religious person who believes that God has made it so that certain facts provide us with reasons to act. In particular, he believes, on account of his reading of the bible, that God has made it the case that the fact that a particular action will ease the suffering of a person provides a reason to perform it. Finally, like Mary, he desires to be admitted at the University and believes that volunteering increases

his chances of admission. Pete, therefore, recognizes two incentives to volunteer: the fact that doing so increases his chances of being admitted at the University and the fact that doing so is a way of acting according to reason. He is moved to act by the corresponding belief-desire pairs. Nevertheless, according to the suggestion under consideration, Pete acts with a view solely to an altruistic goal. Given the motive-goal link, it follows that it is incorrect to claim that “Pete volunteered because doing so would increase his chances of being admitted at the University” or that “Pete volunteered because doing so was a way of acting according to reason” or a combination of both, for his action does not aim at the corresponding goals. But one of these motive attributions should be the correct one: these are the only incentives he recognized and these are the considerations that actually motivated him. Claiming otherwise, while holding on to the hydraulic model, is to accept the possibility of a radical disconnection between one’s motivational states and one’s motive.

Deliberation beyond Means-End Reasoning. Bernard Williams, in his seminal paper “Internal and External Reasons”, claims that our motivational set is highly plastic. Not only can we create new derivative desires by drawing means-end relations and suppress desires by showing that they rest on false beliefs, we can also, for instance, lose or acquire desires by exercising our imagination to get a more concrete sense of what would be involved in satisfying it (Williams, 1981, p.104-5). Deliberation, Williams claims, while being controlled by the agent's motivational set, can change it dramatically – it may add and exclude non-derivative elements from it, and, since it can do that, there should be no difficulty in admitting that it can change the relative strength of these elements.

Smith claims, in a very similar fashion, that the states that can explain our actions as goal-directed behavior are belief-desire pairs²⁸ and that practical reflection can produce as

28 In his *The Moral Problem*, Smith denies a hydraulic image of our motivational psychology (1994, p.101-2). He claims that this image is committed to the idea that belief-desire pairs cause our actions and that a Humean theory of motivation is not committed to causalism. Nevertheless, he defends the thesis that “R at t constitutes a motivating reason of agent A to Φ iff there is some ψ such that R at t consists of an appropriately related desire of A to ψ and a belief that were she to Φ she would ψ ” (Smith, 1994, p.92). That is, every pro- Φ pair of an agent provides her with a motivating reason to Φ . And a motivating reason explains an action by making it “intelligible in terms of the pursuit of a goal” (Smith, 1994, p.104). It seems, then, that according to Smith, the goals at which a particular action Φ aims are directly determined by the agent’s pro- Φ pairs – exactly what must be denied if we are to account for multiple-incentives cases. The fact that he conceives of desires as sets of dispositions to act in a certain way in certain circumstances (1994, p.113-5) is of no help here. This claim entails only that when we have more than one desire prompting us to perform a particular action, we have sets of dispositions that overlap at a particular point. That is of no help in specifying the goal at which the action aims. Furthermore, in a more recent paper, he seems to revert to a

well as suppress non-derived desires. According to Smith, what we have normative reason to do is, roughly, what we would desire that we do if we were completely rational (Smith, 1994, p.150) or ideal agents (Smith, 2013, 19). To believe that we have a reason to do *A* is to believe that we would desire to do *A* if we were deprived of cognitive limitations and rational failings and had a perfectly unified set of psychological states. Smith thinks that coming to have that belief can create or suppress desires because being rational involves a disposition or capacity to coherence; given that the set of attitudes that includes the belief that one has reason to do *A* and the desire to do *A* is more coherent than the set that includes that same belief but not the correspondent desire, agents, to the extent that they are rational, will display a tendency to transition from the latter set to the former (see Smith, 2003, p.32-35).

If that its correct, then agents have an active role to play in determining the configuration of their motivational sets and, through it, their actions. That, however, is not enough to account for the possibility of multiple-incentives cases. If we are moved by belief-desire pairs, the fact that our motivational set is highly plastic does not contribute to the explanation of how one can be moved to perform action *A* by a desire for *G* and yet not perform *A* in order to bring about *G*.²⁹ And this teach us an important lesson: in order to account for multiple-incentives cases it is not enough to assume that agents can actively determine their own behavior; rather we have to assume that agents can actively determine the goals at which their actions aim (instead of passively letting those goals be determined by the belief-desire pairs they happen to have).³⁰

causal conception of belief-desire explanations of actions. He continues to conceive of desires as dispositions to be moved in a certain way (2012, p.393), but now holds that belief-desire pairs cause actions (2012, p.387). He also holds that we are moved by the *stronger* disposition at play (2012, p.395). He is committed, therefore, to the idea that our behavior is determined by the power struggle between our desires. That is the essence of the hydraulic model and it brings with it all the problems we have been discussing.

29 Both William's and Smith's account of rational deliberation presuppose that an agent deliberation is controlled by her motivational set, in such a manner that it is possible for two agents that deliberate properly to come to different conclusions about their reasons only if they started with different motivational sets. This, however, is not the feature that makes these accounts unable to account for multiple-incentives cases. Even if our practical reasoning is capable of creating and suppressing non-derivative pro-attitudes without being controlled by other elements in the motivational set, the problem persists.

30 It is not immediately clear what deliberation, as Williams conceives of it, can accomplish. One could suggest that it could go so far as to *silence* certain considerations, preventing certain belief-desires from moving us. However, that would not help us account for multiple-incentives cases. As long as we hold on to the idea that we are moved by motivational forces, the idea of silencing would have to be understood as the temporary suppression of a motivational force. The resulting position would be identical to the position considered and rejected in section 5. Surely, the notion of silencing or disregarding a consideration that could be taken as reason for acting is relevant to our motivation. Its relevance, however, can only be properly appreciated when we reject the idea that we are moved by motivational forces. Even then, it is not clear that it can, by itself, account for the possibility of multiple-incentives cases – as I discuss in section 4 of the next chapter.

8. Why do we need the notion of Will?

In a multiple-incentives case, the agent recognizes more than one incentive to act but the correct motive ascription mentions only one of these incentives. The trouble is to explain why the correct motive ascription does not incorporate the other incentives. Given that the agent recognizes these incentives and, therefore, has the corresponding belief-desire pairs, the only available explanation seems to be this: as a matter of fact, her action did not aim at the goals associated with these incentives. For instance, given that Mary recognizes the fact that volunteering will increase her chances of being admitted at the University as an incentive to volunteer, the only way to explain why a motive ascription that incorporates the fact that volunteering will increase her chances of admission is incorrect in her case is to claim that as a matter of fact her action did not aim at the goal of being admitted at the University. That answer is unavailable as long as we assume that the goals at which our actions aim correspond to the goals associated with the incentives to act we recognize. The alternative proposals about how our belief-desires pairs determine our actions' goals explored in the previous sections were refused because they lead to arbitrary motive attributions. In order to account for multiple-incentives cases, therefore, we have to reject the idea that the goal at which an action A aims is passively determined by the pro-A pairs the agent happens to have. Rather, I will now suggest, we should conceive of agents as endowed with the capacity to actively determine the goals at which their actions are directed and, consequently, to actively determine the motive for which they act.³¹

I refer to such capacity as the “will”. An agent’s will determines the agent’s goal in performing a particular action by forming or acquiring intentions. The content of these intentions can be expressed as “I intend to do A, in circumstances C, in order to G”. The G-slot specifies the goal action A aims at when the intention is executed. To say that an agent has the capacity to actively determine the goal her action aims at is simply to say that the content of G-slot is not determined by her pro-A belief-desire pairs but rather by an exercise of the will. In particular, in multiple-incentives cases, an agent may form the intention of performing action A in order only to G even though she recognizes a further incentive to perform A, associated with goal E. Mary, for instance, forms the intention expressed by “I intend to do A in order to contribute to relieving the suffering of people in need” and not an intention that

31 And not only which action they performed: a position such as William’s accounts for the capacity to actively determine one’s action but not the capacity to actively determine one’s motive.

incorporates the goal of increasing her chances of being admitted at the University, even though she recognizes both incentives and has the corresponding belief-desire pairs. According to this view, an incentive becomes an agent's motive to act only to the extent the goal associated with it is incorporated into the agent's intention.³²

This view, of course, is incompatible with any view that reduces intentions to predominant desires or some combination of predominant desires with other desires or beliefs. There are two reasons for that. First, if we identify intentions with non-overridden desires to perform the intended action then intentions do not have the right kind of content to specify the goal at which the action is directed – they specify only the intended action not what it aims at. Ridge, for instance, holds such a view. He reduces the intention to do *A* to a combination of a non-overridden desire to do *A* and a desire not to deliberate further about whether to do *A* (Ridge, 1998, p.163). It becomes clear that one cannot account for multiple-incentives cases if one conceives of intentions in this manner once we take into consideration that agents that perform exactly similar actions in exactly similar circumstances can both have non-overridden desires to perform those actions and a desire not to deliberate further and yet act with a view to different goals and, therefore, for different motives. Mary and Pete both have a non-overridden desire to volunteer but Mary aims at an altruistic goal, despite recognizing a self-interested incentive, whereas Pete aims at a self-interested goal. In order to account for multiple-incentives cases we need to ascribe to agents *end-directed intentions*, i.e., intentions that specify the end the intended action aims at.

Audi provides a reductive account of end-directed intentions. According to him, an agent intends to bring about *G* by doing *A* (which is the same as having the intention expressed by “I intend to do *A* in order to *G*”) if, and only if, (i) she wants to bring about *G* by doing *A*, (ii) that want is not opposed by a stronger or equally strong want or set of wants and (iii) she believes that she will bring about that *G* by doing *A* (Audi, 1973, p.395). Although intentions thus conceived have the right kind of content to account for multiple-incentives cases, they cannot do so. If intentions are thus conceived, then the goals our actions aim at are determined by our motivational set, particularly by the power struggle between the elements in that set, and we learned from our previous discussion that we cannot account for multiple-

32 The thesis that incentives only become motives when they are incorporated into the agent's intention has a Kantian ring to it. Kant notoriously held that an incentive moves one to action only in so far as it is incorporated into one's maxim (RGV, 6: 24, 73). Herman claims that an incentive becomes a motive only when incorporated into the agent's maxim (Herman, 1993, p.11-12). See also Allison (2011, p.114-5).

incentives cases while we hold on to that supposition. Indeed, given Audi's characterization of intentions, all such cases are reduced to compound-motive cases. Consider once again Mary's case. She may well have a desire to increase her chances of being admitted at the University by volunteering, and it seems correct to claim that this want is not overpowered by any opposing want, that she believes that she will volunteer and that volunteering will increase her chances of being admitted at the University. According to Audi's account of end-directed intentions, therefore, it would be case that she intended to volunteer in order to increase her chances of being admitted at the University. Consequently, her motive would be a compound one: her motive for volunteering would be that it would promote the well-being of others and increase her chances of admission. We would be unable to specify one of these as her sole motive for acting. But that is exactly what is necessary in order to account for multiple-incentives cases. Audi could argue that although Mary wants to increase her chances of being admitted at the University and believes that volunteering is a way to do that, she does not want to increase her chances of admission by volunteering. But what could that mean? One way to understand that is as the claim that the agent does not want her action of volunteering to aim at increasing her chances of admission. Given the desire-goal link, that could only be the case if she was not moved to volunteer by her desire to increase her chances of being admitted at the University. To say that the agent does not want her action to aim at increasing her chances of admission would be, then, to say that she has a second order desire not to be moved to volunteer by her desire to increase her chances of admission. But, as we saw when discussing Frankfurt's account of second order volitions, introducing second order desires is not enough to account for multiple-incentives cases, at least as long as one holds on to the view that our belief-desire pairs produce motivational forces that move us to action. Another option is to understand the claim that Mary does not want to increase her chances of admission by volunteering as the claim that the belief-desire pair composed by the desire to increase her chances of admission and the belief that volunteering would do so does not motivate her to volunteer. That would have to mean that the belief-desire pair in question remained inoperative in this case. But we already rejected the claim that one can account for multiple-incentives cases in terms of inoperative belief-desire pairs in section 5.

The existence of multiple-incentives cases entails, therefore, that agents are capable of actively determining the goals at which their actions aim, which is to say that they are not determined by the belief-desire pairs they happen to have. To give content to this idea

we have to develop a conception of the will and the intentions it produces. That is the task of the next two chapters.

2. Willing, Weighing, Planning

1. Introduction

In order to account for multiple-incentives cases, we have to reject the idea that the goal with a view to which an agent acts is determined by the desires or pro-attitudes the agent happens to have at the time of the action. We must, rather, conceive of agents as endowed with the capacity to actively determine the goals their actions aim at and, therefore, to actively determine the motive for which they act. I refer to such capacity as the “will”. The will determines the agent’s goal in performing a particular action by forming intentions. The content of these intentions can be expressed as “I intend to do *A*, in circumstances *C*, in order to *G*”. The *G*-slot specifies the goal action *A* aims at when the intention is executed. To say that an agent has the capacity to *actively* determine the goal her action aims at is to say that the content of *G*-slot is not determined by her motivational set but rather by an exercise of the will. In particular, in multiple-incentives cases, an agent may form the intention of performing action *A* in order to *G*, *and only G*, even though she has a pro-attitude towards *E* and knows that performing action *A* would promote *E*. For instance, even though an agent has both a concern for the well-being of others and a desire for a good reputation and knows that performing a charitable action *A* is an effective means to promote a good reputation, she may form the intention of performing *A* in order to promote the well-being of the beneficiaries of her act and not in order to promote their well-being *and* her personal interest.

Of course, the terms “will” and “intention” are, so far, no more than place-holders for an actual explanation of how agents can determine the goals their actions aim at. Now we need to flesh out a conception of the will and the intentions it produces that can perform the theoretical role of accounting for multiple-incentives cases.

Given the claim that the content of an agent’s intentions is actively determined by the agent and not merely a function of the desires or pro-attitudes the agent happens to have, it may seem natural to conceive of the will as practical reason. After all, reasoning is something we do, not something that merely happens to us and one may think that if the goals our actions aim at are not determined by our desires or pro-attitudes they can only be determined by the deliverances of our reason. The content of this conception of the will depend, of course, on how we conceive of practical reason. In order for it to be a viable conception of the

will we should not conceive of practical reason as instrumental reason or as deliberative reason (in the manner in which, for instance, Bernard Williams conceives of it as a capacity to deliberately modify our motivational set). If practical reasoning can only change our motivational set (whether by producing new derivative pro-attitudes by discovering means-end relations or by altering that set in a more significant way) then the will cannot be identified with practical reason, for, thus conceived, it will be unable to actively determine the goals our actions aim at. A more promising option is to conceive of practical reason as the faculty to identify certain considerations as *pro tanto* reasons to act or to refrain from acting in a particular way and to reach, in light of these considerations, a verdict about what we have, all things considered, reason to do. According to this conception intentions are to be identified with (or at least, determined by) all-out normative judgments about what we have more reason to do (or some similar normative judgment). In what follows I will use “judicative reason” to refer to this capacity to weigh reasons in order to reach a normative judgment.

The idea that our behavior is determined by the activity of judicative reason is, of course, widespread and it can take different forms. Davidson, for instance, can be read as defending a version of this view, despite his insistence on the claim that for an agent to perform an action *A* with a view to *G* is for action *A* to be caused in an appropriate way by a desire for *G* and a belief that doing *A* is a way of promoting *G*. The fact is that, according to Davidson, when we act intentionally the belief-desire pair that causes our action must rationalize it (Davidson, 1980, p.77-8). For a belief-desire pair to rationalize an action, under a particular description, is for the performance of the action to be reasonable in light of the pair. And Davidson explains what it is for an action to be reasonable in light of a set of desires and beliefs in terms of the notion of reasoning. According to him, we should think of our desires and beliefs as providing the premises for an argument whose conclusion is that the action is desirable or possess some other positive quality (Davidson, 1980, p.77). Consider, for instance, the case of someone that adds sage to the stew with a view to improving its taste. This person does that because she desires to improve the taste of the stew and believes that adding sage to it will do just that. This desire and this belief rationalize that action because they provide the starting point for an argument to the conclusion that adding sage to the stew is desirable. The belief provides the premise that “adding sage to the stew will improve its taste”. The desire, Davidson proposes, provides the evaluative premise that “it is desirable to improve the taste of the stew” (Davidson, 1980, p.78). From these premises we can conclude

that it is desirable to add sage to the stew. This conclusion rationalizes the agent's action. Things are not so simple, however, because this reasoning is defeasible. If the agent in the example above also wants her son to enjoy the stew and knows that he dislikes the taste of sage, she may, on that account, conclude that adding sage to the stew is undesirable. What this shows is that all we can conclude from the premises provided by a belief-desire pair is that the action is desirable in a certain respect. That is, the conclusion of the argument must take the form of a *prima facie* evaluative judgment (and so must the premise corresponding to the desire). A *prima facie* judgment, however, cannot directly rationalize an action because the claim that the action is, all things considered, undesirable is compatible with such a judgment. To determine if an action is reasonable in light of set of desires and beliefs, we have to weigh the various *prima facie* judgments that can be derived from this set. If the *prima facie* judgment that are favorable to the action override the judgments that testify against it, then we can form the unconditional judgment expressed simply as "this action is desirable" (Davidson, 1980, p.87). Davidson equates intentions with these unconditional evaluative judgments (Davidson, 1980, p.88). Presumably, when we are faced with more than one way in which to act, we can weigh the *prima facie* judgments for and against each option so as to determine which is more desirable.

Other, more recent, versions of the view that our behavior is determined by judicative reason do not postulate such a close relation between our desires and the reasons we acknowledge. Scanlon, for instance, holds that a reason is simply a consideration that counts in favor of performing a particular action (Scanlon, 1998, p.17 and 2014, p.44). Dancy holds the same view (2000, p.1 and 2004, p.29). According to these philosophers we are capable of identifying some (putative) facts as counting in favor of performing an action and others as counting against performing. We can then weigh these considerations in order to determine what we have most reason to do.³³ Desires are not what underpins these reasons, rather they are to be understood in terms of reasons. According to Scanlon, for instance, to desire that p, in part, to have one's attention "directed insistently toward considerations that present themselves as counting in favor of p" (Scanlon, 1998, p.39). And according to Dancy, to desire that p is simply to motivated to bring about that p, where the motivation is produced by a belief or the consideration of a putative fact (Dancy, 2000, p.85).

33 Both Scanlon and Dancy hold that weighing is not the only thing we can do with reasons. We can perform more complicated operations: an agent can also disregard a particular reason or decide in light of a particular reason. I consider these kinds of operations of practical reason in section 4.

Yet another version of this view is put forward by Sergio Tenenbaum who holds that for an agent to desire something is for that something to *appear* good for the agent from a certain perspective (Tenenbaum, 2007, p.14). These appearances are the “building blocks” from which the agent tries to reach an all-out evaluative judgment to the effect that performing a particular action is good (2007, p.12). The kind of reasoning involved here is not simply the weighing of considerations for and against a particular action, but rather an effort of bringing together, in a single, unified conception, several disparate and possibly conflicting appearances of the good (Tenenbaum, 2007, p.53-4) - what may require, for instance, that one reject certain appearances as illusory. When it comes to actions, though, the conclusions this kind of reasoning can lead to are of the form “in these circumstances, the best thing to do is this”. One reaches such an “unconditional judgment” by moving from one’s set of appearances of good to judgments about what is valuable and then weighing these judgments so as to reach an “unconditional judgment” about what is the best thing to do in a particular situation. Intentions are then identified with these unconditional judgments.

All of these views differ in important ways, but they agree in holding that our behavior is determined by our capacity to move from certain considerations (considerations one takes as *pro tanto* reasons, *prima facie* evaluative judgments or appearances of the good) to all-out judgments (of the form “this is what I should do”, “this is what I have most reason to do”, “this is the most desirable option available”, “this is the best thing to in this situation”, or the like). For this reason, I group them all together as views according to which our behavior is determined by judicative reason.

For ease of exposition, I will say that according to each one of these views, the basic operation of judicative reason is to move from a set of *pro tanto* reasons to a normative verdict about what one has, all things considered, more reason to do. There are two ways in which to conceive of the view that our behavior is determined by judicative reason. According to what I will call the *reasons-to-motivation model*, the *pro tanto* reasons one takes into account produce a motivation in the direction of the action they favor whose strength corresponds (at least when the agent is not akratic) to the weight ascribed to that reason. This ensures that the agent will be most motivated to perform that action she takes herself to have the most reason to perform, and, thus, that she will act accordingly. I will argue that this model is simply a variation on the hydraulic model, and thus that it fails to account for multiple-incentives cases. According to what I will call the *reasons-to-judgment model*, we

should do away with talk of motivational forces. We are not beings that are moved by motivations. Rather we are beings that move from the consideration of *pro tanto* reasons to intentions (which are either identical to judgments about what we have, all things considered, most reason to do or directly determined by such judgments) and then execute these intentions. I will argue that intentions thus conceived are not end-directed intentions and, therefore, do not allow us to account for multiple-incentives cases. One could claim that the goal at which an action aims is fixed not by the intention of the agent, but rather by the reasons the agent took to favor the performance of the action in question. I argue, however, that this proposal also fails to account for multiple-incentives cases. Finally, I argue that in order to account for multiple-incentives cases we should take intentions to be more than normative judgments to the effect that an action is the best thing to do or what we have most reason to do. If we conceive of intentions as plans (or parts of plans) then they can fix the goal at which a particular action aims in the way required to account for multiple-incentives cases. I conclude, therefore, that we should conceive of the will as a capacity to form and adopt plans of action in response to certain considerations.

2. Reasons-to-Motivation Model

In the previous chapter I argued that multiple-incentives cases show that the goals at which our actions aim are not fixed by the incentives to act we recognize and, given the desire-goal link, that our desires are not motivational forces (such that our behavior is determined by the power struggle between opposing forces). According to the reasons-to-motivation model, our behavior is determined not by motivational forces that originate in our desires, but rather by motivational forces that are produced by an operation of judicative reason. This view can be formulated as a set of thesis: (a) to take a consideration as a *pro tanto* reason to perform an action A involves moving, in practical thought, from the consideration to a motivation to perform A; (b) at least in non-akratic cases, the strength of the motivation produced corresponds to the weight the agent ascribes to the reason in question; (c) when an agent takes several considerations to provide *pro tanto* reasons to perform A they combine to produce a stronger motivation to perform A and (d) whenever an agent acts, she does what she is, at the time, most motivated to do.

There are strong hints of such a view in the writings of many philosophers who reject the claim that we are moved by desires. Dancy claims that a desire is a “state of being

motivated” produced by the consideration of what one takes to be a fact. Combine that with the claim that “the only thing necessary to take us from motivation to action is the absence of contrary motivation, or the fact that contrary motivations were 'weaker' than this motivation” (Dancy, 2000, p.85) and we seem to have an expression of the reasons-to-motivation model. Setiya also seems to commit to this view when he claims that a consideration is a reason for one to perform action *A* just in case it is a good disposition of practical thought to pass from the belief in that consideration, perhaps combined with other psychological states, to a motivation (with a particular strength) to do *A* (Setiya, 2014, p.229). And it seems as if the same can be said of Scanlon in light of his claim that “the only source of motivation lies in my taking certain considerations [...] as reasons” although the “*strength* of this motivation varies depending on what happens—for example, on the degree to which I attend to a given consideration, focus on it, and ignore others” (Scanlon, 1998, p.35, my emphasis). In this passage Scanlon is discussing cases of akrasia, in which the agent acts in a way that does not accord with her judgment about what she has most reason to do. But the passage strongly suggest that what happens in this case is that certain considerations produce a motivation that is disproportionate to the weight of the reason it provides and that that is what ultimately leads the agent to perform the akratic action. The normal, non-akratic case, therefore, would be one in which the consideration produces a motivation whose strength corresponds to the weight of the reason it provides and, consequently, the agent’s assessment of the relative weight of the reasons she acknowledges corresponds to the strength of the motivations she has – in such a way that the agent is moved by her strongest motivation to perform the action she believes is recommended by the weightier set of reasons.

Now, whenever an agent takes the fact that an action *A* is conducive to a goal *G* as a reason to perform *A*, it follows that the agent takes *G* as in some sense desirable, valuable or worthy of being pursued. The fact that an action leads to a worthless outcome, by the agent’s own lights, could never be in itself a reason for one to perform that action. Even if the reason the agent recognizes is fully expressed as “action *A* is conducive to *G*”, believing the goal to be desirable is a condition for the agent to recognize that putative fact as a reason to perform *A* and, according to the reasons-to-motivation model, for her to be motivated accordingly. One could claim, then, that according to this view the source of the motivation is not a belief-desire pair, but rather a pair of beliefs, one of which has an evaluative content (see Dancy,

2000, p.13). This pair of beliefs underlies the recognition of a reason to act, in exactly the same way in which a belief-desire pair underlies the recognition of an incentive to act.

Now, this view is simply a variation in the hydraulic model. As such it faces exactly the same problems. In particular it cannot account for multiple-incentives cases. In the volunteer work example, Mary took the fact that volunteering would further her goal of being admitted at the University as a reason for volunteering. According to the reasons-to-motivation model, therefore, Mary had a motivation pointing in the direction of volunteering that was stronger than competing motivations and the consideration that volunteering would further her admission at the University contributed to the strength of that motivation.

We can show that this description of Mary's case is incompatible with the claim that hers is a multiple-incentives case if we can show that supporters of the reasons-to-motivation model are committed to a principle analogous to the desire-goal link. Given that the hydraulic model and the reasons-to-motivation model share the same structure, it should come as no surprise that that is the case. Trivially, for an action *A* (as performed by *P*) to aim at goal *G* is simply for *P* to choose to perform *A* as a way of pursuing goal *G*. But if one's behavior is determined by motivational forces, what is it to choose to perform *A* as a way of pursuing goal *G*? To choose to do *A*, according to this thesis, is simply for the reasons one acknowledges to produce a motivational force in the direction of *A* that is stronger than competing motivations (perhaps this motivation should be accompanied by the judgment that *A* is the action one has most reason to perform, but this will not make any difference here). In the reasons-to-motivation model there is no such thing as being *motivated* to do *A as a way of pursuing G*. That can only be understood as the claim that one was motivated to perform *A* by the consideration that that *A* is conducive to *G*. To choose to perform *A as a way of pursuing goal G* must be, then, to be motivated to perform *A*, at least in part, by the consideration that *A* is conducive to *G*, in combination with the belief that *G* is desirable or valuable. But then it follows that:

MOTIVATION-GOAL LINK: If the consideration that action *A* is conducive to *G* combined with the belief that *G* is desirable is part of what motivated agent *P* to perform action *A*, then *A* aims at goal *G*.

When combined with the reasons-to-motivation model, the motivation-goal link renders multiple-incentives cases impossible. It entails, for instance, that Mary's action of

volunteering aimed at furthering her goal of being admitted at the University. And that, I take, is a good reason to reject the reasons-to-motivation goal.

This argument, of course, has exactly the same structure as the argument against the hydraulic model I put forward in the previous chapter. By now we are already familiar with the kind of reply that could be offered to such an argument. One could hold, for instance, that in a case such as Mary's we claim that she acted with a view only to promoting the well-being of others because altruistic considerations would produce a motivation strong enough to move her even if she did not believe that volunteering would increase her chances of being admitted at the University or did not believe that being admitted at the University was a desirable or valuable outcome. Or one could claim that we ascribe a purely altruistic goal to her because we judge her action praiseworthy in light of the fact that altruistic considerations would produce a motivation strong enough to move her even in the absence of self-interested reasons. Or finally, one could claim that in a case such as Mary's the pair of beliefs constituted by the belief that volunteering would increase her chances of admission at the University and that being admitted at the University is desirable remained inoperative, in the sense that it did not produce any motivation, even though Mary recognized the corresponding reason to volunteer. All these replies will fail, for exactly the same reasons the corresponding replies on behalf of the hydraulic model failed. The fact is that if my arguments against the hydraulic model are successful, they also show Pure Cognitivism to be false.

Therefore, multiple-incentives cases show not only that we are not motivated to perform a particular action A by the belief-desire pairs that underlie the incentives we recognize to perform A but also that we are not motivated to perform A by the considerations we take to provide reasons to do A.

But if we are not motivated by our desires nor by our reasons, what motivate us? I believe the answer should be that we are not beings that are moved by motivations. The whole idea of motivations conceived of as motivational forces, with a particular strength, that dispute the determination of our behavior with other motivational forces is, I am contending, misguided and should be abandoned, regardless of what we take the source of these motivations to be. There are no such things as motivations that can stack up and make you more motivated to this or that. The conclusion to be drawn is that rather than being moved by motivations (regardless of whether they assail us or are produced by us) we are beings who

decide to act in one way or another and then act, if everything goes well. To say that a consideration moved us to perform an action is simply to say that we decided to act in light of that consideration.

3. The Reasons-to-Judgment Model

It seems then that the reasons-to-judgment model, which does away completely with talk of motivational forces, is best suited to account for multiple-incentives cases. According to this view, we are beings who simply form intentions in light of certain considerations, which we take to provide *pro tanto* reasons, and then execute these intentions. But even when we reject the notion of motivational forces, multiple-incentives cases continue to pose a challenge for those who would like to identify the will with judicative reason.

If the all-things-considered judgments issued by judicative reason are fully expressed in the form “action *A* is the best thing to do in this situation” or “action *A* is what I have most reason to do in this situation” then the intention that is identified with such judgments may well be fully expressed in the form “I intend to do *A*”. Intentions that are fully expressed in this way are not end-directed intentions and cannot specify the goal at which action *A* aims. Two agents can come to the conclusion that performing an action *A* is the best thing to do in the circumstances they find themselves but perform *A* with a view to very different goals. For instance, both Mary and Pete may believe that volunteering is the best thing to do in the situation they find themselves, but Mary may volunteer with a view to helping people in need whereas Pete, who is terribly selfish, volunteers only with a view to getting something out of it. Therefore, if the goal at which a particular action aims is to be determined by the activity of judicative reason, then it must be determined not by the conclusions judicative reason draws but rather by the premises from which these conclusions are drawn – i.e., by the reasons the agent took to favor the performance of the action. According to this suggestion, agents with the same intention act with a view to different goals because the premises from which they derive this intention (conceived of as an all-things-considered normative judgment) are different. One could claim, for instance, that an action *A* aims at goal *G* when the conclusion that *A* is the best thing to do in the current circumstances is derived from a set of *pro tanto* reasons that include the putative fact that *A* is conducive to *G*.

Indeed, if one accepts the reasons-to-judgment model and identifies the will with judicative reason, one *must* take this path. For trivially, for an action *A* (as performed by *P*) to aim at goal *G* is simply for *P* to choose to perform *A* as a way of pursuing goal *G*. But according to the reasons-to-judgment model, to choose to do *A* is simply to move from the consideration of a set of *pro tanto* reasons to the intention of doing *A*, i.e., to the all-things-considered judgment to the effect that *A* is the best thing to do in the current circumstances. What is it to choose to perform *A as a way of pursuing G* then? It seems that the answer must be this: to choose to do *A* as a way of pursuing *G* is to move to the intention of doing *A* from a set of reasons that includes the consideration that *A* is conducive to *G*. But then it must be the case that:

REASON-GOAL LINK: If agent *P* judges that *A* is the best thing to do in circumstances *C* in light of the consideration that *A* is conducive to *G* and is moved to do *A* by this judgment, then *A* aims at goal *G*.

This principle, however, renders multiple-incentives cases impossible, in the exact same way in which the desire-goal link and the motivation-goal link did. In multiple-incentives cases, the reasons in light of which the agent reaches the judgment that doing *A* is the best available course of action may very well include both the consideration that *A* is conducive to *G* and the consideration that *A* is conducive to *E*. In these cases, nevertheless, the action aims at only one of these goals. If the goals with a view to which an agent acted are fixed by the considerations she acknowledged as reasons for acting as she did, these cases would be impossible: they would all be reduced to compound-motive cases.

That can be made clear if we consider an ordinary case of employment of judicative reason: suppose you are wondering whether or not you should accept a job offer in another state. Several facts testify in favor of accepting the job (you will get a significant raise, the work you will be doing is more stimulating, etc.) but several facts testify against accepting the job as well (you will be away from your friends and family, the city to which you will have to move is quite violent, etc.). You must weigh the various considerations at stake to arrive at a judgment about what is the best choice. Suppose that, in the end, you come to the conclusion that, all things considered, it is better for you to accept the job and that is what you do. What is the intention with which you act in this case? It surely is inappropriate to claim that you accepted the job aiming only at earning a better salary or only at having a more stimulating job. The fact of the matter is that in accepting the job you are aiming at

getting a better salary *and* having a more stimulating job *and* whatever else you took in consideration. The activity of judicative reason in this case cannot specify a particular goal (corresponding to only one of the *pro tanto* reasons taken into account) as *the goal* at which the action it recommends is directed. Admittedly, this is not a multiple-incentives case – we are perfectly comfortable with the claim that in this case the agent has a compound-motive. But if that is judicative reason’s *modus operandi*, why should the result be any different in multiple-incentives cases, where the agent performs an action that she takes to be conducive to more than one valuable or desirable goal?

In the job offer case the agent does not take any of the available *pro tanto* reasons to accept the job to be sufficient on its own to recommend that action given the available *pro tanto* reasons not to accept the job. Perhaps practical reason could specify *the goal* at which an action aims in a multiple-incentives case by declaring a particular *pro tanto* reason to be sufficient to recommend the action. One could say that an action *A* aims at goal *G* and goal *G* alone, even though the agent recognizes several different incentives to perform *A*, if, and only if, the agent takes the fact that *A* is conducive to *G* to provide a *pro tanto* reason to perform *A* that is not outweighed by the set of *pro tanto* reasons not to perform it. According to a plausible interpretation, this is Davidson’s considered view: he holds that judicative reason issues judgments to the effect that a particular desirable characteristic of an action is “enough to act on” (i.e., that the consideration that the action has this particular characteristic provides a *pro tanto* reason to perform it that is not outweighed by other considerations) and this “allow us to give the intention with which the action was performed” (Davidson, 1980, p.87-8).

This proposal can account for some cases of multiple-incentives but not all. In many such cases, more than one incentive may provide the agent with a sufficient *pro tanto* reason to act and yet the goal with a view to which the agent acts may correspond to only one such reason. For instance, the circumstances in which Mary acts in the volunteer work example may be such that she has very few and very weak reasons not to volunteer and, thus, that both the fact that the charitable action will promote the well-being of others and the fact that it will further her goal of being admitted at the University may provide a *pro tanto* reason to perform it that is not outweighed by opposing considerations. And, notwithstanding, she can act only with a view to the well-being of others and not with a view to both the well-being of others and her own personal interest.

Another option for the supporters of the reasons-to-judgment model is to claim that our motive attributions are guided by our moral assessment of the action. One could claim that we judge the performance of an action praiseworthy when the agent acknowledges a moral reason as sufficient reason to act even if she also recognizes other sufficient reasons to perform the same act (because what matters for us when we are concerned with evaluating actions is whether the agent acknowledges certain reasons) and that when we judge an action praiseworthy we are inclined to claim that it was aimed at a moral goal (such as relieving the suffering of others) and not at a personal goal. For instance, when assessing a charitable action, what matter for us is whether the agent took the appropriate reason to be charitable (namely, the fact that the action in question would promote the well-being of others) as a sufficient reason to act. If that is the case, then we judge the agent's action praiseworthy. And then, in light of this positive assessment of the action, we claim that the agent acted with a view to promote the well-being of others, even though she also took the fact that the action would further a personal goal of hers as a sufficient reason to act. This is to propose a criterion for acting with a view to a commendable or moral goal that takes this form: "An agent S performs an actions A with a view to, and only to, a moral goal M (and, consequently, her action is more praiseworthy than exactly similar actions performed with a view to another goal) if, and only if, S performs A and takes the fact that A is conducive to M to be a sufficient reason to perform A ".

We already discussed a similar proposal in the previous chapter, and we should expect this proposal to be vulnerable to same objections. It leads to arbitrary moral or evaluative distinctions. Consider two agents, S_1 and S_2 both of which judge that they have, all things considered, a reason to perform a charitable action A and do perform it. S_1 acknowledges the fact that action A is conducive to moral goal M (say, the promotion of the well-being of people in need) as a reason R_1 to perform A , she also acknowledges the fact that A promotes end E on which she has a personal interest as a reason R_2 to A and recognizes some reasons not to do A . S_1 does not take R_1 alone to be sufficient to perform A , although she takes R_1 and R_2 together to be sufficient. S_2 acknowledges a very similar set of reasons: she acknowledges R_1 and R_2 and ascribe them the same weight S_1 does (this could be ascertained by putting S_1 and S_2 in a variety of counterfactual situations, in which their interest, concerns and inclinations would provide them with reasons of varying weights, and checking what their all things considered judgments would be in those circumstances). Nevertheless, S_2 takes R_1 to

be sufficient reason to perform A . That means that S_2 ascribes less weight to the set of reasons opposing A than does S_1 . Let us assume that this difference is explained by the fact that S_1 has a personal interest E_2 that is frustrated by the performance of A and, thus, provides a *pro tanto* reason not to perform A . S_2 does not have an equivalent interest and thus her set of reasons against doing A has less weight. According to the proposed criterion, S_2 's action is more praiseworthy than S_1 's. But, by stipulation, the only relevant difference between them is that S_1 acts at the expense of one more personal interest than does S_2 . That, if anything, should make S_1 's action more praiseworthy. We should, therefore, reject the proposed criterion for the same reason we rejected the sufficiency criterion in the previous chapter.

The failure to account for multiple-incentives cases when conceiving of the will as judicative reason should not come as a surprise. According to this proposal, the goals our actions aim at are a function of the relative weight of our reasons. In the same way, according to hydraulic model, the motives for which we act and, therefore, the goals at which our actions aim are a function of the relative strength of our pro-attitudes. Given the structural similarity between these views, it is not surprising that they should face similar problems. Just as according to the hydraulic model, several desires combine in determining our behavior, according to the reasons-to-judgment model, several *pro tanto* reasons combine in determining our intention. Because they are so combined, we need a criterion to specify which among them determines the goal at which the action in a multiple-incentives case aims. But no satisfactory criterion can be found.

4. Disregarding Reasons

One could hope to account for multiple-incentives cases while identifying the will with judicative reason by removing from the set of *pro tanto* reasons that are taken into account in the agent's reasoning those considerations that provide a sufficient reason to act but do not correspond to the goal at which the action aims. Thus, in the volunteer work example, we could claim that Mary acted with a view only to the well-being of others provided she had not taken into account in her reasoning the fact that the action would promote a personal interest of hers. We could conceive of an agent whose capacity for practical reasoning is engaged when her attention is drawn to the urgent needs and suffering of another person; she then starts reasoning about what the best course of action is in her circumstances, immediately considers the fact that action A would promote the well-being of

the other person, declares that this fact provides a sufficient reason to perform *A* (dispensing additional considerations in favor of the action) and brings her reasoning to an end before having an opportunity to consider the fact that *A* will also promote a personal interest of her. In this case, it would be appropriate to describe this agent as acting with a view only to the well-being of the beneficiary of her act, and not with a view to the promotion of a personal interest of hers. And I believe something along these lines comes close to capture an important truth about virtue: that some facts are salient for the virtuous agent and that her modes of reasoning are particularly sensitive to such facts. Nevertheless, not all cases of multiple-incentives can be accounted for in this manner. Consider this slight variation in the volunteer work example:

VOLUNTEER WORK – TAKE 2: Mary is a really benevolent person. She cares for the well-being of others and does what she can to promote it. One day she is riding the bus when she sees a group of people distributing meals to homeless people in a park. That gets her thinking that she could spare a few hours a week to do something to help people who are down on their luck. She does not know who were the people distributing meals. She decides then to do some research online to figure out where and how she could help. Sometime later she does that. During her research she discovers two volunteer work organizations that prepare and distribute meals to homeless people in her town – call them Alpha and Beta. There is not much difference between the two organizations. Both seem to be run by people who take their job seriously and the good Mary would be able to bring about by joining either of them is the same. She has pretty much decided that she will volunteer in one of these two organizations. She then discovers that a friend of hers is volunteering at Alpha. She likes very much to spend time with this friend. Mary then decides to volunteer at Alpha.

Why did Mary choose to volunteer at Alpha rather than Beta? Because doing so would allow her to spend time with her friend. That is a consideration she explicitly considered while making her decision. That, however, does not seem to preclude the possibility of claiming that the motive for which she volunteered was simply that doing so would contribute to relieving the suffering of people in need and not that doing so would contribute to relieving the suffering of people in need *and would allow her to spend time with her friend*. And that means that her goal in volunteering was simply to help people in need and not to spend time with her friend. Recall the remark that a truly virtuous person performs virtuous actions for their own sake. That, I suggested, is best understood as the claim that a virtuous agent performs the virtuous action with a view to its proper end – she performs

courageous acts in order to protect what is valuable, and benevolent acts in order to promote the well-being of others, and not in order to enjoy the praise or some other advantage she will get. If that is correct, to insist that Mary's action, in the example above, must be aiming at the promotion of her goal of spending time with her friend is to insist that her choice is incompatible with the true virtue of benevolence. But that would be preposterous. Being a truly benevolent person does not involve choosing the least attractive way of helping people whenever one has a choice.

But this is bad news for the reasons-to-judgment model. This model is committed to the view that the goals at which an action A aims are determined by the reasons on which the agent grounds the judgment that A is the best available course of action or the one best supported by reasons. And the fact that volunteering at Alpha would contribute to her goal of spending time with her friend surely is among the reasons that grounded Mary's conclusion that volunteering at Alpha was the best option. Worst still, this problem cannot be solved by painting a more complex image of judicative reason which includes the ability to disregard certain reasons altogether.

Consider Scanlon's proposal.³⁴ According to Scanlon (1998, p.17), to take a consideration as a reason to perform an action is simply to take that consideration as counting in favor of performing the action. Our basic reasoning abilities are, therefore, the ability to recognize certain facts or putative facts as reasons to act in a particular way and to weigh these reasons in order to determine what we have overall reason to do. But things can get a little more complicated:

“[...] reasons can be related to one another in more complex ways. I may, for example, judge one consideration, C, to be a reason for taking another consideration, D, not to be relevant to my decision whether or not to pursue a certain line of action. [...] The reason-giving force of C not only competes with that of D; it urges that D lacks force altogether (at least in the given context). Often, our judgment that a certain consideration is a reason builds in a recognition of restrictions of this kind at the outset: D may be taken to be a reason for acting only as long as considerations like C are not present.” (Scanlon, 1998, p.51)

According to Scanlon, therefore, judicative reason can not only weigh reasons, it can also disregard certain reasons in light of others – i.e., it can, in light of a consideration, refuse to ascribe any weight to a consideration that, in other context, would count in favor or

34 To be clear, Scanlon is not concerned with accounting for multiple-incentives cases but is rather making a claim about the ways in which reasons can relate to each other. My point in what follows is simply that noting the relations between reasons to which Scanlon draws our attention is not enough to account for multiple-incentives cases. Something else is needed.

against acting in a particular way. In a very similar vein, McDowell characterizes the virtuous person as someone who sees her situation in such a manner that "some aspect of the situation is seen as constituting a reason for acting in some way" and not only a reason to be weighed against opposing reasons but one that complete *silences* opposing reasons (McDowell, 1979, p.335). The claim that opposing reasons are silenced is naturally read as the claim that opposing reasons are disregarded, that no weight is ascribed to them, even though they would have some weight in different circumstances.

In light of these claims, one could suggest that it is possible to account for multiple-incentives cases while holding on to the reasons-to-judgment model by claiming that in these cases the agent takes the relevant reason (the one that fixes the goal with a view to which she acts) as a sufficient reason to act and disregards (i.e., ascribes no weight to) the other reasons that she acknowledges and that would normally count in favor of the action she performed. The variation on the volunteer work case shows, however, that this suggestion will not work. Mary not only recognizes the fact that volunteering at Alpha will further her goal of spending time with her friend as a reason to volunteer at Alpha but also ascribes weight to it – so much so that she judges the option of volunteering at Alpha to be better than the option of volunteering at Beta.

Another possibility is to hold that even when we acknowledge several reasons to do A, we can ground the judgment that doing A is the best available course of action on only one of these reasons (which we take to be a sufficient reason to do A). In light of this suggestion, multiple-incentives cases can be understood as cases in which although the agent acknowledged several reasons favoring the action she performed, she grounded her normative judgment on only one of the relevant reasons (the one that fixes the goal with a view to which she acts).

This suggestion also fails to account for cases as the one above. In our example, Mary volunteers at Alpha with a view to promoting the well-being of others and not with a view to spending time with her friend. But in light of what reason does she come to the conclusion that volunteering at Alpha is a better option than volunteering at Beta? It seems that the *only* available reason is that doing so would further her goal of spending time with her friend. So, according to the suggestion under consideration, her action should aim at furthering that goal – and that is exactly what we would like to deny.

I conclude we should deny the identification of the will with judicative reason and, consequently, the identification of intentions with normative judgments.

5. Reasons-to-Decision Model

If we are to identify the will with reason in its practical application, we should not think of reason as judicative reason (the capacity to move from *pro tanto* reasons to normative judgments which issue in action). Rather we should think of it as genuine *practical* reason. Practical reason thus conceived is not a capacity to issue judgments about practical matters but rather a capacity to decide in light of reasons. Its activity concludes not in a normative verdict but on a decision or intention. I believe this suggestion is on the right track. However, whether or not this suggestion succeeds in accommodating multiple-incentives cases depends on how we conceive of intentions. In particular, if we take intentions to be simply the resolve or determination to perform a particular action (so that their content is exhaustively expressed by “I intend to do A”) then this suggestion represents no improvement with regard to the view that identifies the will with judicative reason. The problem, again, is that an intention fully expressed by “I intend to do A” is not an end-directed intention and as such it cannot specify the goal at which action A aims. The goal must then be specified by the reasons on which the decision (the act of adopting the intention) is grounded. And that brings with it the same problems we have been discussing.

It could be suggested that by removing the normative judgment from the equation, we open up the possibility for a more fine-grained selection of the reasons implicated in a decision. Dancy, for instance, holds that it can be the case that an agent took a consideration to favor acting in a particular way, decided to act in that way and yet did not decide to act *in light of that consideration*. Furthermore, he thinks that distinction is all we need in order to account for multiple-incentives cases. Here is what he has to say (when discussing the claim that we should accept Davidson’s causalism because it allegedly allows us to account for multiple-incentives cases):

“The most direct response to Davidson, however, is just to say that the difference between those reasons for which the agent did in fact act and those for which he might have acted but did not is not a difference in causal role at all. It is just the difference between the considerations in the light of which he acted and other considerations which he took to favour acting as he did but which were not in fact the ones in the light of which he decided to do it. This is admittedly not very informative, since we have to allow that we have offered no analysis or philosophical account of the ‘in the light of’ relation. I suspect, however, that no such analysis or account is available to be given, without therefore supposing that this has

any tendency to show that the relation concerned does not exist. It is what it is, and not another thing; and if it cannot be analysed, so much the worse for the more global pretensions of analysis.” (Dancy, 2000, p.163).

Instead of claiming that we have the capacity to drain some considerations of their reason-giving force in light of other considerations (as Scanlon does), Dancy claims that we are capable of recognizing several reasons to act in a particular way while at the same time deciding to act *in light* only of some of these reasons. He does not provide an explanation of what it is to *decide to perform an action in light of a consideration* nor does he have anything to say about how it differs from taking a consideration to favor the performance of a particular action, but he holds that these are different relations and that we can grasp the difference between them.³⁵ In light of these claims, one could suggest that multiple-incentives cases are simply cases in which although the agent acknowledged several reasons favoring the action she performed, she decided to act in light *only* of the relevant reason (the one that fixes the goal with a view to which she acts).

For this suggestion to work, however, one would have to hold that in the example above Mary did not decide to volunteer at Alpha in light of the consideration that doing so would allow her to spend more time with her friend. But that is clearly the case. First, that seems to be the only available reason for her to choose Alpha and not Beta. Second, it is correct to claim that Mary decided to volunteer at Alpha because that was, from her point of view, the best available option. Indeed, her decision process consisted in evaluating the two options so as to determine which one was the best. Part of what made the option of volunteering at Alpha better than the alternative was the fact it would allow Mary to spend time with her friend. This consideration, therefore, cannot be excluded from the set of considerations in light of which she decided. This, of course, does nothing to show that there is a problem with the distinction between reasons the agent merely acknowledges and reasons in light of which she decides.³⁶ The point is rather that if an agent chose a course of action because she thought it was the best course of action, then the considerations in light of which

35 It is hard, though, to see how the claim that we are moved by the stronger motivation (something which Dancy seems to accept, as discussed above) can be reconciled with the claim that one can decide to act in light of only some of the reasons one took to favor the action one performed. If all the reasons one took to favor an action contribute to one’s motivation, then it seems that one was moved by the whole set of reasons. It seems that one would be able to decide in light of a restricted set of reasons only if the other reasons one recognized did not contribute to producing the motivation. That would lead to same problem faced by the view that in multiple-incentives cases most of the agent belief-desire pairs remain inoperative, which was what motivated the rejection of Davidson’s account of multiple-incentives cases in the previous chapter.

36 This distinction will play a central role in the argument of the next chapter.

she decided coincide with the considerations on which she grounded her normative judgment (even if the will is not identified with judicative reason).³⁷

The second take at Mary's case give us good reason, therefore, to hold that in order to account for multiple-incentives cases we need a more robust conception of intentions. One could resist this conclusion by arguing that we can do just fine with the idea that intentions are fully expressed by "I intend to do A" and that the goals at which our actions aim are fixed by the reasons on which the intention is grounded as long as we introduce a distinction between end-decisions and means-decisions. The goals at which our actions aim, the suggestion would go, are determined by the reasons that ground the end-decision – the decision to pursue a certain goal. The reasons that ground a mere means-decision (a decision regarding which means to adopt in the pursuit of a pre-established goal) do not fix the goals at which the action aim. In Mary's case, it seems as if she had already reached the end-decision to do something to promote the well-being of others before comparing the options of volunteering on Alpha and on Beta. The consideration that volunteering at Alpha would allow her to spend time with her friend grounds only a decision regarding the means for promoting the well-being of others – as such it is not part of the set of reasons that fix the goal of the action. I believe this suggestion correct, but it already presupposes a more robust conception of intention. What I called a means-decision also has to result in an intention. That intention, however, cannot be fully expressed by "I intend to do A", for one has decided not only to do A but to do A with a view to a particular, predetermined goal. This intention is fully expressed by "I intent to do A in order to G". The content of the G-slot of this intention, however, is not determined by the reasons that ground it. As Mary's case shows, one can, for instance, form the intention of volunteering at Alpha in order to help others in light of the fact that

37 Setiya makes a suggestion that is somewhat similar to Dancy's. According to Setiya we are capable of choosing the reason for which we act among the several reasons to act we recognize. To choose a reason as *my* reason is to make it the case that it is the reason *for which* I act (Setiya, 2007, p.30 and p.39-2). The way in which I choose *my reason* for acting is by forming an intention, which Setiya takes to be a desire-like belief to the effect that I am doing A (or that I am going to do A) for the reason that P (Setiya, 2007, p.42). Adopting one such intention makes it the case that my reason for doing A is that P, even though I might acknowledge several other reasons to do A. Again, in light of these claims, one could suggest that multiple-incentives cases are simply cases in which although the agent acknowledged several reasons favoring the action she performed, she took *only* the relevant reason (the one that fixes the goal with a view to which she acts) as *her reason* to act. This suggestion fails to account for Mary's case for the same reasons Dancy's suggestion fails. Surely the fact that volunteering at Alpha would allow her to spend time with her friend is one of *her reasons* for choosing Alpha (and not simply a consideration she saw as counting in favor of volunteering at Alpha). Furthermore, Mary decided to volunteer at Alpha because she thought it was the best choice. In part, it was the best choice because it would allow her to spend time with her friend. That consideration, therefore, cannot be excluded from the set of considerations that provide Mary's reasons for choosing Alpha.

volunteering at Alpha will allow me to spend time with a friend – the goal specified in the reason and the goal specified in the intention differ. What we have here is, therefore, an end-directed intention. And we need a more robust conception of intentions to account for it.

If this is correct, then the existence of multiple-incentives cases has a number of interesting consequences. First, these cases show that we are not moved by motivational forces, whether they are the product of our desires or of the reasons we recognize. Second, they show that we are not moved by the normative judgments that are the conclusions of judicative reason. Third, they show that even if we identify the will with practical reason, we need a robust conception of intentions – one according to which it is the content of the intention and not the reasons on which it is grounded that determine the goals at which the action that executes it aims. In the next section I offer a view of intentions that fits the bill.

6. The Will as the Capacity for Planning

Let us take stock. One of the conclusions of the previous chapter was that in order to account for multiple-incentives cases we should hold that the goal at which an action aims is determined by the intention the agent executes when she performs the action. In this chapter I considered the possibility of identifying the will with judicative reasons. I argued that the all-out normative judgments that are the conclusions of judicative reasoning do not have the proper form to determine the goal at which a particular action aims. If this proposal is to work, we should think of the goal at which an action aims as determined by the premises of such reasoning, i.e., by the reasons the agent took to count in favor of performing the action. This approach, however, fails to account for multiple-incentives cases. Finally, even if we conceive of the will as a capacity for practical reasoning that concludes directly on intentions, we cannot account for multiple-incentives cases if we do not accept the idea that the goal at which an action aims is determined by the content of the corresponding intention. What we need, therefore, is a conception of intentions according to which they have a content that is rich enough to specify the goal at which the intended action aims.

If intentions determine the goals our actions aim at and, as I have argued, the goals at which our actions aim are not determined by the belief-desire pairs we happen to have, nor by the normative judgments we happen to hold, the reasons we happen to recognize or the reasons in light of which we decide, it seems as if we should hold that intentions are not reducible to desires, beliefs, normative judgments or some particular combination of these

kinds of mental states. We should take intentions to be a different, irreducible kind of mental state.

We may get a better grasp of intentions by characterizing them in functional terms – i.e., in term of its typical relations with other mental states, with our behavior and our reasoning patterns. For instance, according to Bratman’s influential account, intentions are *conduct-controlling* and *stable* (Bratman, 1999, p.21-22). Intentions are conduct-controlling in that in normal conditions and if they are not revised, they lead the agent to act. And they are stable in that (i) we are disposed not to revise them in the absence of new information, (ii) we are disposed to engage in means-end reasoning where the ends are fixed by our intentions and (iii) our current intentions, combined with our beliefs, determine which options are to be considered admissible in practical reasoning (Bratman, 1999, p.34-5). The discussion of multiple-incentives cases has led us to ascribe a further function to intentions: they fix the goals at which our actions aim. And it is fair to demand an explanation of how they do that. If intentions are a mental state properly expressed by “I intend to do *A* in order to *G*” it cannot be fully characterized by its tendency to control our behavior in such a way that we perform *A* and by its tendency to lead us to form intentions of performing preliminary steps to *A* as well as to lead us to declare the formation of intentions to perform actions that are incompatible with *A* inadmissible, for the intention expressed by “I intend to do *A* in order to *E*” will present exactly the same functional profile. The phrase “in order to *G*” has to express itself somehow in the agent’s behavior, in her mental states or in her patterns of practical reasoning. Otherwise, it will seem as if it is merely a form of words the agent holds before her mind as she executes her intention to *A* – as if an agent could make it the case that she is doing *A* with a view to *G* and not to *E*, and thus acting for, say, a moral motive instead of a self-interested one, simply by repeating to herself “I am doing this in order to *G*”.

Now, if the intention of doing *A* in order to *G* cannot fully express itself in the agent’s action nor in the agent’s pattern of instrumental reasoning about how to achieve *A*, what is left? Perhaps this intention constrains further intentions in a different way than the intention of doing *A* in order to *E*. An intention of doing *A* in order to *G* constrains further intention by deeming inadmissible the option of performing any action that is incompatible with *A* or whose execution would render *A* ineffective in bringing about *G*. So, the intention of helping at a soup kitchen tomorrow in order to promote the well-being of others fixes a screen of admissibility that excludes the options of performing an alternative action at the

same time tomorrow as well as the intention of performing any action that will render your action of helping at the soup kitchen ineffective in promoting the well-being of other (for instance, doing something that would prevent any homeless person from getting to the soup kitchen). The same is true, however, of the intention of helping at the soup kitchen tomorrow in order both to promote the well-being of other *and to promote my reputation as a charitable person*. How are they different (as they must be in order for us to account for multiple-incentives cases)? Perhaps if you really intend to help at the soup kitchen in order only to promote the well-being of others, then you will consider inadmissible the option of spreading the word about your volunteer work, for that would show that you were aiming at promoting yourself. But that is not necessarily true – perhaps you came to believe that it would be good to share your experience at volunteer work in order to encourage other people to volunteer as well. Maybe then you will consider inadmissible the option of spreading the word about your volunteer work in order to promote yourself. But there is no reason for the agent to declare that option inadmissible. Surely the performance of the action in question does not frustrates the intention of promoting the well-being of others by helping in the soup kitchen. If it is to be deemed inadmissible the reason cannot be that it fails to pass the screen of admissibility created by the agent’s prior intention but rather that the agent does not take it to be a worthy option. I see no reason to assume that to be the case, however. The agent may very well think that having acted for a commendable motive (executing the intention of helping at the soup kitchen in order only to promote the well-being of others) there is no reason not to collect the social reward available to her. It will not do as well to suppose that the agent who really intends to help at the soup kitchen in order only to promote the well-being of others will not as a matter of fact form the intention of spreading the word about her volunteer work in order to promote herself – not only can she take that to be an eligible option having already acted for a moral commendable motive, she also may simply give in to vanity later on (which does not entails that her prior action was performed even in part for a vain motive). The only option left is to hold that the intention of helping at the soup kitchen with a view only to the promotion of the well-being of others is simply incompatible with the intention of spreading the word about one’s volunteer work in order to promote oneself in the sense that the two cannot co-exist. That may very well be true, but if it is then it requires an explanation. If they are incompatible that is not because the intended actions are incompatible or because performing one of them prevents the other from achieving its goal. Furthermore, if that is the

only content we can give to the claim that the intentions to do a charitable action with a view to promoting the well-being of other and with a view to promoting the well-being of others and one's reputation are different, it becomes hard to understand why we should care whether one acted with the former or the latter intention. After all, one can act with the compound-goal intention without having the intention to spread the word about one's volunteer action (one may assume it will spread itself) or any other intention that is excluded by the single-goal intention. What is the difference then between acting with the single-goal intention and acting with the compound-goal intention?

We can understand the difference if we assume that to have an intention is the same as settling on a *plan of action*. We may think of plan as an ordered pair {goal; strategy} where the content of strategy-slot represents steps to be taken, in a particular order, to achieve the goal. Having an intention is simply a matter of having a particular attitude directed towards a plan.³⁸ The attitude in question is the attitude of being settled on the plan. Plans of course come in all kinds. They can be quite detailed but usually we have only partial plans. To use one of Bratman's examples (Bratman, 1999, p.31): I may be settled on going to a concert tonight; I know that in order to achieve that goal I have to buy tickets and find a way to get to the concert house; I have, therefore, a vague idea about the strategy component of my plan but in order to implement it I will have to fill in a lot of gaps; I can fill in some or most of the gaps before starting to execute the plan or, as soon as I have a more detailed idea of where to start, I may begin to implement the plan and fill in the gaps as I go. Plans can also be very complex. The strategy-slot may include sub-plans (sub-goals and strategies to achieve those) as well as alternative strategies to accomplish the same goal. One can also be settled on a goal without having settled on any strategy on how to bring it about. In this case one intends simply to bring about *G* without yet intending to perform any particular action. Finally, one can have as a goal simply the performance of an action, in which case performing the action in question will be part of one's strategy.

When an agent executes a plan, the contents of the strategy-slot guide her action. The possibility of this guidance is dependent upon the agent's ability to monitor her action in light of the plan. In order to be able to follow a plan you have to be able to determine whether or not you have performed the step you are endeavoring to perform and whether it achieved the sub-goal at which it was aimed and you must be able to adjust your behavior in light of the

38 This view of intentions is defended by Mele (see 1992, p.150).

feedback you receive (deciding, for instance, whether to go on to the next step, to try to perform the current step again, to modify your plan, to give up, etc.). When something does not go as planned, as when you find that you cannot execute one of the steps of the plan or it fails to produce the expected effect, the execution of the plan stops in its tracks and you have the option of re-execute the step in question, modify the plan or abandon it. When agents halt the execution of a chain of actions and adopt one of these measures, we have indication that they were monitoring their action in light of a plan and get some insight into the content of the plan. Thus, suppose you have a simple plan: your goal is to read your emails and your strategy is to go to your office, open your notebook, type in your login and password and so on. Suppose you go to your office, grab your notebook, open it and it does not turn on as you expected it would. If you are an able plan-follower you must by now know that you performed a step of your plan and it did not accomplish the expected sub-goal. You must also realize that you cannot move on to the next step and you have to be able to identify the available options: you may try to perform the same step again (opening and closing the notebook), or you could modify your plan (say, adding the step of clicking on the notebook's power button) or you could simply give up your goal of accessing your email. Suppose that being a reasonable person you go with the second option and you manage to achieve the sub-goal of turning on the notebook. Being aware of that, you move on with your plan and finally get to read your emails, just like you planned. Now an able plan-follower should be in a position to realize she has achieved her goal, register that information and declare her plan a success, which means that she can stop monitoring her behavior in light of that plan.

If we conceive of intentions in this way, then the difference between the intention of doing A in order to G and doing A in order to G and E is a difference in the content of the plan on which the agent has settled. If one has the former intention, then the plan one has settled on is represented as $\{G; A\}$, if the latter then $\{G \text{ and } E; A\}$. Now, obviously, as far as the actions performed in executing the plans are concerned, there is no difference between those plans. But there is a real difference in the way these agents monitor their behavior in light of each of these plans and that gives content to the idea that they are aiming at different goals.

Consider the case of the agent who forms the intention of volunteering at the soup kitchen (action A) in order to promote the well-being of people in need (goal G) while believing that this action will also promote her reputation as a charitable person (goal E). This

agent is settled on a plan {G; A} and she believes that executing *A* will also bring about *E* but does not aim at *E* (although she welcomes that result).³⁹ This means that in monitoring her plan the agent need not track her action effectiveness in bringing about *E*. That is, if she executes the action and finds that it achieved its goal of promoting the well-being of people in need, then she is in a position to declare her plan a success and can stop monitoring her behavior in light of that plan. On the other hand, if her intention was to do *A* in order not only to promote the well-being of people in need but also to further her reputation as a charitable person, her plan would take the form {G and E; A}. In this case, having found that by working in the soup kitchen she did help promoting to some extent the well-being of people in need, she would still need to monitor her action, particularly by tracking its consequences. Thus, if she notices that nobody has taken heed of the fact that she volunteered, she will be forced to reconsider her plan (should I do something else to make the fact that I volunteered public? Should I just wait and see if with enough time people will acknowledge my contribution?). Whatever her choice turns out to be, the fact is that she would at this point still be monitoring her action in light of his plan, whereas in the case in which her goal is simply to promote the well-being of people in need she would already have declared her plan a success.

The same kind of explanation applies to any case of multiple-incentives. Suppose that an agent recognizes two incentives to give blood: it saves lives and by giving blood she can get a blood donor card that will give her a discount at the movies. She forms the intention of giving blood in order to do her part in the effort of saving lives. Her plan will include several steps: she will eat something in the morning, then she will go to the blood bank, after she goes through the medical screening, she will follow the nurse's instructions and so on. She will monitor her action in light of this plan. If everything goes well, once she finishes her donation, she is in a position to declare her plan a success. If she had the intention of donating blood in order to do her part in the effort to save lives and in order to get a blood donor discount card, things would be different. Suppose that once she gets at the blood bank she planned to go to she sees a poster informing that they are no longer issuing blood donor discount cards. This will prompt her to reconsider your plan: she will have to decide whether

39 Bratman explains the difference between intending do *A* and merely expecting it to bring about *E* and intending to do *A* in order to *E* in terms of how the agent monitors his action inf light of her plan, but he discusses only double-effect cases, in which an agent intends do an action *A* in order to *G*, knows that it will bring about *E* but *does not* want *E* to be the case (see Bratman,1999, p.114-5).

to look for another blood bank or go on with her plan accepting partial failure. The agent with the single-goal intention, in contrast, was not monitoring the effectiveness of her actions in getting her a discount card. Registering the fact that the blood bank no longer issues discount cards will not prompt her to reconsider her plan. Moreover, even if the blood bank was still giving discount cards the agent with the compound-goal intention would not be in a position to declare her plan successful after completing her blood donation. She would still have to monitor the consequences of her act, namely, whether she actually gets a discount card or not. If after a while she does not get the card (she assumes that it is an automatic process and that there is no need to make a requirement after donating blood, otherwise her plan would be different from the plan of the agent that has the single-goal intention), then she will have to consider what to do: should she wait longer, should she re-execute some steps in her plan, should she add a step of making a requirement for the card? This shows that the agent is still monitoring the effects of her action, at a time in which the agent with the single-goal intention has long ceased monitoring her action in light of her plan. This of course is compatible with the supposition that the agent with the single-goal intention will be satisfied if she gets a discount card, for her desire for the card was never suppressed.

So, we should think of the will as a capacity to settled on plans. That gives us a good grip on the idea that an agent can prize *G* and *E*, believe action *A* to be conducive to both and form the intention of doing *A* in order to bring about *G* and not to *E*.

7. Choosing a means

When discussing the second take at Mary's case in section 5, I noted that the most natural interpretation of her case appeals to the distinction between end-decisions and means-decisions. We would like to say that even though she decided to volunteer at Alpha instead of Beta in light of the consideration that doing so would allow her to spend time with her friend, her goal in volunteering was simply to help people in need and not to spend time with her friend. We explain that intuition by saying that by the time the question of whether to volunteer at Alpha or Beta came up, she had already decided to do something to help people in need. That is, by that time she had already made an end-decision to help people. The decision between Alpha and Beta is a mere means-decision – a decision regarding which mean to adopt in order to pursue that pre-established goal. She chose to volunteer at Alpha instead of Beta *as a means* for helping people in need. The reasons that ground a mere means-

decision do not fix the goals at which the action aim. And that is why it can be true both that Mary decided to volunteer at Alpha in light of the consideration that doing so would allow her to spend time with her friend and that she did not volunteered with a view to spending time with her friend. The idea that intentions are plans and that decision making is a matter of developing and settling on a particular plan of action was introduced to account for this reading of the case. Clarifying how it does so will complete my argument.

First, this idea allows us to reject an argument that would show our preferred reading of Mary's case to be impossible:

(a) Trivially, for an action *A* (as performed by *P*) to aim at goal *G* is simply for *P* to choose to perform *A* as a way of pursuing goal *G*;

(b) For *P* to choose to perform *A* is for *P* to move from the consideration of a set of considerations to the intention of doing *A*; therefore, for *P* to choose to perform *A* as a way of pursuing goal *G* is for *P* to form the intention of doing *A* in light of a set of considerations that includes the consideration that *A* is conducive to *G*;

(c) Therefore, if *P* decides to perform action *A* in light of the consideration that action *A* is conducive to *G*, then *A* aims at goal *G*.

Our preferred reading of the second take at Mary's case is incompatible with (c). Given that (a) is a triviality, we should deny (b). In order to do so, we need to be able to explain what it is to choose to perform *A* as a way of pursuing goal *G* without reference to the act of deciding to perform *A* in light of the consideration that *A* is conducive to *G*. The conception of intentions as plans allow us to do just that. With the notion of a plan of action in view, we can deny (b) because we can explain what it is for *P* to choose *A* as a way of pursuing *G* in the following manner: to choose to perform action *A* as a way of pursuing goal *G* is simply to settle on a plan that includes *A* as a strategy to pursue *G*. This opens up the possibility of denying the equivalence between choosing *A* as a way of pursuing *G* and choosing *A* in light of the consideration that *A* is conducive to *G*. It becomes conceivable that an agent could choose to perform *A* as a way of pursuing *G* in light of the consideration that *A* is conducive to *E* – for that can be understood as the claim that the agent settled on the plan {*G*; *A*} in light of the consideration that *A* was conducive to *E*. And that is exactly what seems to be happening in cases such as Mary's.

Second, with the idea that intentions are plans in view we can give a clear account of what is the difference between end-decisions and means-decisions. We want to say that Mary had already decided to volunteer when she contemplated the question of whether to volunteer at Alpha or Beta. In that sense, she had already made an end-decision. What that means is that she had already settled on a *partial* plan whose goal was to help people in need and whose strategy was to do that by volunteering at some organization. In asking whether to volunteer at Alpha or Beta, that is, in trying to reach a means-decision, she was asking how to further develop that plan. To make a means-decision is to settle on a particular way of filling a partial plan to which one is already committed.

I believe there can be little doubt that we can make end-decisions and means-decisions. In particular, I believe there should be no difficulty in holding that Mary had already decided to volunteer in light of the fact that doing so would help people in need when she asked herself whether to volunteer at Alpha or Beta and, thus, that the goal of her action was already fixed. What I hope to have shown is that if we accept this, we have very good reason to take intentions to be plans of action and decision-making to be a matter of gradually developing and settling on a plan of action. If that is correct, in deciding we do not simply choose among alternative actions, but between alternative plans – packages that includes not only actions, but also aims.

So far, I have argued that multiple-incentives cases show that we are not moved by motivational forces, be they the product of desire that assail us or of the reasons for acting we acknowledge. They also show that the goals at which our actions aim are not determined by the reasons for acting we recognize, nor simply by the reasons in light of which we decide. Rather, the goals at which our actions aim are determined by the content of our intentions. In order to account for multiple-incentives cases we should conceive of intentions as plans. The picture that emerges is one according to which we are not beings who are moved by motivational forces. Rather we are beings who adopt certain plans in light of certain circumstances and then execute those plans.

3. The Will as a Capacity for Practical Reasoning

1. Introduction

The goal of this chapter is to provide a positive account of the will. So far, I have argued that we cannot conceive of our behavior as determined by the clash between opposing motivational forces. Rather we should conceive of ourselves as capable of setting intentions that determine the goals at which our actions aim. In the previous chapter I argued that we should not conceive of intentions as normative judgments concerning what we have most reason to do. If that was the case, the goals at which our actions aim would have to be determined by the premises in the reasoning leading to the normative judgment and then we would have trouble making sense of multiple-incentives cases. Rather, we should conceive of intentions as commitments to plans of action. This may suggest that the making of a decision, i.e., the forming of an intention, cannot be the result of a piece of practical reasoning. And that would be a troublesome conclusion, for it seems that we can reach a decision by engaging in practical reasoning.

I will resist the conclusion that we should distinguish between the will and practical reason. If we take practical reason to be our capacity to issue normative judgments concerning what we have most reason to do, then this conclusion would be inevitable. That is shown by some ordinary phenomena, such as what I call cases of *normatively underdetermined decisions* and cases of *akrasia*. These cases seem to suggest that the will is an executive capacity – tasked with converting practical reason’s verdicts into intentions but capable of going astray. This conception of the will, however, obscures the sense in which we form intentions *for reasons*. In a broad enough understanding of reasoning, intentions, despite not being judgments, are the conclusion of some form of reasoning, to the extent that we decide to act in light of certain considerations. What cases of normatively underdetermined decisions and cases of *akrasia* really show, I will argue, is that we should distinguish between what may be called theoretical reasoning concerning our reasons, which concludes in a judgment, and practical reasoning, which concludes in an intention (that is, in the commitment to a plan of action). In section 2, I defend the view that the will should be identified with our capacity for practical reasoning in this sense.

After defending that view and dealing with some objections and concerns, I turn to the relation between the will and desires. If the will is our capacity for practical reasoning,

then it is clear that desires can affect that kind of reasoning in deep ways. In order to complete the image of our motivational capacities, we have to present an account of the relation between the will and desires. In section 3, I suggest that desires conceived as what Scanlon has called “desires in the directed-attention sense” fit well with the conception of the will developed here and allow us to account for the influence of desires over the will. Finally, in section 4, I address the lingering concern that by rejecting the idea of motivational forces we lose grasp of the way in which desires can interfere with our decision making. In particular, the concern is that we can no longer account for the way in which desires press us to act in such a way that self-control becomes a hard task, as illustrated by cases of *akrasia*. In response I argue that talk of motivational forces offers no advantage at this point – in particular, it fails to capture the distinctive way in which desires can press us to act.

2. The Will as the Capacity to Choose for Reasons

It seems natural to conceive of practical reason as the capacity to issue normative judgments about what we should do (or what we have most reason to do, or what is best to do or what is to be done). If that is what we take “practical reason” to refer to, then some ordinary phenomena indicate that the will, conceived of as our capacity for making decisions, is distinct from practical reason.

One such phenomenon is what I will refer to as “cases of normatively underdetermined decisions”. These are cases in which we display an ability to make decisions about how to proceed even when the reasons we acknowledge are (or seem to us to be) equally balanced, incommensurable or inconclusive. Suppose you are offered an attractive job in a distant city. The job you are offered is stimulating and you will earn more than your current salary. However, you will have to move, meaning that you will have less contact with your friends and family. You reflect about what to do but fail to arrive at a definite conclusion – the considerations in favor of each option seem equally compelling to you. Even if as a matter of fact there is a correct judgment about what you should do all things considered, you are unable to reach a verdict.⁴⁰ That means you cannot arrive at a practical judgment of the form “in light of these facts, I should accept/decline the offer”. Now, if the will just is

⁴⁰ It is reasonable to assume that in many cases there is not a single correct judgment to be made. It may be that two or more options are equally well supported by reasons or that one simply cannot weigh the reasons favoring each of the options because they are incommensurable. See Raz (1999, Ch. 3) for a defense of incommensurable reasons. We need not bother to defend this assumption for the problem presented above depends only on the incontestable claim that sometimes agents are unable to reach a definite judgment about what option they should choose.

practical reason conceived of as the capacity to issue judgments of this kind, you should be unable to form a determinate intention in this case. But surely you can do that. You can and must decide what to do, even in the face of uncertainty about what is the best option. And this suggests that our capacity to decide what to do, i.e., to settle on a plan of action, is different from our capacity for making normative judgments.

A natural suggestion about how to cash out this distinction is this: the will stands in relation to practical reason as an executive power stands to a legislative power. The idea is that in cases of underdetermined decision, practical reason lays before the will a set of acceptable options, equally well supported by reasons, and the will then exercises itself in choosing (*at will*) one of the options and forming a corresponding intention. This is what Joseph Raz takes cases of underdetermined decisions to show: the will's typical activity is to choose among the options deemed eligible by reason (Raz, 1999, p.48).⁴¹

This executive capacity seems to take a life of its own in cases of *akrasia*. These are cases in which agents act contrary to their judgment about what they should do.⁴² If one performs an action intentionally and for a reason even though one believes that one has a sufficient reason not to perform it or if one fails to perform an action one believes one should perform, then one is said to be akratic. Akratic actions are all but uncommon: after careful reflection an agent comes to the conclusion that she should stop smoking but she cannot bring herself to do it or an agent sincerely judges that she should not cheat on her partner but decides to do it anyway.⁴³

41 Watson also holds that cases of normatively underdetermined decision show that the will comes into play only "when intention is not completely scripted in advance by reasons" (Watson, 2003, p.182). His comparison of the will with an executive power that "have latitude for its own operation within a legislative framework to which it is subordinate" (Watson, 2003, p.182) suggest that he shares Raz's view. However, the claim that "'Deciding to' typically involves shaping priorities among a structure of reasons and thereby giving certain considerations a special reason-giving force" (Watson, 2003, p.182) suggest that his view is closer to Chang's view discussed in what follows.

42 Cases of *akrasia* are commonly referred to in the literature as cases of weakness of the will. Holton (2009, ch.4) has argued persuasively that weakness of the will is a different phenomenon from *akrasia*. I follow Holton in distinguishing the two phenomena.

43 The existence of cases of clear-eyed *akrasia* (in which the agent really believes, at the time of the action, that she has most reason to do A and yet decides to do something else for a reason she herself takes to be outweighed) is disputed. Those who deny the existence of clear-eyed *akrasia* must redescribe such cases so that the agents in them do not hold the normative judgments they seem to hold (i.e., that contrary to appearances akratic agents do not judge that they should not act as they do) or so that they do not decide to act for an outweighed reason. The former strategy is illustrated by Davidson (1980, p.39) and Buss (1997, p.36). The latter by Hare (1963, p.78-9) and Watson (1977). I do not believe these strategies can be successful. As for the latter suggestion, the claim that akratic agents do not decide to act as they do for a reason flies in the face of ordinary experience. It also fits poorly with the fact that an akratic agent answers for her action in such a way that it is appropriate to ask for the reason for which she acted (in the sense to be discussed in what follows). Finally, once we reject the view that desires are or produce motivational forces that drags us into action, it is hard to make sense of the claim that an agent can be led to act for a desire

The existence of cases of akrasia shows that reaching a normative verdict is not sufficient for an agent to make the corresponding decision. Cases of normatively underdetermined decision show that reaching such a verdict is not necessary. It is reasonable to conclude that we have two separate capacities: a capacity for reaching normative verdicts and a capacity to choose or to decide to act. If we take “practical reason” to refer to our capacity for normative judgment and “the will” to refer to our capacity for decision, we must conclude that the will is not to be identified with practical reason.

I believe this conclusion to be correct, but it has to be taken with a grain of salt. In particular, the tendency here to think of the will as an executive capacity (tasked with executing commands issued in the form of normative judgments but capable of going astray in certain cases) may obscure the sense in which deciding to act is something we do for a reason – and, therefore, something that involves the exercise of our rational capacities. This, I take it, is the main point of Hieronymi’s paper “The Will as Reason” (2009).

In her paper, Hieronymi distinguishes between *theoretical reasoning about practical matters* and *practical reasoning* (2009, p.206-7). These kinds of reasoning are distinguished in terms of the questions they aim at settling. Theoretical reasoning is directed at the question whether p , where p is a proposition. Theoretical reasoning about practical matters, in particular, is directed at the questions such as whether I ought to do A or whether A is what I have most reason to do. To settle one such question is to come to a particular judgment. *Practical reasoning*, in contrast, is directed at the question whether to do something. That kind of reasoning concludes not in a judgment but rather in an intention. Hieronymi argues for the view that the will should be identified with our capacity for practical reasoning in this sense by drawing our attention to the way in which we are liable to answer for our decisions.

This answerability is displayed in the kind of question that can be properly posed to us once we have decided to act in a particular way. As pointed out by Anscombe, intentional actions are distinguished by the fact that they are actions to which a certain sense of the question “Why?” applies (Anscombe, 1963, p.9). That is, when a person does A

without deciding to so act. As for the first strategy of re-description, I believe that what I call cases of *mitigated akrasia* (which I shall discuss in section 4) provide good evidence that akratic agents actually do hold normative judgments against which they act. I will not press these points here, however. My primary goal in this chapter is to explore how the conception of the will at which we arrived in the previous chapter (namely, the conception of the will as a capacity to settle on plans of action in light of certain considerations) accommodates phenomena such as *akrasia*, in which desires seem to play a central role. In what follows, then, I shall simply assume that genuine cases of clear-eyed akrasia exist.

intentionally, we may rightly ask “Why are you doing (or did) A?”. In the same way, if someone intends to do A but has not yet acted on that intention we can ask “Why do you intend to do A?”. The relevant sense of the why-question is that in which an answer to the question gives the agent’s reason for doing A. So, if you accept the job offer, you can be asked why you did. If you reply that you accepted it because it was a more stimulating job or because it would pay you more than your previous position, your answer satisfies the Anscombean why-question because it provides the consideration in light of which you decided to act as you did. In contrast, the question “why did you yawned?” is not an Anscombean why-question. It can properly be answered by “because the person next to me yawned”. But that is not a consideration in light of which you decided to yawn. That shows that the question is not looking for the reason for which you decided to yawn. Yawning is not the kind of thing you do for a reason, because it is not the kind of think you decide to do.

Furthermore, Anscombe’s why-question only applies to the action under a particular description, namely, the description under which the agent intended to perform it. If you are asked “Why are you using the good china?” the question is *refused application* by the answer “I did not know that I was” (Anscombe, 1963, p.11). The question is refused application in this case because it is shown to be inapt. And it is shown to be inapt because it is asking why you decided to use the good china where, as it turns out, you have not decided to use the good china at all (rather you decided to do something else, maybe to pour your tea in any available cup, but happened to pick the good china). The why-question is shown to have an unsatisfied presupposition in this case. Just as the question “how much money do you have in your pocket?” is shown to have an unsatisfied presupposition, and thus refused application, when one answers “I have no pockets”.

These distinctive features of the Anscombean why-question can be accounted for if we take intentional actions to be actions performed for a reason, where an agent performs an action for reason R only if she decides to perform the action in light of the consideration that R. If an agent decides to act in light of a consideration, we can ask what consideration that was. And if an agent does not know that she is doing something, then she cannot have decided to do it, and then the why-question has no application. However, the why-question applies even when one decides to perform an action for no particular reason. If I answer “for no particular reason”, I do not refuse application to the question, in the same way in which I do not refuse application to the question “how much money do you have in your pocket?” if I

answer “none”. This shows that while the Anscombean why-question presupposes that you have made a decision to act, it does not presuppose that you decided in light of a particular consideration.

With that additional point in view, we can account for the distinctive features of the Anscombean why-question if we take it that to decide to A is to settle the practical question whether to do A, where A is the description under which one intends to perform the action (Hieronymi, 2009, p.204). Given that settling questions is the kind of activity done for a reason, if one has settled a practical question of whether to do A, then one can be asked for one’s reasons for doing so – and that is just what the Anscombean why-question does. The question is denied application when the agent is not aware of what she is doing because in this case she did not settle the question of whether to do what she is doing. Since the why-question looks for the consideration in light of which the agent settled that question, it assumes that the agent has settled the question. It is inapt when that supposition is false, i.e., when the agent has not settled the question. It is apt, however, when the agent has settled the question, and thus decided to act, for no particular reason.

I take that to show that we should think of the will as our capacity to decide how to act, where that amounts to settling a practical question of the form “whether to do A?”. Settling questions is the kind of activity done for a reason. According to Hieronymi, that shows that the will is reason “in its practical employment” (2009, p.208). The claim that the will is the capacity for practical reasoning in this sense is, of course, compatible with cases of normatively underdetermined decisions and cases of akrasia. In the former, one fails to settle the theoretical question about what one ought to do (i.e., fails to reach a normative judgment) but manages to settle the question of what to do. In the latter, the agent reaches a normative verdict but, when reasoning practically, employs reasons that are different from the ones she acknowledged when reasoning theoretically, settling the question in a way that is not aligned with her answer to the theoretical question.

To be clear, the claim that the will is a capacity for practical reasoning in this sense amounts to the claim that deciding (i.e., settling on a plan) is the kind of thing we do for a reason. A capacity for reasoning in this sense can be exercised in a completely unreflective manner, without any explicit process of deliberation taking place. The sense of reasoning in play here is, therefore, that which Jonathan Way captures when he claims that “in the most general sense, any psychological process by which we come to form, revise, or sustain an

attitude for a reason – because of or in light of some consideration – counts as reasoning” (Way, 2017, p.252).

Once that point is clear, nothing turns on whether or not one wants to reserve “reasoning” to refer to a more demanding, reflective activity. What matters is that in exercising the will, we form an intention (settle on a plan) in light of some consideration which is, in that case, the reason for which we decided. We can get a better sense of what is involved in this claim by considering what is involved in rejecting it.

A blunt denial of the claim that the will is our capacity for practical reasoning in the sense identified above would amount to the claim that forming an intention is not the kind of thing we do for a reason. According to this view, the will is a merely executive power that makes the transition from normative judgments to intentions. This is a capacity that should, and usually does, follow the dictates of our capacity for normative judgment but which may rebel. But if this transition from judgment to intention is not something we do for a reason, then it is a mere causal process. We simply wait and see whether our normative judgments will lead to an intention – whether or not our executive power will behave as it is supposed to.⁴⁴

The main problem with this conception of the will is that it cannot account for the applicability of the Anscombean why-question to our decisions. To be sure, even if we conceive of the will in this manner it may still make sense to hold a person accountable for her intentions, and, therefore, her actions. A person may be accountable for the misbehavior of her will in the same way she can be responsible for the malfunction of the brakes in her car. It falls to her to make sure that her brakes are working properly and she is negligent if she fails to do what is in her power to ensure that. In the same way, one can be accountable for the operations of one’s will conceived as a part of the mechanism of one’s agency. But this falls short of the manner in which we answer for our decisions. In particular, it makes no sense to ask for *my* reasons for the malfunction of my car’s brakes. It makes sense, however, to ask for my reasons to decide against my normative verdict (Hieronymi, 2009, p.210-1). That is, the Anscombean why-question applies to our decisions, even when the will misbehaves. But if settling on an intention is not something we do for a reason, why the question has application?

One could hold that the Anscombean why-question applies to operations of the will, even though reaching a decision is not the kind of thing we do for a reason, because the

⁴⁴ Silverstein (2007, p.361) makes the same point.

will is supposed to follow our normative verdicts. Given that normative judgments are the kind of thing we hold for a reason, when one's intention follow one's judgment it makes sense to ask for one's reason for having so decided because these are just the reasons in light of which one has come to a normative verdict. Clearly this will not work in cases of akrasia. According to the suggestion we are considering, if in a particular case the will has ignored the agent's best judgment and formed an intention at odds with it, it will be inappropriate to ask the Anscombean why-question with respect to the intention itself or to the action that executes it. After all, according to the suggestion, the question should apply to intentions only when they are the downstream effect of a normative verdict. And that is not the case when someone acts akratically.

Another option is to hold that even though forming an intention is not the kind of thing we do for a reason, there is still a sense in which we can still ask for the agent's *motivating reason* for acting as she did. In a case of *akrasia*, the agent's motivating reason, in this sense, cannot be the reasons in light of which she reached her normative verdict, because she did not act in accordance with that judgment. It also cannot be the considerations in light of which she settled on her intention, because we are assuming that forming intentions is not the kind of thing one does for a reason. The only option left is to hold that an agent's motivating reason is something that explains her action. It is often assumed that we can explain an action in this way by pointing to a desire or pro-attitude of the agent in combination with an appropriate means-end belief. One may argue that even if an agent has not settled on her intention for a reason, there can still be a reason for her acting as she did, namely, the desire-belief pair that explains it. The problem with this proposal, however, is that the agent's *motivating reason* thus conceived is not, by hypothesis, something in light of which she settled the practical question of whether to act as she did. By providing the agent's motivating reason we are simply pointing out the psychological states that produced her action. These states provide the *agent's* reason for acting only in the sense that they are states of the agent – features of her psychology. In the same sense of “reason”, I could point out someone's reason for, say, blushing: “she blushed because she thought of an embarrassing moment”. But blushing is not the kind of thing to which the Anscombean why-question applies. Indeed, the same explanation can be provided in a case in which the agent is not aware of blushing, and, as already remarked upon, the Anscombean why-question does not apply when the agent is not aware of what she is doing. This goes to show that the fact that a

motivating reason in this sense can be provided is not enough to explain why the Anscombean question applies.⁴⁵

We have very good reason, then, to hold on to the view that forming an intention or settling on a plan is the kind of thing we do for a reason. But one may still feel inclined to resist the claim that the will is a capacity for practical *reasoning*. In particular, one may hold that there is no sense that can be given to the claim that the will is a capacity for reasoning where the relevant kind of reasoning is different from theoretical reasoning about practical matters. Someone who presses this objection would claim, first, that the kind of reasoning that settles the practical question of whether to do A is, in many cases, theoretical reasoning concerning practical matters and, second, that when that kind of reasoning fails to settle the practical question what settles it is not another distinctive kind of reasoning, but rather an irreducible act of the will.⁴⁶

As for the first claim, one can draw our attention to the fact that in many cases there seems to be no gap between the normative judgment that is the conclusion of an episode of theoretical reasoning concerning practical matters and the corresponding intention. It seems that in these cases one settles the practical question of what to do *in* settling the question of what one ought to do. There is no need for an additional step in reasoning to move from the judgment to the intention. So, there is no role for the will to play in this case.

But the claim that there is no gap between judgment and intention in these cases can be interpreted in more than one way. One way to understand it is as the claim that in these cases the normative judgment one reaches as a result of theoretical reasoning simply produces

45 Davidson (1980, p.4) e Smith (1994, p.131) use “primary reason” and “motivating reason”, respectively, to refer to a desire-belief pair that explain an action. But they do not hold, in opposition to the view discussed in this paragraph, that intentions are not the kind of thing we adopt for a reason. According to Davidson, a primary reason explains an action by rationalizing it. It rationalizes the action by showing what the agent saw in the action. And in doing so, Davidson holds, it allows us to see the reasons in light of which the agent decided to act (1980, p.98-9). Indeed, it seems that Davidson position is close to one defended in this chapter. He distinguishes our capacity to reach judgments of the form “all things considered, doing A is desirable (or more desirable than doing B)” from our capacity to form intentions. But he claims that intentions simply are all-out judgments of the form “doing A is desirable (or more desirable than B)”. To decide to act is to reach one such all-out judgment. What leads him to this view is the idea that an intention or decision can be the conclusion of a piece of practical reasoning (Davidson, 1980, p.96). An akratic agent, on his view, is one who reaches an all-out judgment of the form “doing A is more desirable than doing B” while judging that “all things considered, doing B is more desirable than doing A” (Davidson, 1980, p.39-40). The view defended here differs from Davidson’s only in that it holds that intentions, despite being the conclusions of pieces of practical reasoning, are not to be identified with normative judgments (for the reasons presented in the previous chapter). This allow us to avoid the need to claim that there is a sense in which the akratic agent judges her weak-willed action to be preferable to the course of action that she takes to be better supported by reasons.

46 This two-pronged objection is pressed by Carey (2019, p.4-5).

the intention. But then the connection between judgment and intention is merely causal, and this will lead to the problems discussed above. On another way of understanding the claim, it means simply that the agent settled the theoretical question in light of certain considerations and simultaneously settled the practical question in light of the same considerations. In this case, the agent answers both question in parallel. This is perfectly compatible with the claim that the will is our capacity for practical reasoning. It only shows that in a set of familiar cases practical reasoning and theoretical reasoning about practical matters mirror each other – as we should expect from beings that are reasonably rational most of the time.

In pressing the second claim, one would argue that in cases in which there is a gap between judgment and intention, what is needed to close it is not reasoning. One could think that when one has already reached a verdict concerning the balance of reasons, there is nothing left to reason about. Again, some care is needed here. Sure, one has already reached a normative judgment, but there is still need for reasoning in the minimal sense identified above. One still needs to settle on an intention in light of certain considerations. Of course, that one has failed to settle the practical question despite having settled the theoretical question regarding the practical issue at hand shows one not to be fully rational. But that does not change the fact that one still has to settle the practical question, and that doing so requires reasoning. One may call the settling of this question the making of a *choice* or *decision*, but as long as this is understood as something that we do for a reason, the disagreement here is merely verbal.

Here our interlocutor is likely to take issue with our use of “reasoning”. We have been talking of reasoning as the settling of a question in light of certain considerations. That is something that we can do in a completely unreflective way. It is also something that we can do momentarily. We can distinguish reasoning in this sense from deliberation, understood as the conscious, temporally extended cogitation or process whose end-point is the settling of a question.⁴⁷ So, the objection we are dealing with can be understood in another way. The claim that, in many cases, the kind of reasoning that settles the practical question of whether to do A is theoretical reasoning concerning practical matters can be understood as the claim that the deliberation leading to the settling of a practical question usually takes the form of the weighing of reasons in order to determine what we have most reason to do. And the claim that

⁴⁷ This distinction between reasoning and deliberation is the same distinction Harman draws between the revisions of one views or intentions and the process of reflection that leads to this revision. See Harman (1986, p.2).

after the theoretical question about the balance of reasons has been settled there is no more room for reasoning may be understood as the claim that the weighing of reasons is the only thing we can do in trying to settle a practical question and, therefore, that after the theoretical question has been settled there is no thinking to be done – one simply has to decide.

Let's start by discussing the latter claim. If we take it a face value, it is implausibly strong. That we do engage in purely practical deliberation, when we have already settled the relevant theoretical questions, and that at least sometimes we do it rationally, should not be controversial. Consider the job offer case again and suppose you came to the conclusion that you have sufficient reason to go either way, accepting or rejecting the position. Now you have to decide. But this is a big decision. Not one to be made on a whim. One has to think about it, and think carefully. But given that one has already settled the relevant theoretical questions regarding one's reasons, what can this thinking consist in? It seems that it can only be practical deliberation – the kind of thinking directed at settling a practical question.

The only way to resist this conclusion is to hold that even in one such case the agent is still weighing reason, trying to reach a normative judgment that will settle the issue. This is Ruth Chang's view. She defends what she calls "hybrid voluntarism", the view according to which some of our reasons are created by acts of the will. According to Chang, when our reasons fail to fully determine what we have more reason to do we may "will some consideration to be a reason" (Chang, 2013, p.180). When our reasons run out and fail to determine what is the single best course of action available to us, we can create "voluntarist reasons that may make it the case that for that agent, she has most all-things-considered reasons to choose one way rather than another" (Chang, 2013, p.181). If Chang is correct, then the process by which we arrive at normative judgments has two stages: first we consider our given reasons (i.e., reasons we have not created by an exercise of the will); if we conclude that in light of these reasons we have multiple equally eligible options, then we move on to the second stage where we engage in deliberation in order to create reasons that tip the balance of reasons one way or the other (Chang, 2009, p.255-6). In this latter stage we have two options: we can either take as a reason a consideration that was irrelevant to our choice and, therefore, not considered in the first stage of deliberation or take as a reason a consideration that was taken in account in the first stage of deliberation and make a voluntarist reason out of it, so that the same consideration provides two different reasons to perform a particular action (Chang, 2009, p.256-7). The deliberation that follows the

conclusion that our reasons are equally balanced, incommensurable or inconclusive, then, still aims at settling the theoretical question of what we have most reason to do, although it does that by coming up with new reasons.

I take this view to be highly problematic. The claim that we take as reasons considerations that are, by our own lights, irrelevant to our decision seems to me to directly falsified by the common experience of deliberation. And the idea that the same consideration can provide two different reasons leads to highly counterintuitive results. One would have to take seriously claims like “I had two reasons to take the job: that I would get a raise *and* that I would get a raise”. Furthermore, if one does change the balance of reasons in favor of, say, option A, that means that choosing B would be irrational. But that seems wrong. Sure, if one has settled on a plan it may be irrational to simply change one’s mind out of the blue, for that defeats the practical purpose of plans. And if one is already engaged in executing that plan, expending time and resources, as well as restricting one’s choices in light of the plan, these facts may provide reason not to change one’s mind now. Especially when it comes to important choices, as one’s life becomes more and more entangled with the project one chose, facts about how one’s life is organized will provide all kinds of reasons to keep on track. But according to Chang, it becomes irrational to choose differently at the moment one has settled on a decision. And that is counterintuitive: from the point of view of the agent, even after making the decision, choosing the alternative may seem like a perfectly reasonable option.

Even if we ignore these difficulties, however, Chang’s approach does not provide a genuine alternative to the claim that we sometimes engage in purely practical deliberation. Surely, in a case in which the reasons in favor of doing A and B are equally balanced, one would will an irrelevant consideration or a consideration one has already taken into account to be a further reason to do A (thus tipping the scales in favor of A) only if one has already decided to go for A. It is not as if one decides to take an irrelevant consideration as a further reason to A and then goes through the process of weighing reasons again only to discover, to one’s surprise, that one has most reason to do A and then decides accordingly. Rather, in willing an irrelevant consideration to be a further reason to A one is *ipso facto* deciding to A. It can only be because we are going with the plan of accepting the job offer that we take an irrelevant conclusion to be a further reason to accept the job, not the other way around. What we are trying to do, in the first place, is to reach a decision, knowing (or taking) the reasons in favor of the alternatives to be balanced or inconclusive. The kind of thinking we are engaged

in is not at all concerned with weighing reasons in order to reach a normative verdict (even if, as Chang insists, after reaching a decision we revise our judgment). Rather, we are engaged in the kind of thinking that concludes in the formation of an intention in light of certain considerations.

So, there is such a thing as purely practical deliberation. Surely, if one engages in this kind of thinking after having reach a decisive normative judgment, that indicates a flaw in rationality – for one is failing, even if momentarily, to decide according to one's normative verdict. In this sense, this practical deliberation cannot be an exercise of *Reason*. But pure practical deliberation also can, and must, take place when we fail to reach a normative verdict. Being unable to decide in a situation in which reasons are equally balanced can also be a display of irrationality. And, in cases such as the job offer example, one can also display some degree of irrationality by making a decision without thinking carefully about the practical question one faces. Pure practical deliberation is not always at odds with rationality. Sometimes rationality requires it.

Enough about pure practical deliberation. What about the other claim, that practical deliberation usually takes the form of theoretical deliberation about what we have most reason to do? Notice first that saying that practical deliberation *takes the form* of theoretical deliberation can be misleading. As long as we are weighing reasons with an eye to settling the practical question of what to do, the kind of deliberation we are engaged counts as practical deliberation. The fact is that theoretical deliberation is usually nested inside practical deliberation. That is to say, it seems that when we explicit approach the practical question of what to do in a reflexive manner we usually do so by asking ourselves what we should do. And, at least in a familiar set of cases, concluding that piece of theoretical deliberation also brings the practical deliberation to an end. Indeed, that seems to be the normal form of explicit, conscious practical deliberation, one from which we depart only in the special cases in which we cannot seem to reach a normative verdict or when we succumb to temptation. The problem here is to explain why that is so.

Part of the answer has to do with the structure of practical deliberation. Practical deliberation aims at settling the question of whether to do A. If one explicitly and reflexively entertains that question, the salient considerations to which one's attention will be directed are those that bear on the question, namely, reasons for and against doing A. These considerations strike a rational agent as counting in favor or against doing A and as having a certain relative

weight (although it may be hard to accurately estimate that weight at first). But that means that explicitly considering the practical questions prompts the agent to consider the theoretical question about how the reasons at play balance out.

So, practical deliberation prompts theoretical deliberation regarding the practical matter at hand. That seems reasonable enough. But it is not enough to account for the way in which theoretical deliberation is usually nested in practical deliberation. We do not simply engage in both kinds of deliberation at the same time but rather approach the practical question by engaging in theoretical deliberation. In order to explain that we have to explain not only why we usually engage in theoretical deliberation when we engage in explicit practical deliberation but also why the theoretical deliberation takes over, so to say, in such a way that we suspend the settling of the practical question until the theoretical deliberation is concluded and that concluding the theoretical deliberation settles the practical question as well.

Here we should follow Scanlon in holding that the explanation appeals to the idea of a rational agent. A rational agent is one who is capable of thinking about reasons for action and reaching normative judgments about what one has most reason to do, and one whose judgments about reasons make a difference to how she acts. In particular, a perfectly rational agent always acts in accordance with her own judgment about what she most reason to do (Scanlon, 2014, p.54-55). To the extent that we are reasonably rational, then, we can expect that settling the theoretical question about our reasons will lead us to settle the practical question as well, and in accordance with our normative judgment.⁴⁸

48 One could rightly put pressure on me at this point to explain more clearly the relation between normative judgment and practical reasoning. Surely, it is no mere accident that rational agents always decide to act according to their normative judgments. There must be some kind of internal connection between practical reasoning and theoretical reasoning regarding practical matters. In divorcing the two kinds of reasoning, have I not severed that connection? I do not think so. We can understand the connection between the two kinds of reasoning if we take theoretical reasoning about practical matters to be reasoning *about practical reasoning*. According to this view, practical reasoning is the most basic kind of reasoning and theoretical reasoning is to be understood by reference to it. When we think about what we should do or about what we have most reason to do in circumstances C we are not engaging in practical reasoning but rather thinking about what would be the *correct or sound* course of practical reasoning in C (this view was also defended by Silverstein, 2017, p.373). For instance, in trying to determine if the fact that the job I am offered will pay much more than my current job provides a sufficient reason for me to take the job, I am trying to determine whether the course of practical reasoning that moves from that consideration to the decision to accept the offer is a sound piece of practical reasoning. To judge that I should take the job in light of the fact that it pays more is to say that that piece of practical reasoning is sound. According to this view, the fact that we engage in theoretical reasoning about practical matters is a reflection of the fact that we are capable of reflecting about our decisions and wondering whether we should have decided differently in a certain situation. If we understand theoretical reasoning about practical matter in this way, then it is no accident that rational beings decide to act in accordance with their normative judgments. For rational beings are coherent, and there is a kind of incoherence in thinking that a certain course of practical reasoning is the only sound course of

This provides a partial explanation of why practical and theoretical deliberation usually combine as they do. Practical deliberation prompts theoretical deliberation, and to the extent we are rational, the conclusion of the theoretical deliberation in a normative judgment should lead us to settle the practical question, thus concluding the practical deliberation as well. But this explanation is still incomplete. We still need an explanation of why in normal cases we wait for the conclusion of the theoretical deliberation before deciding how to act.

Consider once again the job offer example. And suppose that the agent in this example tells us something like “I’m still thinking about whether or not I should accept the job offer, but I already decided to accept it”. This is a very odd claim. We expect one to suspend one’s decision as long as one is still engaged in theoretical deliberation concerning what one should do. That can be easily explained if we assume that practical deliberation simply consists in theoretical deliberation about practical matters. In light of that assumption, practical deliberation can be brought to a conclusion and a decision made only if the theoretical deliberation is brought to a conclusion, since they are the same. But we have rejected the view that practical deliberation can be identified with theoretical deliberation about practical matters. Why then do we expect the former to be concluded only when the latter comes to a conclusion?

Again, I believe the answer is that we expect each other to demonstrate a certain degree of rationality. The expectation that we suspend our decision until theoretical deliberation has been completed, however, cannot be explained by pointing to the fact that we expect agents to decide in accordance with their normative judgments. The anomaly here is not that of an agent deciding against her best judgment, but of an agent deciding before making up her mind about what is the best option. Why do we expect agents to make up their minds about the balance of reasons before deciding? We are comfortable with the idea that a perfectly rational agent can decide to act in an unreflective manner, without entertaining any normative judgment. So, our problem is not with the idea of unreflective choice. And we are

practical reasoning in circumstances C and engaging in a different course of practical reasoning (one which is defective by one’s own lights). This understanding of normative judgments also explain why they have, so to speak, normative authority over practical reasoning. An agent should regard her practical reasoning as defective if it does not conform to her normative judgments, simply because these judgments are judgments about correct or sound practical reasoning. Furthermore, this understanding of normative judgments is compatible with a number of different meta-ethical theories. If there are facts (natural or non-natural) about what is the correct course of practical reasoning in certain circumstances, as cognitivists hold, then our theoretical reasoning about practical matters aims at capturing those facts. If non-cognitivism is correct, then we should take our normative judgments to be mere expression of our preferences regarding practical reasoning or prescriptions regarding how we should engage in practical reasoning. Be that as it may, the relation between practical reasoning and theoretical reasoning about practical matters remains the same.

also comfortable with the idea that a perfectly rational agent may find herself in a position in which she has to make a decision before reaching a normative verdict, perhaps because the decision can no longer be postpone. So, we do not simply expect agents to always make their mind about what they have most reason to do before acting, rather we expect agents to make up their mind about this theoretical question before deciding *when they are already explicitly considering the question whether to act and there is no pressure to make a decision at this moment*. We can account for that expectation as derived from the expectation that agents display a certain degree of rationality if we assume that there are rules of prudence embedded in our conception of rational agency. Suppose the agent in the job offer example decided to accept the job before reaching a verdict about how the balance of reasons pro and against accepting the job. If she has genuinely decided to accept the offer, then she is committed to performing a number of other actions. She has *ipso facto* decided to quit her current job, to move to another city, etc. If she were to do any of these things and later were to come to the conclusion that what she had most reason to do was to refuse the offer (as she might if her theoretical deliberation extended for a long time) she would regret her action. This shows that her decision would have been rash and reckless. If she does not regret her action, if it turns out that was, by her own lights, the right decision, that is a matter of luck. Her decision is no less reckless for that reason. A prudent agent does not make rash and reckless decisions like this. If there is no pressure to make a decision right now, the prudent agent takes her time. That is, she suspends the decision until a normative verdict has been reached or until the decision can no longer be postponed. And, I'm suggesting, a perfectly rational being is prudent in this minimal sense. To the extent, then, that we expect agents to be reasonably rational we expect them not to decide before reaching a normative verdict when they have explicitly brought up the question of whether to A.

What emerges from this discussion is the following explanation of why explicit practical deliberation usually has theoretical deliberation nested in it despite the fact that practical deliberation does not necessarily involves theoretical deliberation. On account of its nature and subject, explicit practical deliberation prompts theoretical deliberation. To the extent that one is prudent in the minimal sense we expect from rational beings, one will (when the circumstances allow it) suspend decision until a normative verdict has been reached. To the extent that one is rational, if one is successful in reaching a definite normative verdict, one will decide in accordance with that verdict, thus putting an end to practical deliberation.

Practical deliberation takes this standard form when agents proceed as we expect rational agents to proceed. But these expectations may not be met. If one is impetuous, one may decide to act without taking enough time to think about whether or not to act. If one is reckless, one may engage in theoretical deliberation but decide before arriving at a final normative verdict. And if one is akratic, one may continue to deliberate practically even if one has settled on a definite normative verdict.

The fact that theoretical deliberation is usually nested in practical deliberation in the way described above is the main source of the intuition that practical reasoning must consist in the activity of weighing reasons with a view to reaching a normative judgment about what one has most reason to do. I hope to have done enough to undermine the strength of that intuition. I have suggested, instead, that we take practical reasoning to consist in the activity of settling on an intention in light of a consideration and practical deliberation to be the thinking that concludes in that kind of reasoning. The will, I am contending, just is our capacity for practical reasoning in this sense.

3. Desires and the Will

I have argued that we should conceive of the will as our capacity for practical reasoning. But there is no doubt that desires have an influence over the will. One way in which desires can be relevant to our decisions is as a source of suggestions about what to do. If one desires something one is likely to ask oneself whether to act so as to get what one wants. Another way desires can be relevant to our decisions is for us to take the fact that we have a particular desire or the fact that it affect us somehow as considerations that are relevant to our decision. That is what happens, for instance, when one decides to satisfy a distracting desire so as to be able to focus on one's work. But surely desires can have a more profound effect on our decision making. That is illustrated by the role desires can and often play in cases of akrasia. Often it is because one desires something, and because that desire is particularly intense, that one lingers on in practical deliberation even though one is convinced one should not indulge in the desire and eventually settles on an akratic intention in light of what is, by one's own lights, an outweighed reason. Desires, to put it simply, can make it hard to decide according to our normative verdicts. In order to complete the volitionalist account of motivation we have been developing, we need to provide an account of how desires can influence the will and what their role in the production of decisions is.

It is important to provide one such account because one may take the very fact that desires can tempt us and interfere with our decision making to lend support to the idea that the operations of the will are determined by a dispute between motivational forces. After all, talk of motivational forces seem to give us a reasonable grasp of what this influence consists on. First, it allows for a straightforward explanation of why an agent decides against her best judgment in a case of *akrasia*. Namely, because the desire that was tempting her produced a stronger motivational force than the reasons she recognized to act differently. Second, if we conceive of the operations of the will as determined by motivational forces in dispute, it may seem that we can account for the observation that even when one manages to resist a temptation, it is hard to do so. One may hold that even when the motivational force produced by the desire is stronger than opposing forces, resistance may still be possible. We may conceive of the agent as exercising a capacity of self-control (call it strength of the will) that counteracts the motivational pull of the desire. The stronger the desire (that is, the stronger the motivational force it produces), the harder it will be to resist. The reason for that is that a more intense and effortful deployment of our capacity for self-control is necessary in order to counteract a stronger motivational force.

If we reject the idea that a desire's influence over us takes the form of a motivational force, pushing us to act, how are we to understand that influence? First, we must ask what we mean by "desire" here. Philosophers use the word in more than one sense and, in some of these, desires are not the kind of thing that can have an influence over the will and can tempt us. In one sense of "desire", the term refers to any pro-attitude. A pro-attitude is any attitude that in combination with a means-end belief can rationalize an action by showing what the agent saw in the action. In this sense, the class of desires is extremely broad. It can include such ordinary desires as hunger and thirst, but also far more sophisticated attitudes as the love for one's children, a taste for company, habits, a sense of duty or loyalty, moral convictions and values, normative judgments and intentions. This use of "desire" fails to distinguish between states that have an influence on the will (such as, say, hunger) and states that are the product of the exercise of the will (an intention). In Thomas Nagel's terminology, this sense of "desire" does not distinguish between unmotivated desires (which are states in which we simply find ourselves, regardless of deliberation) and motivated desires (which are the product of our agential capacities).

In the sense that interest us here, desires are the kind of state that influence the

will. But the kind of state I have in mind has another mark as well: these are states with regard to which we are *passive* in an intuitive sense. In contrast with states such as intentions or normative judgments, these states “come upon us” or “assail us”, much like a perceptual experience. They are also not directly responsive to our normative judgments or decisions. We can have desires we judge we should not have and fail to have desires we judge we should have. A desire may persist even if we decide not act on it and we may fail to have a desire to act as we have decided to act. That is to say, we cannot change our desires at will. On account of this passiveness, we do not answer for our desires as we answer for our intentions or judgments. Desires, in this sense, are not the kind of thing we have for a reason.

Motivational pulls and pushes fit the bill. But we have rejected the idea that desires are or produce this kind of motivational force. What alternatives are there? One popular view holds that desires are dispositions to act in a particular way.⁴⁹ But there is a problem with this suggestion as well. A disposition to act in a particular way can be ascribed to an agent whenever she acts in that way. In that sense, whenever one acts, one has a desire to so act. But it is not true that whenever an agent acts in a particular way she had a desire in the sense that concerns us here – that is, a desire of the kind that could tempt her to act against her best judgment.

This distinction is helpfully presented by Schueler (1995, p.29). In one sense of “desire”, whenever you act intentionally it can be said that you have a desire to do what you did. In his sense of “desire” it is impossible to intentionally do something one does not want or desire to do. But the reason why that is so is that to ascribe a desire in this sense to an agent is simply to register the fact that the action was intentional and as such aimed at a goal to which the agent was not indifferent. In a second, ordinary sense of “desire”, it is perfectly possible to do something one does not desire to do. One can attend a meeting even though one has no desire to do so. One would much rather stay at home, for instance. But in the first sense of “desire”, you had a desire to go to the meeting. Nothing forced you to go to, you did that on your own accord.

The claim that desires are dispositions fits very well with the first sense of “desire”. If one acted intentionally, then it is reasonable to claim that one was disposed to so act (in particular, one was disposed to decide to so act in light of certain considerations). But the sense of “desire” in which we are interested here is the second. Call this the *substantive*

⁴⁹ See, for instance, Smith (2012) or Hyman (2015, ch.5).

sense of desire, for this is the sense in which it makes sense to ask whether or not an agent desired to act as she did. The agent that goes to a meeting even though she much rather stay at home has, by hypothesis, a disposition to so act, but she does not have a desire in the substantive sense. That shows that having such a desire cannot consist simply in having a disposition.

An alternative is to conceive of substantive desires as what Scanlon has called “desires in the directed-attention sense”. Scanlon holds that normative judgments can on their own be a source of motivation and explain action. We can see that they do so by considering the idea of a rational agent. “A rational agent”, Scanlon tells us:

is, first, one that is capable of thinking about the reasons for certain actions or attitudes, and for reaching conclusions about which of these are good reasons. Second, a being is a rational agent only if the judgments that it makes about reasons make a difference to the actions and attitudes that it proceeds to have. A perfectly rational agent would always have attitudes and perform the actions that are appropriate according to the judgments about reasons that he or she accepts. (Scanlon, 2014, p.54).

In particular, a perfectly rational being would always form an intention that corresponds to her normative judgment about what she should do. And intentions will, all things equal, lead to action. To the extent that an agent is rational, then, we expect her to act in accordance with her normative judgments. And that is why pointing to the agent’s normative judgment or belief can make sense of her action and explain it - “because the action is what one would expect of a rational agent who accepted that judgment” (Scanlon, 2014, p.55). There is no need here to postulate a further source of motivation that ensures that the agent conforms to her judgment. Surely one can fail to be moved by one’s normative judgments, but that is not due to a lack of an independent form of motivation, but rather to a failure in the processes by which a rational agent moves from the consideration of reasons to an intention (a failure, therefore, that amounts to failure in rationality).

Even when one acts or decides to act in a way that reflects a failure in rationality (as in a case of akrasia), Scanlon argues that there is not an independent form of motivation that accounts for the action:

On the contrary, when I examine these cases it seems to me that in all of them the only source of motivation lies in my taking certain considerations—such as the pleasures of drinking, of eating, of hearing from a friend—as reasons. The strength of this motivation varies depending on what happens—for example, on the degree to which I attend to a given consideration, focus on it, and ignore others—but these reasons remain the only motivating factors. (Scanlon, 1998, p.35)

We cannot take very seriously the talk of strength of motivation, on pain of falling back of a version of the hydraulic model. Rather, the claim that taking a consideration as a reason to act resulted in a stronger motivation than the motivation associated with what the agent took to be better reasons to act differently is simply a way in which to register the fact that the agent decided to act in light of what she herself took to be outweighed reasons. What Scanlon is proposing in these passages is then, in my terms, that motivation always takes the form of an exercise of the will, in which one forms an intention in light of considerations one sees as bearing on the question of what to do. To say that the strength of the motivation derived from these considerations varies, for instance, to the extent that one's attention is focused on a particular consideration is simply to say that focusing one's attention on a consideration is the kind of think that can lead one to decide in light of it even when one knows it to be outweighed. That is the only content that can be given to talk of motivational strength once we moved away from the hydraulic model.

But what about the motivational relevance of desires in the substantive sense? Scanlon proposes that "desiring something involves having a tendency to see something good or desirable about it" (Scanlon, 1998, p.38). But that cannot be all. In one sense, I can have a tendency to see something good about drinking a glass of foul-tasting medicine when I am sick and know that doing so would relieve my symptoms. But that does not mean that, in light of that knowledge, as soon as I get sick I develop a substantive desire to drink foul-tasting medicine. That is exactly the kind of thing we do without wanting to do. This is what leads Scanlon to the idea of desire in the directed-attention sense. An agent has a desire in the directed-attention sense that p "if the thought of p keeps occurring to him or her in a favorable light, that is to say, if the person's attention is directed insistently toward considerations that present themselves as counting in favor of p" (Scanlon, 1998, p.39).

Scanlon emphasizes that for a consideration to *present* itself as a reason is not the same as judging it to be a reason. "One can have a strong and recurrent tendency to see something as a reason for acting (under one's present circumstances) even though one's firm considered opinion is that it is not (under the circumstances) such a reason" (Scanlon, 1998, p.40). He distinguishes, therefore, between our ability to make judgments about our reasons and our ability to see a consideration as a reason and holds that they can come apart. We can continue to see something as a reason even when we are convinced it is not (just as a stick in the water continues to seem crooked even though I am perfectly aware that it is not). This

move is necessary to account for the passiveness of desire, or as he puts it, to capture the familiar idea “that desires are unreflective elements in our practical thinking—that they ‘assail us’ unbidden and that they can conflict with our considered judgment of what we have reason to do” (Scanlon, 1998, p.39).

To have a desire in the directed-attention sense is to have one’s attention drawn, in an insistent and recurrent manner, to considerations that seem to the agent to count in favor of acting in a particular way or having a particular attitude. And a consideration can seem to be a reason even when one is convinced it is not.⁵⁰ This conception of desire seems to fit well with the phenomenology of the influence of desire on our decision making. It also accounts for the passiveness of desires: the fact that they assail us, that they can be unresponsive to our judgments and that we do not answer for them. Finally, desires in this sense are independent enough of the will to be the kind of state that can influence it, sometimes in unduly ways. They fit very well, therefore, with the volitionalist account of motivation. But how exactly do they influence the will? How can a desire in the directed-attention sense lead an agent to decide against her normative judgment about what she has most reason to do? Clearly, the stronger a desire the more likely we are, other things equal, to decide to indulge in it even though we judge we should not. In order to explain the influence of desires over the will, then, we need a conception of the strength of desires.

A desire in the directed-attention sense manifests itself not as a motivational force but rather in the direction of one’s attention to the object of desire and, in particular, to those features of the object that are practically salient and seem to the agent to count in favor of acting somehow. The strength of a desire in this sense is a matter of the intensity with which one’s attention is directed to these features when one is under the grip of the desire. A particular strong desire is one which tends to dominate the agent’s conscious experience. One’s thoughts keep coming back to the object of desire. One can’t help but to think about it. One finds oneself constantly imagining what it would be like to satisfy the desire, anticipating that experience, even fantasizing or daydreaming about it. One comes up with possible plans

50 Similar views of the nature of desires, that also rely on an analogy with perception were defended by Stampe (1987, p.326), Wallace (1999, p.641-2), Oddie (2005, p.42) and Tenenbaum (2007, p.39). Schapiro (2009) and Schafer (2013) agree in conceiving of desires as normatively charged modes of presentation of certain objects or states of affairs but hold that the content of desires is not properly articulated using normative concepts. I will ignore this criticism here for even if we accept Schapiro’s or Schafer’s conception of desire, the explanation of the influence of desire over the will should take the same form as the one presented below. For our purposes, then, exactly how to understand the nature of desires once we agree that they are normatively charged modes of presentation of their objects is irrelevant.

for satisfying the desire, exploring different options, anticipating obstacles, etc. When the desire is very strong, its urgency may manifest in the form of dysphoric sensations (such as pain or discomfort) or emotions like anxiety and excitement.

This gives us what Wallace has called a “phenomenological conception” of the strength of a desire. A desire is not strong in virtue of the motivational force associated with it but rather on account of the “way things seem experientially to the person who is in their grip” (Wallace, 1999, p.643). It is easy to see how being in the grip of a strong desire in this sense can interfere with one’s theoretical deliberation about what one has most reason to do. Having one’s attention constantly directed to what seem to be considerations that count in favor of indulging in the desire may make it difficult to think clearly. That is likely to encourage the agent to overestimate the weight of these considerations and to ignore or discount considerations that provide reasons against indulging in the desire. One’s ability to think clearly will be further impaired if the desire manifest itself in the form of dysphoric sensations and turbulent emotions. And, finally, the highly vivid presentation of the considerations speaking in favor of indulging combined with the discomfort involved in refusing to satisfy the desire is likely to prompt the agent to engage in rationalization, so as to convince herself that overall there is no reason not to indulge.

But the desire must also be able to affect the agent’s practical deliberation even when she manages to reach and keep firmly in mind a normative judgment to the effect that she should not act on the desire. How can we understand this kind of influence? As already noted above, Scanlon speaks as if by taking a consideration to be a reason an agent would produce in herself a motivational force and that the strength of this force can vary on the degree to which one focus on the consideration and ignores others (Scanlon, 1998, p.35). I suggested that this talk of motivational forces cannot be taken seriously in light of the argument of the previous chapters. It is best to read Scanlon as claiming simply that focusing one’s attention on a consideration that counts in favor of acting in a particular way is the kind of thing that can lead one to decide to act in light of it even when one knows that one has more reason to act in another way.

To insist in this idea is to take the motivational effect of temptation as basic. Temptation is the psychological condition that facilitates the decision to act against one’s best judgment by directing one’s thoughts onto what seems to be attractive features of the action. To be under the influence of a strong desire in the directed-attention sense is to be subjected to

temptation. The claim here is that the motivational effect of temptation is not to be explained in terms of some more fundamental features of our motivational psychology, such as motivational forces, but rather to be taken as a basic motivational phenomenon. It is simply a fact that when a strong desire takes over the course of one's thoughts and presents one with highly vivid alternatives for action under an attractive light, it makes it difficult for one to decide according to one's normative verdict. Holding on to one's normative conviction may require concentration, effort, strength of will. For someone committed to the phenomenological conception of the strength of desires, the effort in question is the effort of deciding according to one's judgment (or holding on to a previous decision) in a situation in which an alternative keeps appearing to one under a highly attractive light. What we call strength of the will is the capacity to uphold this effort. It is a fact that this is indeed an *effort*, that it is something we do, that is hard and that consumes energy. But that is not something we read off the notion of desire. It is something we find out in experience.

4. Akrasia and Motivational Forces

I will now consider an objection to the account I have provided of the role of desires in our decisions. We know that desires tempt us, that it is hard to resist temptation and that the stronger the desire the harder it will be – more strength of the will, concentration and effort will be necessary to resist. All of these claims are compatible with the phenomenological conception of the strength of desires. But the phenomenological conception does not explain them. Or rather, the explanation provided runs out quickly – that is how the focusing of attention on attractive features of alternatives affect us and we are all familiar with this phenomenon. This may seem as a short coming of the volitionalist model. One may feel that a deeper explanation was in order. And it may seem as if a view that postulates motivational forces has a decisive advantage at this point. In particular, one may think that the notion of motivational forces can give us some account of why the focusing of attention on attractive features of the object of desire makes it harder to decide according to one's best judgment. When we focus our attention on an attractive feature of the object and are tempted by it, we feel pressured to act by our desire. That pressure, our interlocutor would hold, is best understood as a motivational force acting upon us. If we are convinced that we should not act on the desire, we will resist it. That resistance consists in opposing the motivational force of the desire. The difficulty involved is the difficulty of resisting a motivational force. The stronger the force the harder we have to push back. And the stronger

we have to push back the harder it is, the more effort is required.

I will argue, however, that the view that the notion of motivational forces gives us a deeper grasp of how desires tempt us is an illusion. The way in which a desire can pressure us to act is not properly captured by the notion of motivational force, for that pressure can have effects upon our decision making that are not the kind of effect a motivational force can have. First, the way temptation affects our decision-making manifest itself in the course our practical thinking takes. Motivational forces, however, can play no role in explaining why our practical thinking took the course it took. Second, and more importantly, the pressure of temptation can get us to decide in ways that do not correspond to the way we would decide if our resistance was defeated by a stronger motivational force. In particular, we may give in only partially to temptation but there is no such thing as being partially moved by a motivational force.

Consider a particular case of akratic action. An agent that has come to believe that, all things considered, she should not eat meat because modern farming methods impose too much suffering on animals; while trying to get through a family barbecue by eating just salad, she experiences a strong temptation to eat meat; there is some veal available; that is a meat she enjoys very much; her mind keeps coming back to the possibility of having some veal; she anticipates the pleasure she would experience if she were to have some; resisting that temptation is becoming quite hard for her; nevertheless, she reminds herself of the terrible conditions under which veal is produced and renews her decision to resist; but then it occurs to her that there are also some chicken legs available, and she knows for a fact that these came from free range chickens (which, she supposes, live a much better life than calves that are killed for veal); now she is considering the possibility of having the chicken legs; she does not enjoys chicken as much as veal, but the thought that I would not be that bad to have some free range chicken legs keeps occurring to her; finally she gives in and decides to have some.

This is a case in which an agent has settled on a decision that is sub-optimal both from the point of view of desire and the point of view of reason. The agent seems to strike a compromise between her normative conviction and temptation. Here is how what is going on can be explained if we conceive of desires in the directed-attention sense. The desire for meat focus the agent's attention on the attractive features of having some veal, namely, the pleasure one would experience; that gets the agent considering that possibility of action under an attractive light; the agent dwells on that possibility and feels drawn to it; now she is tempted

to have some veal; but the thought that having some veal would involve benefiting from the suffering imposed on calves makes her retract from that possibility and she manages to resist; because she is still on the grip of desire, however, she quickly becomes aware of another alternative, one to which her moral objection is weaker; once again her attention is focused on the attractive features of this alternative, she is tempted and she eventually gives in.

Now, it is in part because the agent in the example is pressured by her urge to have some meat that her practical thinking takes the course it takes. That is why she dwells on the thought of having some meat, anticipating the pleasure she would get. And that is why after rejecting the possibility of having some veal, her thoughts shift to the possibility of having some chicken. But if that pressure is understood as the manifestation of a motivational force, we should not expect it to have these effects. How could it? One option is to hold that the movements in practical thought that can be identified in the example above correspond to changes in the dynamic of motivational forces. But that will not do. Our understanding of the notion of a motivational force is exhausted by the claim that when deciding to act an agent chooses that option which she is more motivated to perform. That is not a notion suitable to explain why our practical thought took the course it took in a particular case. For instance, the fact that in the example above the agent's motivation to abstain from veal was stronger than her motivation to have some veal entails that she would abstain from veal. But knowing that does not help explain why she dwells on the consideration of the possibility of having some veal. One could suggest that she is dwelling in that thought, anticipating the pleasure, because she is inclined to have some meat. But what does "inclined" means here? It cannot mean that the motivation to have some veal is stronger than the motivation to abstain, for we are supposing that is not the case. Or are we to suppose that at this point the motivation to have some veal was the only motivational force in play? That is absurd. If that was the case, what was stopping her from acting on that motivation? On the other hand, if the motivation to abstain was already in place and its stronger, then why is she dwelling on the thought of indulging? Perhaps we are to suppose that the force of the two opposing motivations is varying – at one time the motivation to indulge is stronger and the agent is inclined in that direction, at another the motivation to abstain dominates. But surely, if the strength of these motivations is changing, that is because the agent is focusing on one or another consideration, and not the other way around. Furthermore, knowing the balance of the agent's motivational forces at this point in her practical deliberation does not allow us to understand why her

subsequent practical thinking takes the course it takes – why she moved away from the thought of having some veal when she did or why she moved on to consider the possibility of having some chicken instead. Motivational forces simply are not the appropriate theoretical tool to explain why our practical thinking takes the course it takes. Pointing to the agent's desires can render intelligible to us why a stretch of practical thinking took the course it took, but that is because we have an independent understanding of how desires influence our practical thinking and not because we think of desires as motivational forces.

A supporter of the idea of motivational forces will likely say that motivational forces were never meant to account for the course a stretch of practical thinking takes. Rather, our interlocutor would insist, we need the notion of motivational force to account for the fact that desires can *pull* me to a course of action – a feature of desires that is aptly characterized as an urge or, in extreme cases, a craving. The idea of a force pushing us, against which we struggle to resist, does seem capture the kind of internal conflict an agent experiences when tempted. But, contrary to appearances, the notion of motivational force does not capture the way in which urges can press us to act.

In the example above, the agent has an urge to eat meat. That urge presses her, enticing her to have some veal. She eventually gives in to urge, but only partially – she goes for a compromise. If we take the pressure of the urge to be the manifestation of a motivational force, how are we to understand this case? At first, the possibility of having some chicken has not even occurred to the agent. The motivational force that is operating upon her can only be a motivational force in the direction of having some veal. Now, this motivational force is either stronger or weaker than the motivational force produced by the moral considerations to which the agent is sensitive. If it is stronger, then we should expect the agent to act according to it. That is not what happens. If it is weaker, we should expect the agent not to give into the urge. But that is not what happens either. Rather, what happens is that the agent finds herself inclined towards another course of action. That is, in the end, the effect the urge has on her. But if the urge is a motivational force, how does that happen?

It is not the case that the agent has two urges, one to eat veal and another to eat chicken. Rather, she just has an urge, even a craving, for meat and that is what is pressuring her. So perhaps that urge manifests itself in the form of two different motivational forces, one to have veal, the other to have chicken. But, according to proposal we are considering, we experience opposing motivational forces as an internal conflict. That means that if for some

reason the agent in the example above could only choose one kind of meat to eat, even if she had no moral qualms with eating meat, she would feel conflicted, pulled in two different directions by two opposing motivational forces. But there is no reason to suppose that that would happen. And that is not how we experience urges. If she had an urge for meat, she would simply go for the one she enjoys the most, without any conflict. One could insist that she would not experience any conflict because the motivational force in favor of, say, having chicken would not be sufficiently strong to have an effect on her. But we have to suppose that the motivational force in favor of having some chicken is strong enough to win over the motivational force of her moral reservations, which is not negligible. So, the option of supposing that the urge produces two (or more) motivational forces is not available.

Another option is to suppose that when the motivational force associated with the urge is blocked (for instance, by a moral reservation) it can be channeled in another direction. But recall that according to the proposal we are considering, temptation is to be understood as the effect of a motivational force. If in the example above the motivational force in favor of having some veal is simply channeled towards the option of having some chicken, then there is no longer a motivational force pressuring the agent to have some veal. The agent should no longer feel tempted to have some veal. But that is not the case. Even after having decided to have some chicken, the option of having some veal, a meat she enjoys much more, may still tempt her.

Perhaps then we could conceive of the motivational force in favor of the option of having some chicken as being the product of the interaction between the motivational force produced by the urge and the motivational force produced by moral considerations. Just like two physical forces operating over an object can produce a resulting force at an angle, two conflicting motivational forces can result in a motivational force pointing towards a compromise. But if motivational forces behaved like that, we should expect agent's always to go for a compromise when faced with a significant temptation. Plain resistance to a strong temptation, such that the agent does not give up any ground to desire, would not be possible were a compromise option to be available. But it is.

What this shows is that the idea of a motivational force does not capture the way in which an urge presses us to act when we are tempted. In giving in to an urge, I may decide for a compromise, so to say, between reason and desire. But a motivational force is not the kind of thing that can pressure us into a compromise. Surely, we can make the fact that the

agent decided for a compromise intelligible by pointing to the fact that she was feeling the pressure of an urge (although it is probable that more information would be required as well), but that is because we have an independent, pre-theoretical understanding of how urges and cravings can affect or decision making (and the course our practical thinking takes). The content of the idea of the *pressure* produced by an urge is exhausted by this understanding. Talk of motivational forces does not add anything to it. Indeed, any grasp we seem to have of how these forces are arranged at a given point during deliberation and how they will play out (whether, for instance, the stronger one will prevail and lead to a decision, or whether one will be deflected or channeled in another direction) is tributary of this independent understanding. In sum, nothing is to prevent one from saying that in the example above the urge for meat is a motivational force that pressures the agent into settling for a compromise. But that is no more informative than the claim that the agent was tempted, drawn to the possibility of having some meat, and eventually gave in partially.

One can still insist that we need the notion of motivational force to account for the fact that the agent made the particular decision she made instead of going with another available alternative. We could explain that by saying that the agent decided to do what she was most motivated to do. But at this point motivation has been reduced to the logical correlate of intentional action. The claim that the agent was more motivated to do A than B has no more content than the claim that she decided to do A rather than B. Explaining why she was more motivated to do A rather than B is simply a matter of making the decision for A rather than B intelligible. And to do that we can only appeal to our understanding of how desires and normative considerations influence our decision making. If talk of motivational forces could not contribute to that understanding, it cannot help us here as well.

This discussion started with the claim that talk of motivational forces could give us a deeper grasp of why temptation makes it hard to decide according to our normative judgments. The idea was that if we take a motivational force to underlie temptation, we can explain the effort involved in deciding according to our judgment as the effort of opposing that force. I hope to have shown that the way in which urge press us does not correspond to the influence of a motivational force. If that is the case, then the effort of resisting cannot be the effort of opposing a motivational force. What this shows, I hope, is that a model that postulates motivational forces does not provide a deeper account of temptation. It has no explanatory advantage over the volitionalist model with respect to the phenomena of

temptation and *akrasia*. And, therefore, these phenomena do not provide us with a reason to reintroduce the idea of motivational forces. To be clear, the point here is not that the volitionalist model provides a better, deeper account of the phenomena at stake. It does not. It presents temptation as an unanalyzed phenomenon; it does nothing to explain why it should be hard to decide when our attention is focused on an attractive alternative – that is simply how it is; and it does not explain what strength of the will is or how it can be developed. The point is simply that a view that postulates motivational forces does no better.

5. Taking Stock

The image of our motivational capacities that emerges from our discussion so far is this: the will is a capacity for practical reasoning in a particular sense. It consists in our capacity to decide to act in light of certain considerations. To decide is to settle on an intention to act, which I have argued, is properly conceived of as the commitment to a plan of action rather than a normative judgment. Practical reasoning in this sense can be an unreflective, even automatic process. One may be constituted so that, in certain circumstances, one immediately decides for a course of action upon becoming aware of certain considerations, without any thinking. As long as the Anscombean why-question has application, a piece of practical reasoning can be ascribed to the agent. But we can also explicitly engage in the process of trying to reach a decision. That reflective process is what I have called practical deliberation. We expect practical deliberation to usually take the form of theoretical deliberation about what we have most reason to do (which concludes in a normative judgment) because that is how a perfectly rational being would deliberate and we are reasonably rational beings. But practical deliberation can deviate from this ideal. One can linger in practical deliberation after having reached a normative verdict and eventually reach a decision that is not aligned with that verdict. Frequently that is caused by the influence of a desire that tempts the agent. This influence, however, cannot be conceived of as the work of a motivational force. Rather, I have suggested, the desire influence the agent's deliberation by focusing her attention on features of an alternative course of action that make it look attractive and that makes it hard to decide according to the judgment.

According to this picture, desires have an important but not central role in moving us to action. By directing our attention, they can suggest possible courses of action, their effects on us are frequently relevant for our decisions and they can also unduly influence our decision making. But they are not the fundamental motivational element. We can decide to do

something while having no desire to it and we can decide to do something while desiring strongly to do something else. What get us acting is the will, our capacity to make decisions. It is often claimed, however, that desires have a central role to play in the explanation of actions. That is what I turn to in the next chapter.

4. Desires, Beliefs and the Explanation of Action

1. Introduction

So far, I have argued that we should reject the idea that we are moved by motivational forces and hold, instead, that we are endowed with a will. We decide to act in light of certain considerations, form a corresponding intention and then execute that intention in action. Practical reasoning, the activity of the will, plays a central role in the etiology of action. I have presented the view I am defending as entailing the rejection of the claim that we are directly moved by desires. But one may take issue with this way of putting things. One may claim that even if it is true that I showed that we are not moved by desires conceived of as motivational forces, I have not showed that we are not moved by desires. For one can hold that we are moved by desires but reject the idea that we are moved by motivational forces. One could simply hold that we can explain intentional actions by pointing to a belief-desire pair (as we surely can) and that these explanations are causal explanations – that is, that when one such explanation is correct, the belief-desire pair in questions causes the action. In that sense, one could argue, we are moved by desires,⁵¹ but holding this view does not commit one to the idea that there are motivational forces.

I will refer to the view that rejects the notion of motivational forces (and with it the hydraulic model) while holding that whenever we act intentionally our action is caused by a belief-desire pair as the *standard model*.⁵² Thus understood, the standard model is not a view about what underlies the process by which we come to a decision. It tells us nothing about how our often-conflicting desires, attitudes, values, etc., interact in getting us to act. According to this view, a belief-desire pair can explain why an agent did what she did, but in doing so it does not explain why this particular pair and not another available pair caused the action. That, the supporter of the standard model holds, is simply a matter of how the complex neurological, biological and physical goings-on that ultimately result in action play out. It is not at all a matter of the pair that moved the agent producing a stronger motivational force than other available pairs. Holding so would be to fall back on the hydraulic model. Of course, if someone ask “why did she do A rather than B?”, we can answer “because her desire for A was stronger than her desire for B”. Once we reject the hydraulic model and the idea of

51 In this chapter, as before “desires” is used to refer to pro-attitudes in general, not only desires in the substantive sense.

52 This is how Velleman (2000) refers to the view. Smith refers to it as the “standard story of action” (2012).

motivational forces, however, the only content that can be given to the claim that a particular desire or belief-desire pair was stronger than another is that as a matter of fact the desire got the agent to act despite the presence of the opposing desire (and, perhaps, that it would do so on a number of counterfactual situations).⁵³ The strength of a desire cannot be understood as an amount of something (motivation) associated with it. Something that could be measured by considering the desire on its own, apart from its relation to other desires, and could explain why it interacts as it does with other desires. Rather the notion of the strength of a desire is an essentially comparative notion, understood only in relation to other desires. Insofar as there are facts about the relative strength of desires, these facts are fully captured by claims about how the agent would behave in a number of situations. Claims about the strength of desires, therefore, are not claims about a factor that could explain why the agent was moved by this rather than that desire. To make a claim about the relative strength of desires is simply to register that the agent was moved (and would be similarly moved in a number of counterfactual situations) by this rather than that desire.

If we understand the standard model in this way then it is not vulnerable to the arguments I presented against the hydraulic model. Furthermore, it could easily account for multiple-incentives cases by holding that in such cases only one of the agent's belief-desire pairs causes the action. (This is another possible interpretation of Davidson's view). At first, then, it may seem as if the standard model offers an alternative to the volitionalist model I have been advocating – an alternative I failed to rule out. The goal of this chapter is to show that this appearance is misleading.

It is an undisputed fact that we are capable of acting for reasons. Any plausible view about motivation must account this fact. Supporters of the standard model acknowledge this and usually present their views as a reductive account of what it is to act for a reason. According to this view, to act for the reason that *p* is to have one's action caused by the belief that *p*, together with a suitably related desire. If this reductive account was successful, then the standard model would be a rival alternative to the volitionalist model. I will argue, however, that this reductive project fails. The reason it fails is that once we have rejected the

53 Another option is to understand talk of the strength of desires as talk about the phenomenological strength of desires in the sense discussed in the previous chapter. However, only substantial desires have phenomenological strength and the desires which figure in the explanation of actions according to the standard model are not substantial desires. One can have a substantial desire not to do something and do it anyway. The desire that causes the action in this case is stronger than the desire not to act, even if it is phenomenologically weaker.

hydraulic model, desires (in the broad sense in which supporters of the standard model use the word) have to be understood as dispositions to decide to act in light of certain considerations. The very notion of desire presupposes and can only be understood by reference to the notion of deciding to act in light of a consideration and, thus, to the notion of acting for a reason. The latter, therefore, cannot be reduced to the former. Or so I will argue on section 3.

If the reductive account of acting for a reason fails, how else can we understand the standard model? In section 4 I argue that the only option is to understand it as the view that desires cause our actions by causing us to decide to act in light of a consideration. This view already presupposes, however, that we are capable of engaging in practical reasoning and that the process in which practical reasoning consists plays an irreducible role in the production of action. This is exactly what the volitionalist model I developed in the previous chapters holds. At this point the standard model is no longer an alternative but a version of the volitionalist model. One could still insist that the standard model differs from the volitionalist model in that it holds that we are moved by desires. But, given that the desires that figure in the standard model are simply dispositions to decide to act in light of a certain consideration, the claim that we are moved by desires is reduced to the claim that whenever we decide to act in light of a consideration we manifest a disposition to decide to act in light of that consideration. This is perfectly trivial and in no way conflicts with the volitionalist model. I conclude that the non-reductive reading of the standard model collapses on the volitionalist model.

Before moving on to discuss these claims, however, we need to say something about the notion of a reason for action. Supporters of standard model usually refer to the belief-desire pair that explains an action as *the reason why the agent acted*. This suggests that reasons for action are psychological states that explain the action. This understanding of reasons clashes with the natural view that reasons for action are not psychological states of the agent but considerations that count in favor of acting. In order to properly understand the standard model we have to clarify in what sense the psychological states that explain the action are reasons and how that relates to the idea that reasons are considerations that count in favor of acting. That is business of the next section.

2. Reasons for Action

Davidson opens his “Actions, Reasons and Causes” with the following question: “What is the relation between a reason and an action when the reason explains the action by

giving the agent's reason for doing what he did?". He goes on to say that giving the reason why an agent did something is a matter of naming a belief-desire pair. He calls one such pair a "primary reason" (Davidson, 1980, p.3) and holds that explanations of actions in terms of primary reasons are causal explanations.⁵⁴ This suggests that the agent's reason for doing what she did and the reason that explains why she did what she did are one and the same thing, namely, psychological states that explain and cause action. However, this way of putting things confuses different senses of "reason for action" that have to be distinguished.

Suppose a firefighter rushes into a burning building because he believes there is someone trapped in there and, being a good firefighter, he wants to save that person. The firefighter has a primary reason in Davidson's sense. But suppose he was mistaken – nobody was trapped in the building. If that is the case, it makes perfect sense to say that he had *no reason* to rush into the building and risk his life. What we mean by that is that in light of what the facts were, there was nothing to be said for the firefighter acting as he did. If there was indeed someone trapped in the building, then there would be something to be said for the firefighter risking his life. In this case, he would have had a reason to act as he did, namely, that someone was trapped in the building.

In this sense, reasons are facts that count in favor of an action or an attitude.⁵⁵ The fact that I promised to meet you is a reason to do so. The fact that the result of the elections will have a great impact in our lives is a reason to vote. The fact that the subway is far away may be a reason to take a cab. The fact that doing something will get me what I want may be a reason to do so. The fact that you betrayed me may be a reason to be angry at you. Statements about reasons for action in this sense usually take the form "*F* is a reason for *P* to do *A*". They claim that a three-place relation *R* ("is a reason for") holds between a fact *F*, a person *P* and an action *A*.⁵⁶ Reasons in this sense are *normative reasons*. These are the reasons

54 In the same way, Smith (1994, p. 131) uses "motivating reasons" to refer to the belief-desire pairs that explain actions. As I explain below, I will reserve the phrase "motivating reason" to refer to the considerations that moved the agent, rather than the psychological states that explain her action.

55 This intuitive view is upheld by Scanlon (1998), Raz (1999), Dancy (2000) and Parfit (2001). To claim that reasons are facts is not to beg the question against the Humean thesis that the reasons an agent has are grounded in some features of her psychology, such as what she desires, wants or cares about. A sophisticated Humean view about reasons holds that a fact is a reason for an agent to perform a particular action only because the agent has certain desires. The psychological fact that an agent has certain desires explains why a particular fact *R* gives her a reason to act in a particular way, but the psychological fact itself is not part of the reason. This is Schroeder's view (2007, p.57).

56 The fact that reason statements express a three-place relation is not immediately evident in every reason statement. Some such statements have the form (i) "*P* has a reason to do *A*", (ii) "*F* is a reason to *A*" or (iii) "there is a reason to *A*". Despite the appearances, however, all these statements can be interpreted as making a claim about the three-place relation *R*. We get (i) when we existentially quantify into the fact-place in

that bear on the question of what to do.

Now, reasons are not only considerations that bear on the question of what to do. We can also act for reasons. If the fact that I promised to do something is a reason to do so, then I can do it for the reason that I promised to do so. But it is an unfortunate fact that we can act for reasons that are not good reasons. In the example above, the firefighter rushed into the burning building because he thought someone was trapped in there, but that was not the case. If no one was trapped in the building, then nothing actually counted in favor of running into a burning building – as I said, there was no reason for the firefighter to risk his life. But he did not act on a whim or for no reason. He acted for a reason. So, we can ask for an agent's reasons for acting even when there is no normative reason for so acting. It may seem as if an agent's reasons in this sense could be identified with primary reasons in Davidson's sense, but that would be a mistake. It is possible for an agent to act for a reason that counts in favor of so acting. That is, it is possible for the reason for which the agent acted to correspond to a normative reason. An agent's reason for acting in this sense is, then, a consideration *R*, such that it makes sense to say “*R* was a reason for her to do *A* and her reason for doing it was that *R*”. When we ask for the reasons for which someone acted in this sense, we are asking for the considerations in light of which she acted – we are asking the Anscombean why-question discussed in the previous chapter. Reasons in this sense, the reasons for which the agent acted, are often referred to as *motivating reasons*.⁵⁷

We can think of motivating reasons as the considerations that motivated the agent. Given that we have rejected the notion of motivational forces, however, we should be careful with what we mean by that. To say that a consideration motivates the agent is not to say that it produces a certain amount of motivation. To say that the consideration motivated the agent can only mean that the agent was persuaded to act (or dissuaded from acting) by the consideration. If you are persuaded to act by a consideration, then you decide to act in light of it. The motivating reasons for which an agent acted are, therefore, the considerations in light of which the agent decided to act, that is, the considerations in light of which she settled the question of whether to act. We can say that an agent was more strongly motivated by this rather than that consideration, but that can only mean that despite taking both considerations

relation *R*. We get (ii) when we universally quantify into the agent-place. And we get (iii) when we existentially quantify over the fact-place and universally quantify over the agent-place. See Schroeder (2007: 17-9). See also Scanlon (2014: 30-1).

57 See Dancy (2000), Allan Gibbard (1990, p. 162), Derek Parfit (2001) and Schroeder (2007). Scanlon refers to the considerations in light of which the agent acted as the agent's “operative reasons” (1998, p. 19).

to be relevant to the question of whether to act, she found one to be more persuasive than the other and ultimately decided in light of it.

We frequently explain actions by pointing to the agent's motivating reason. If an agent's motivating reason for doing something is that she promised to do so, we can say "she did it because she promised to do so". At first view, then, it may seem that to explain an action done for a reason we can simply lay out the agent's motivating reasons, the reason for which the action was done. But the possibility of deciding to act in light of a false consideration spells trouble for this view. If the firefighter's reason for rushing into the burning building was that there was someone trapped in there but as it turns out no one was trapped, then we cannot appeal to the fact that there was someone trapped in the building to explain the action – there is no such fact. Something that is not the case cannot explain why someone acted.⁵⁸ To explain the firefighter's action we must cite some fact. And a natural candidate for explaining the action is the psychological fact that he *thought* someone was trapped in the burning building. Presumably, a more complete explanation would say that he rushed into the building because he thought someone was trapped in there and wanted to rescue this person. What explains the action in this case is a belief-desire pair – what Davidson calls a primary reason. If one such belief-desire pair can explain an action when the agent decided to act in light of a false consideration, then it could also explain the action had the agent decided in light of a true consideration. We can always explain an action, therefore, by pointing to a primary reason.

The primary reason that explains the action, however, cannot be identified with the agent's reasons for acting, i.e., with her motivating reason. To use one of Dancy's examples (2000, p.125), suppose you see someone violently shaking his boots and you are told he is doing that because he believes there are pink rats living in his boots and wants to get rid of them. These psychological facts may very well explain his action but they are not the agent's reason for shaking his boots. His motivating reason for shaking his boots (the consideration that motivates him to shake his boots) is not that he believes there are pink rats living in them, but rather that *there are pink rats living in his boots* – which is something he believes. This is not to say that psychological facts cannot be motivating reasons. They sure can. Our agent could decide to see the doctor in light of the worrying fact that he believes

58 Dancy once held that "a thing believed that is not the case can still explain an action" (2000, p.134). More recently he came to accept that only something that is the case can explain an action, although he still resists the view that what explain actions are belief-desire pairs (Dancy, 2014, p.89-90).

there are pink rats living in his boots. The fact remains, however, that it is a mistake to describe the primary reason that explains an action as the agent's (motivating) reason for acting. Therefore, what Davidson calls primary reasons are neither normative reasons for action nor motivating reasons for which the agent acted, but rather *explanatory reasons* – reasons that explain why the action took place.

We can distinguish, then, between three kinds of reasons for action. There are normative reasons (considerations that count in favor of an action or attitude), motivating reasons (considerations that motivate the agent) and explanatory reasons (psychological facts that explain why the agent performed the action).⁵⁹ These different senses, however, do not reflect a mere ambiguity in our use of “reasons for action”. There are close relations between them. The relation between normative and motivating reasons is clear. The motivating reason for which someone acted can be a good reason for acting, that is, it can be something that speaks in favor of so acting. In that case, the agent's motivating reason *is* a normative reason. But there is also a close relation between explanatory reasons and motivating reasons. As Davidson puts it, a primary reason not only explains the action, it explains it in a particular way: “by giving then agent's reason for doing what he did”. That means that when a primary reason explains an action, we can (perhaps with a little ingenuity) recover from the content of the belief and the desire “some feature, consequence, or aspect of the action the agent wanted, desired, held dear, thought dutiful, beneficial, obligatory, or agreeable” (Davidson 1980, p. 3). That is, given the belief-desire pair that explains the action, we can reconstruct the agent's motivating reason.

With these distinctions in mind, let us return to the standard model. According to this model, whenever we act intentionally, with a view to a goal, our action is caused by a belief-desire pair. This pair is an explanatory reason and as such it must be suitably related to the agent's motivating reason. Now, the agent's motivating reason for acting is the consideration in light of which she decided to act. But how is the claim that an action is explained and caused by a belief-desire pair related to the claim that the agent decided to act in light of a particular consideration?

In his defense of the standard model, Smith addresses exactly this question. His answer is that the considerations “that motivate agents are fixed by the contents of the desires

⁵⁹ This taxonomy of reasons is also defended by Hieronymi (2011, p.411), although she uses a different terminology, and Alvarez (2016).

and means-end beliefs that cause those agent's actions in the right way" (Smith, 2012, p.392). But what exactly does it mean to say that the considerations that motivate the agent are fixed by the contents of the belief and desire that cause the action?

There are two ways in which to understand this claim. One option is this: to claim that the belief-desire pair that causes the action fixes the consideration in light of which the agent decided to act is to say that there is nothing to being moved by a consideration beyond having one's action caused in the right way by the appropriate belief-desire pair. According to this proposal, the standard model is a reductive account of what it is to act for a (motivating) reason. It holds that to decide to act in light of the consideration that p is to have one's action caused by a suitably related belief-desire pair. Call this the *reductive reading*.

If one denies that deciding to act in light of a consideration can be reduced to having one's action caused by a belief-desire pair, then the claim that the belief-desire pair that causes the action fixes the agent's motivating reason must be understood as the claim that the belief-desire pair in question in some way makes it the case that the agent decides to act in light of a particular consideration. According to this view we are capable of performing the activity of deciding in light of a consideration. The psychological process of making a decision plays an irreducible role in the causal etiology of the action, but the course it takes is determined by the causal influence of the belief-desire pair that causes the action. That is, the pair causes the action by inducing the psychological process of deciding and giving it a particular shape. Call this the *non-reductive reading* of the standard model.

Once we have rejected the hydraulic model and have a firm understanding of the relation between explanatory and motivating reasons, the standard model has to take one these two forms. In what follows I will argue that the reductive reading of the standard model is bound to fail. We should, therefore, adopt the non-reductive reading. But this reading of the standard model collapses on the volitionalist model.

3. The standard model as a Reductive View

According to the reductive reading of the standard model, to decide in light of the consideration that p is to have one's action caused by a suitably related belief-desire pair (presumably, the belief that p together with an appropriate accompanying desire). The most common objection to this reductive account arises from the existence of deviant causal chains. The best-know example is Davidson's: a climber is holding his partner on a rope, preventing

him from falling to his death; he wants to rid himself from the weight and danger of holding his partner and he knows that loosening his grip on the rope would accomplish that; this belief-desire pair so unnerve him that he loosens his grip on the rope, dropping his partner to his death (Davidson, 1980, p.79). In this example, the action was caused by a belief-desire pair of the right sort, but the agent did not drop his partner intentionally. He did not choose to loosen his grip nor did he decided to loosen his grip in light of the consideration that doing so would rid himself of the danger of holding on to his partner. The belief-desire pair did not cause the action in the *right way* so that the agent counts as having decided in light of a consideration. If the reductive account of acting for a reason is to work, then we have to provide a reductive account of the normal, non-deviant causal chain from the belief-desire pair to the action. And a number of philosophers, including Davidson himself, do not think this can be done.⁶⁰

I will not push this objection to the reductive account. I believe there are some promising answer to it.⁶¹ Rather, I want to take a step back and focus on a more fundamental issue with the reductive account. If a reduction is to be successful, then we must be able to understand the elements in the reduction base without reference to that which is being reduced. This is why deviant causal chains put pressure on supporters of the reductive account to offer a reductive analysis of non-deviant causal chains. If we propose to reduce acting for the reason that *p* to having one's action caused by the belief that *p* and an appropriately related desire *in the right way* but can only understand "in the right way" as "so that it is the case that the agent acts for the reason that *p*", then we do not really have a proper reduction. Now, the reductive account that concerns us here purports to reduce the activity of deciding to act in light of a consideration and the elements in the reduction base are desires, beliefs, actions and the relation of causation. I believe that the main problem with this reductive account is that the notion of desire it employs cannot be properly understood without reference to the very notion of deciding in light of a consideration.

We can get at this problem by considering what is a desire according to a supporter of the reductive reading. Since we have rejected the hydraulic model, we cannot take desires to be motivational forces nor dispositions to have motivational forces under some conditions. We also cannot understand desires here as substantial desires in the sense

60 "Several clever philosophers have tried to show how to eliminate the deviant causal chains, but I remain convinced that the concepts of event, cause, and intention are inadequate to account for intentional action" (Davidson, 2004, p.106).

61 See, for instance, Setiya (2007, p.32) and Smith (2012, p.398-399).

discussed in the previous chapter. For one can act in a way one has no substantial desire to act. According to the standard model, even in one such case the agent is moved by a desire. So “desire” here means something other than substantial desire. A desire, as the word is used by a supporter of the standard model, is simply a pro-attitude. It covers everything from appetites we share with non-human animals (such as hunger and thirst) to complex attitudes of valuing (such as one’s concern for justice), from a fleeting fancy (like a sudden desire to have a beer just about now) to a permanent character trait (as the love for one’s children or a concern for one’s health).

But what are pro-attitudes? What unifies all these appetites, aversions, values, urges, preferences and so on? What is the trait that gives them such a prominent place in the explanation of action? Indeed, what is this trait that makes it the case that whenever one acts intentionally, there can be no doubt that the action can be explained by identifying one such pro-attitude and a suitably related belief? The most popular answer to this question is to hold that desires or pro-attitudes, in the broad sense that is at play in the standard model, are dispositions to act with a view to a goal.⁶² This is how Davidson presents this view:

“If a person is constituted in such a way that if he believes that by acting in a certain way he will crush a snail he has a tendency to act in that way, then in this respect he differs from most other people, and this difference will help explain why he acts as he does. The special fact about how he is constituted is one of his causal powers, a disposition to act under specified conditions in specific ways. Such a disposition is what I mean by a pro-attitude.” (Davidson, 2004, p.108)

According to this view, the reason why whenever we act intentionally there must be a desire or pro-attitude that can explain the action is that these desires are simply dispositions to act in certain ways when one has certain beliefs. That one has such a disposition follows from the very fact that one has acted. As Smith puts it, the standard story conceives of desires in this manner because “absent such a disposition, motivation is a conceptual impossibility” (Smith, 2012, p.393).

Even if we agree that there is a necessary connection between acting and being disposed to act, this dispositional view of desires, as stated, is clearly incomplete. Desires are not just dispositions to act in a goal-directed way. Very often they manifest themselves in the form of emotional reactions, in the direction of one’s attention and in certain patterns of thought. This fact is easily reconciled with the dispositional view of desires. One can simply

⁶² This view of desires is widespread among supporters of the standard model. See, for instance, Smith (2012, p.393) and Hyman (2015, p. 107).

hold that in addition to dispositions to act, desires can also involve dispositions to have a number of emotional reactions, to have one's attention drawn to certain objects or features of objects and to engage in certain patterns of thought. Desires are sets of dispositions of different kinds. But, a supporter of the reductive reading would claim, what gives a desire its motivational power and its privileged place in the explanation of action is the fact that it is constituted, at least in part, by a disposition to act.

Why stop there, however? Whenever a belief-desire pair explains an action we can identify a consideration as the agent's motivating reason, as the consideration that motivated her. Why not say, then, that desires are also characterized by a disposition to be moved by certain considerations? If that is the case, it may well be true that desires are dispositions to act when one has a suitably related belief, but only because they are dispositions to be moved by certain considerations. This view is intuitive, even to supporters of the standard model. Smith, for instance, characterizes desires as "dispositions to be moved in certain ways, depending on what means-end beliefs we have" (Smith, 2012, p.393). He cannot mean that desires are dispositions to be moved by desires, for that is clearly uninformative. And given his rejection of the hydraulic model, a disposition to be moved cannot be a disposition to be moved by a motivational force. So, he can only mean that desires are dispositions to be moved by considerations. Considerations move us by either persuading us to act or by dissuading us from acting. A disposition to be moved by a consideration is, then, either a disposition to be persuaded to act by a consideration or to be dissuaded from acting by a consideration. And these, in turn, are disposition to decide, that is, dispositions to settle the question of whether to act in a particular way light of certain considerations.

Of course, supporters of the reductive reading of the standard model have a good reason to avoid this understanding of desires. As I have already stated, once we reject the hydraulic model, the claim that an agent has a disposition to be moved by certain considerations can only be understood as the claim that she has a disposition to decide to act or to refrain from acting in light of these considerations. And if we need the notion of deciding to act in light of a consideration to fully characterize desires, then the reductive reading fails. A supporter of the reductive reading has, therefore, two options: either deny that desires can be understood as dispositions to be moved by certain considerations or hold that a disposition to be moved by certain considerations can be reduced to a disposition to act when one has certain beliefs. The first option is not viable. If whenever a desire manifests itself in

an action in the right way the agent counts as having been moved by certain considerations, then desires involve dispositions to decide in light of certain considerations. So, a supporter of the reductive reading must hold that dispositions to decide to act in light of certain considerations can be reduced to dispositions to act given certain beliefs. But this option is also not very promising. There is a good reason to think that a disposition to decide in light of certain considerations is different from a mere disposition to act when one has a suitably related belief. Dispositions can be individuated by the way in which they are *manifested* and dispositions to decide can be fully manifested in situations in which a corresponding disposition to act is not fully manifested. Let me elaborate.

To have a disposition is to be disposed to manifest it under some condition. Thus, for instance, a soporific drug has a disposition to cause one to fall asleep when ingested. This disposition is manifested when it causes a person to fall asleep upon being ingested. I have a disposition to close my eyes and put my hands in front of my face when I see an object moving fast towards my face. This disposition is manifested when someone throws a ball at my face and I close my eyes and raise my hands. In general, a disposition is a disposition to do that which fully manifests the disposition.⁶³

If desires were just dispositions to act with a view to a goal, then desires would be manifested only by goal-directed behavior. This is clearly false. Desires are also manifested “by feeling glad, pleased, or relieved if the desire is satisfied, and sorry, displeased, or disappointed if it is frustrated” (Hyman, 2015, p.107), or by worrying about the desire object or in daydreaming about it. So, desires are not only dispositions to act but also dispositions to experience certain emotional responses and to exhibit certain patterns of thought. This does nothing to upset reductive account we are considering. But desires can and do manifest in other ways.

Consider first the case of an agent that forms an intention to do something but never acts on that intention. As Davidson claims, someone “may intend to build a squirrel house [and] despite his intention, he may never build a squirrel house, try to build a squirrel house, or do anything whatever with the intention of getting a squirrel house built” (Davidson, 1980, p.83). There is no doubt that this can happen. One may form an intention and simply forget about it, for instance. Now suppose that the agent formed the intention of building a squirrel house because he decided to do so in light of certain considerations. In this case, the

63 For a discussion of the notion of manifestation, see Hyman (2015, p.107-108).

agent had a disposition to decide to build a squirrel house in light of certain considerations and this disposition was fully manifested in him deciding to do so. It may well be true that he also had a disposition to build a squirrel house given some suitably related belief. Nevertheless, this disposition was not fully manifested in this case – the agent never acted so as to fulfill his decision. But if the disposition to decide and the disposition to act fully manifest in different conditions, then they are not the same disposition. And then the reductive reading fails.

One could react by reformulating the reductive reading. Instead of holding that desires are dispositions to act given certain beliefs, a supporter of the reductive reading could hold that desires are first and foremost dispositions to form or acquire certain intentions given certain beliefs. According to this view, desires cause actions by causing us to form certain intentions (i.e., adopt certain plans) which we then execute. If the agent in the example above had a disposition to form an intention to build a squirrel house given certain beliefs, then that this disposition was fully manifested in him forming that intention. Therefore, the claim that a disposition to decide in light of certain considerations is simply a disposition to form certain intentions given certain beliefs is perfectly compatible with the example. A supporter of the reductive reading could hold, then, that to decide to do *A* in light of a particular consideration just is for a suitably related belief-desire pair to cause one to form the intention of doing *A*.

But a disposition to decide in light of certain considerations cannot be reduced to a disposition to form an intention either. Consider this example. You want to have a relaxing weekend. You start to consider a plan: spending the weekend all by yourself seems great; your family has a house by the lake; you could stock up on supplies, drive up there and shut yourself from the world for a couple of days; but then you remember the keys are with your annoying uncle; if you ask him, he will want to tag along; that is not good; so, you drop the plan you were concocting. In this case, the agent decided not to go to the lake house. And in doing so, she was guided by a desire. As long as we use “desire” and “want” as supporters of the standard model do, it is clear that if someone were to ask why the agent dropped the plan of going to lake house, the answer could be something like “because she wanted to avoid her uncle”. The decision not to go to the lake manifests, therefore, a desire. It also fully manifests a disposition to decide not to act in light of certain considerations. But no goal-directed action takes place in this example. The agent simply started to develop a plan, found it lacking and dropped it. She did not drop the plan with a view to a goal, she simply dropped it. Nor did she

form any particular intention. She did not settle on a plan to avoid her uncle. In general, we do not form a new intention every time we consider and reject a possible plan of action. So, neither a disposition to act nor a disposition to form an intention fully manifested themselves in this case. That means that the disposition that is manifested in the agent's decision cannot be identified with any of these dispositions. And this shows that we need the non-analyzed notion of deciding not to act in light of a consideration to characterize the desire that manifests itself in the agent's decision in this case. It consists, at least in part, of a disposition to decide not to go to the lake house in light of the consideration that her uncle would be there. The reduction of the act of deciding not to act to the operation of that desire is, therefore, bound to fail.

One could take issue with the fact that this example deals with a negative decision (with a decision *not* to do something) whereas the reductive reading concerns itself with positive decisions (decisions to do something). A supporter of the reductive reading could hold that even if the activity of deciding *not to act* cannot be reduced to the causal operation of belief-desire pairs, the activity of deciding *to act* can. However, this view is very implausible. It entails that the activities of deciding not to act and deciding to act are of different kinds. In the example above, the agent's desire to avoid her uncle manifests itself in the agent deciding to drop the plan in light of the consideration that following through with it would require her to spend time with her uncle. A process or activity of deciding took place in this case and it cannot be reduced to the operation of the agent's desire to avoid her uncle. This activity consists in a piece of practical reasoning – an effort to settle the question of whether to go to the lake house. If the desire plays a causal role in this case, that role can only consist in influencing or shaping the agent's practical thinking so that she decides not to go to the lake house (i.e., settles on a negative answer to the question of whether to go) in light of the consideration that were she to go she would have to spend time with her uncle. Now, the agent in our example could have decided to go to the lake house. If she did not find her uncle so annoying or if she really wanted to go to the lake house, she could have decided to go. Consider a scenario in which she decided to go. In this scenario the agent did not decide not to go to the lake house. But why not? According to the reductive reading, the reason the agent did not decide *not to go* to the lake house is that in this scenario *no* episode of non-reducible practical reasoning took place at all. Rather, it is simply the case that a belief-desire pair played its typical action-inducing role (where this is not to be understood by reference to the

notion of deciding in light of a consideration or the notion of practical reasoning). But this introduces unnecessary complications. It is much more plausible to say that in both scenarios the agent engaged in practical reasoning in the same sense and that in the second scenario she did not decide not to go to the lake house because her non-reducible practical reasoning took a different course – she settled on a different answer to the question of whether to go to the lake house, in light of a different consideration.

If that is the case, then the desires that can figure in the explanation of actions are constituted, at least in part, by dispositions to decide to act in light of certain considerations. Holding that desires involve dispositions to decide does not mean that we have to deny that when a desire explains an action the agent was disposed to act as she did and that this disposition is part of what constitutes the desire. We just have to hold that a disposition to act can be analyzed into further dispositions: a disposition to decide to act in light of certain considerations, a disposition to hold on to one's intention of acting and a disposition to execute that intention in due time. This view has the advantage of providing a unified account of the motivational power of desires. According to it, desires play the same role in the production of action, in the forming of an intention and in mere negative decisions not to act. In all three cases, the desire contributes in the same way, by shaping the agent's practical reasoning.

We have very good reason, therefore, to hold that desires involve dispositions to decide in light of certain considerations and that these dispositions cannot be reduced to dispositions to act. That being the case, we should reject the reductive reading of the standard model.

4. Desires and the Explanation of Practical Reasoning

According to the standard model, whenever we act our action is caused by a belief-desire pair and whenever a belief-desire pair causes us to act in the right way we are moved by a particular consideration, that is, we decide to act in light of a particular consideration. If we reject the view that deciding to act in light of a consideration can be reduced to having one's action caused by a belief-desire pair and hold on the standard model, then it follows that belief-desire pairs cause us to act by causing us to decide to act in light of certain considerations (that is, by causing us to consider the question of whether to act in particular way and to settle it in light of a particular consideration). If that is the case, then the

activity of deciding to act in light of a consideration has a role to play in the etiology of the action. Belief-desire pairs do not cause actions directly, rather they cause us to decide to act in light of a consideration. In deciding to act we form the intention of so acting. And, if everything goes well, we execute that intention in action. A full account of why an agent acted as she did (where this is understood as a request for the explanatory reasons for which she acted and not for her motivating reasons) must mention the complex fact that the agent decided to act in light of a particular consideration, therein intended to act and successfully executed that intention.

To a certain extent, then, any supporter of the non-reductive reading of the standard model must agree with Hieronymi's view that this complex fact can rationalize the action – that is, can explain it by giving the agent's reason for acting (2011, p.421). It can explain the action because it is part of its causal history. The agent acted because she decided to act in light of a particular consideration and followed through with her decision. And in so explaining the action we give the agent's reason for acting because her motivating reason is embedded (as a motivating reason) in the complex fact that explains the action. The agent acted because she settled the question of whether to act in light of a certain consideration. This consideration is the considerations that moved her, that persuaded her to act – her motivating reason.

However, one need not think of this proposal as a competitor to the view that belief-desire pairs can explain and rationalize actions. Rather, one can see it as a way of filling it in, that is, as a way of explaining how belief-desires pair explain actions and how they allow us to grasp the agent's reason for acting. When we explain an action by pointing to the complex fact that the agent decided to act in light of a particular consideration and followed thought with her decision, we explain why the agent acted at the same time that we specify the reason for which she acted. But we do not explain why she acted *for that reason*. Here desires play a central role. The answer to the question why the agent acted for a particular reason (why she decided in light of a particular consideration) usually points to a desire. For instance, the fact that an agent cares for her health can explain why she decided to eat awful-tasting pig's tripes in light of the consideration that they are full of vitamins. And if that is the case and she ate the pig's tripes, then there is a sense in which it is correct to say that the agent ate the disgusting pig's tripes *because* she cared for her health or that her caring for her health is what led her to eat the disgusting pig's tripes. So, the belief-desire pair formed by her

concern for her health and the belief that pig's tripe is full of vitamins can explain her action. It explains the action by explaining why her practical reasoning took the course it took. Furthermore, this belief-desire pair can rationalize her action, because it allows us to reconstruct the course of practical reasoning that resulted in the action.⁶⁴ Granted, it can do so only to the extent that it contributed to the production of the action in the right way. But once we reject the reductive reading, there is no difficulty in explaining what is the right way. A belief-desire pair contributes to the production of an action in the right way when it explains why the agent's practical reasoning took the course it took.

So, once we reject the reductive reading of the standard model, the proper way to understand this model is as the view that actions are caused by complex facts of the kind that Hieronymi's view emphasizes and that belief-desire pairs explain, in a causal way, why these facts obtain (particularly, why our practical reasoning took the course it took).⁶⁵ This non-

64 Hieronymi rejects the view that belief-desire pairs can rationalize actions and holds that only the complex fact that the agent decided to act in light of a particular consideration and followed through with her decision can do so (2011, p.419-421). She holds that a belief-desire pair cannot rationalize an action because, as illustrated by cases of deviant causal chains, even if a belief-desire pair causes an action "it is not yet clear that [the contents of the desire and the belief] were treated, by the agent, as reasons in the standard normative sense, nor, crucially, that the agent's so treating them has any role to play in the explanation of what, in fact, happened—it is not clear that those contents played the role of anyone's operative reasons." (2011, p.419). This objection, however, ignores the fact that supporters of the standard model hold that belief-desire pairs rationalize an action only when they cause it in the *right way*. As noted, specifying the right way is not a problem for the non-reductive reading of the standard model. The belief-desire pair causes the action in the right way when it causes the action by causing the agent to decide in light of a particular consideration (that is related to the content of the relevant belief). Contrary to what Hieronymi holds, then, when a belief-desire causes an action in the right way we can be sure that that the content of the belief (or a suitably related consideration) played the role of the agent's motivating reason.

65 One could object there are other possible formulations of the standard model. In particular, Smith seems to defend a version of the standard model that does not qualify as a reductive reading and does not take the form I just described. He holds that belief-desire pairs cause agents to act but only when they exercise their "rational capacities" in order "to put their desires and beliefs together so as to produce a bodily movement" (Smith, 2012, p.399). According to him, this shows that "that actions of the sort that the standard story is a story about have, as part of their explanation, a distinct [...] exercise of agency for which the standard story is not itself appropriate—namely, an agent's exercise of her rational capacities." (2012, p.399). It is hard to pin point Smith's view. One option is to take "rational capacities" to refer to the agent's capacity for practical reasoning and to understand the claim that in practical reasoning the agent puts a desire and a belief together as the claim that the considerations in light of which we decide are always considerations about our desires and about the means to satisfy these desires. But this is false. We usually decide in light of considerations that make no reference to our desires. And Smith himself acknowledges this in an older paper (Pettit and Smith, 1990). Another option is to understand Smith as holding that the activity of deciding in light of a consideration can be reduced to a more basic activity of putting together beliefs and desires. But this view incurs in the same problem discussed in the previous section: if we need the non-analyzed notion of deciding in light of a consideration to fully characterize desires, then any attempted reduction of the activity of deciding that mentions desires is bound to fail. Perhaps what Smith means by "rational capacities" simply is our capacity for practical reasoning and what he means by the claim that this is a capacity "to put desires and beliefs together" is that whenever we decide in light of a consideration we manifest both a desire and a belief. If that is the case, he is defending the non-reductive reading of the standard model as I presented it. But then we cannot make sense of his claim that the "exercise of our rational capacities" cannot be explained by belief-desire pairs. For once we reject the reductive reading of

reductive reading of the standard model has some advantages over the reductive reading.

First, it is phenomenologically accurate. Consider, for instance, the case of an agent who ate awful-tasting pig's tripe because she has a concern for her own health and believed that pig's tripe are full of vitamins. The agent in this case does not experience her decision as the mere causal operation of that belief-desire pair. Rather, from her point of view, what happens is that her attention is drawn to certain considerations ("eating this is good for my health"), these considerations incline her to act in certain way and eventually she decides to act in light of them. The non-reductive reading of the standard model allows us to take the agent's experience of deciding at face value, as corresponding to a decision process that plays an actual role in the production of the action, while retaining a central place for belief-desire pairs in the explanation of action.

Second, the non-reductive reading of the standard model has an important explanatory advantage over the reductive reading. When an agent acts for a reason, she knows what her reason for acting is without having to find out. When someone ask me why I am doing what I am doing, I do not have to look around for a reason – I already know what my reason for acting is.⁶⁶ If I do not know that, something has gone very wrong. In that case I am alienated from my action. My action becomes incomprehensible to myself and I am likely to stop doing what I am doing until I can figure out why I am doing it. Now, according to the standard model, whenever an action is caused in the right way by a belief-desire pair, the agent counts as having acted for a particular motivating reason. So, whenever an action is caused in the right way by a belief-desire pair, the agent knows what her reason for acting is.

the standard model, there can be no doubt that particular episodes of practical reasoning are explained by belief-desire pairs, because one such pair can only explain why an agent acted by explaining why her practical reasoning took the course it took. It seems likely then that Smith's view is based on a mischaracterization of practical reasoning as an activity that deals with desires rather than an activity in which desires manifest. I do not believe, therefore, that it provides a genuine alternative to the two possible readings of the standard model I laid out.

66 This point is made by Wallace (1999, p.240-241) and Setiya (2007, p.40). It is important to emphasize that saying that when I act for a reason I know the reason for which I act is not to say that whenever I act for a reason I know the explanatory reasons why I act. I know what my motivating reason is. That is, I know in light of which consideration I decided to act. This is perfectly compatible with the Freudian idea that a full explanation of why I acted may point to psychological facts about myself I am not aware of, for even though I know what my motivating reason for acting is I may be unaware of the reason why I decided to act in light of this particular consideration. Suppose, for instance, that I decided to become a medical doctor in light of the consideration that this is a very prestigious career. In this case, my motivating reason is the consideration that the medical career is very prestigious. But perhaps what explains why I decided to pursue this career in light of this particular consideration is the fact that I have a deep seated, subconscious desire to please my mother and that she has made very clear to me that she expected me to be a doctor. If that is the case, the subconscious desire to please my mother is part of the explanatory reason why I decided to become a doctor. Knowing the motivating reason for which I decided to become a doctor does not guarantees that I am aware of the explanatory reason why I decided to become a doctor.

We may wonder why that is the case. If the non-reductive reading is correct, then an explanation is readily available. Whenever a desire causes an action it does so by prompting and shaping an episode of practical reasoning. The agent acts because she decided to act in light of a particular consideration. And because she has decided to act in light of a particular consideration, she knows the reason for which she is acting. On the reductive reading, in contrast, the connection between acting because one had a certain belief-desire pair and knowing the reason for which one acted is mysterious. Unless the right way of causing an action involves the agent engaging in practical reasoning and deciding to act in light of a consideration (in which case the reductive reading fails and desires involve dispositions to decide), it is not clear why an action could not be caused in the right way by a belief-desire pair without the agent knowing the reason for which she acted. Therefore, the non-reductive reading has an explanatory advantage at this point.⁶⁷

Now, the non-reductive reading of the standard model thus understood concedes to the volitionalist model that we are capable of engaging in practical reasoning and that the process in which practical reasoning consists plays an irreducible role in the production of action. It insists, however, that when an agent acts for a reason, the fact that the agent decided to act in light of particular consideration is explained and caused by one of the agent's desires. The question is whether this is enough to distinguish the non-reductive reading of the standard model from the volitionalist model. I believe the answer is no – once we reject the reductive reading, the standard model collapses on the volitionalist model.

67 According to the non-reductive reading, desires involve dispositions to decide in light of certain considerations. If our way of deciding to act in light of a consideration is such that when we act because we have decided to act, we know what our motivating reason is, then desires involve dispositions to decide in light of certain considerations in such a way that one knows the reason for which one is deciding. One could object that this view is too demanding. Non-human animals have desires and can be moved by these desires, but it is not clear that they can decide in light of considerations in such a way that they know the reason for which they are acting. Whether or not non-human animals know the reasons for which they act is a hard question. But even if the answer is negative, I do not think this poses a problem to the view that *our* desires involve dispositions to decide in light of certain considerations in such a way that we know the reason for which we act. Suppose non-human animals do not know the reasons for which they act. In that case, their desires cannot involve disposition to decide in light of considerations in the relevant sense. This would be a problem if our desires and the desires of non-human animals were the same kind of desire. But that cannot be the case. If the objection is to work, the desires of non-human animals are such that they can cause actions in the right way without the animal knowing the reasons for which they acted. If our desires were of the same kind and produced actions in the same way, they could cause actions in the right way without us knowing the reasons for which we act. But that is not possible and supporters of the standard model agree. So, if the supposition that grounds this objection is correct, then our desires have to be different from the desires of non-human animals and have to move us in a different manner. If non-human animals do not know the reasons for which they act, what follows is that our rationality transforms the nature of our desires – so that they are not only brute dispositions to act in response to changes in the environment, but dispositions to decide in light of certain considerations.

Notice first that the claim that whenever we act for a reason we have a desire that is suitably related to the action and can explain it is perfectly compatible with the volitionalist model. According to the non-reductive reading of the standard model, desires are dispositions to decide in light of certain considerations and they explain our actions by explaining why we decided to act in light of a particular consideration. If that is the case, then the claim that whenever we act for a motivating reason we have a desire that explains our action is trivial. It follows from the claim that the agent decided to act in light of a consideration that she is so constituted that she is moved, under certain conditions, by a particular consideration. That is enough for her to count as having the relevant desire. Furthermore, if desires are dispositions to decide in light of certain considerations, then they can explain why an agent decided to act in light of a particular consideration in a trivial sense (in exactly the same sense in which supporters of the reductive reading hold that dispositions to act can explain why someone acted). If that is all there is to the claim that whenever we act for a reason we have a desire that explains our action, then the volitionalist model is perfectly compatible with this claim.

One could argue that the non-reductive reading of the standard model clashes with the volitionalist model in holding that whenever we act for a reason a desire *causes* us to act. But there is no real disagreement here. If we accept the non-reductive reading, the claim that whenever we act for a reason we do so *because* we have a particular desire can only be understood as the claim that (a) whenever we act for a reason we decide to act in light of a particular considerations, (b) that whenever we decide to act in light of a particular consideration we manifest a disposition to decide in light of this consideration under some conditions and (c) that a disposition can explain its manifestation. In claiming that desires cause us to act a supporter of the standard model is simply adding that dispositions cause their manifestation – and that, in particular, dispositions to decide cause the episodes of deciding in which they manifest themselves. That does not alter in the least the volitionalist model's account of the role of desires in our agency. It simply attaches to that account a metaphysical view about the nature of dispositions and dispositional explanations. If that is all, then the non-reductive reading of the standard model is simply a combination of the volitionalist model with the metaphysical claim that dispositions cause (or are part of what reason, causes) their manifestations.

But doesn't the standard model entail that desires move us? And isn't that enough to distinguish it from the volitionalist model? If the claim that desires move us to act amounts

to the claim that desires cause us to act, then the answer is no. As I just argued, if we accept the non-reductive reading, the claim that desires cause us to act is reduced to a combination of the volitionalist model with the claim that dispositions cause their manifestations. Indeed, according to the non-reductive reading, there is nothing to being moved by a desire beyond manifesting a disposition to be moved by a particular consideration. To be moved by a desire in this sense *is* to be moved by a consideration. And to be moved by a consideration is to act in virtue of having decided to act in light of that consideration. The activity of deciding to act in light of a consideration is an activity of the will. So, to be moved by a desire in this sense is to be moved by the activity of the will. If that is all, the claim that we are moved by desires presupposes rather than contradicts the volitionalist model.⁶⁸

I conclude, therefore, that the non-reductive reading of the standard model collapses on a version of the volitionalist model. The former is simply a combination of the latter with the metaphysical view that dispositions cause their manifestations. There is no difference between the two models with regard to their account of what desires are and what role they play in the production of action. Furthermore, according to both models the activity of the will plays a central and unavoidable role in etiology of action and desires (in the technical sense according to which desires always accompany intentional action) can only be understood by reference to that activity.

5. Conclusion

According to the volitionalist model, we are moved by the activity of the will, our capacity for practical reasoning. According to the standard model we are moved by belief-desire pairs that cause us to act. The guiding question of this chapter was whether once we have rejected the hydraulic model, the claim that we are moved by desires conflicts with the

68 The volitionalist model I defended ascribes an irreducible role to the agent in the production of actions. That role is the role of a practical reasoner. It should now be clear that, contrary to what is sometimes implied, ascribing an irreducible role to the agent in the production of actions does not introduce any mysterious breach in the causal order. According to the volitionalist model, desires are dispositions to engage in the activity of the will in a particular way. These dispositions (or the complex physical, neurological and biological facts that realize them) cause the agent to engage in practical reasoning, the process of settling the question of whether to act in a particular way, and shape that process. The agent is the subject of this process. Reasoning practically is something the agent *does*. And that process itself is an element in the causal etiology of the action. The agent's adoption of a particular intention is immediately explained by the fact that she decided to act in light of a particular consideration, that is, by the fact that she settled the question of whether to act in light of a particular consideration. If that intention is not revised nor forgotten, it will, in time, lead to action. Therefore, even though the volitionalist model ascribes an irreducible role to the agent in the production of action there is an unbroken causal chain going from desire to action, passing through the agent's exercise of her will.

claim that we are moved by the activity of the will. I hope to have shown that that is not the case: once we reject the notion of motivational forces, the claim that we are moved by desires is true only in the sense that desires can explain why our practical reasoning took the course it took. In this sense, the claim that we are moved by belief-desire pairs is perfectly compatible with the claim that what moves us is the activity of the will.

The rejection of the hydraulic model is central to my argument. Every one admits that when we are moved by desires in the right way we act for a reason. When we act for a reason, we are moved by a consideration. And because we have rejected the hydraulic model, talk of being moved by a consideration can only be understood as talk of deciding in light of a consideration. We have, then, two options. Either the activity of deciding to act can be reduced to the causal operation of belief-desire pairs or it cannot. If this reduction was possible, the standard model would be a genuine alternative to the volitionalist model. But the reduction fails because once we have abandoned the notion of motivational forces, desires have to be understood as dispositions to decide in light of certain considerations. If the reduction fails, then we must admit that practical reasoning (the activity of the will) plays an actual role in the etiology of action. And once we admit that, the standard model collapses on a combination of the volitionalist model with the metaphysical claim that dispositions cause their manifestations.

5. Acting in Light of a Fact and Acting in Light of a Belief

1. Introduction

We conceive of ourselves as beings endowed with a very special ability: we can identify normative reasons to act in a particular way and act in response to those reasons. This way of thinking about ourselves has some interesting consequences. In particular, it commits us to the claim that we are capable of acting in response to facts about the circumstances in which we find ourselves.

This consequence can be brought to light if we pay attention to the nature of the normative reasons in response to which we act. It is not unusual for philosophers to argue that what provides us with reasons are psychological states (such as desire/belief pairs). But, as discussed in the previous chapter, this claim is false if we take “reasons” to refer to normative reasons. Statements about normative reasons have the form “ F is a reason for P to do A ”. They claim that a three-place relation R (“is a reason for”) holds between a fact F , a person P and an action A . The reasons there are for us to act are facts.⁶⁹ That is compatible with the view that our normative reasons are always provided by psychological facts about ourselves, such as the fact that one has a particular belief or desire. But that is also false. Ordinarily the reasons there are for us to act are not psychological facts about ourselves but rather facts about the circumstances in which we find ourselves. For instance, if I am a firefighter, the fact that someone is trapped in a burning building may be a reason for me to put my life in risk in order to rescue this person. The mere fact that I believe someone to be trapped in a burning building, in contrast, is not a normative reason for me to risk my life – if that belief is false, then there is really no reason for me to risk my life (although it may seem to me as if there is one such reason, in such a way that my behavior may still be reasonable if I act accordingly).

69 One could raise the following objection to this claim: reasons stand in logical relations to each other in a way that facts do not, so reasons cannot be identified with facts. For instance, the fact that someone needs help is a decisive reason for me to help entails the fact that I am late is not a decisive reason for me to help. But there is no similar logical relation between the fact that someone needs help and the fact that I am late. To answer this objection we have to distinguish between normative facts and facts with normative importance (Parfit, 2011, Vol. 2, p.279-280). Normative facts are facts about reasons whereas facts with normative importance are the facts that are reasons. For instance, the fact that someone needs help may be a reason for me to help. It is, then, a fact with normative importance. The fact that the fact that someone needs help is a reason for me is a normative fact. The logical relations to which the objection points, are logical relations between normative facts. Reasons are facts with normative importance, not normative facts. So reasons do not stand in logical relations that facts do not stand in. That may seem to be the case because whenever we have a reason there is a corresponding normative fact that stands in a number of logical relations with other normative facts. But there is no difficulty in holding that normative facts stand in logical relations of a kind that reasons do not stand in. I thank Silvia Altman for pushing this objection.

There are some cases in which the fact that an agent has a particular belief may be a normative reason for her to perform a certain action. For instance, the fact that one believes that there are pink rats living in one's boots is a good reason to see a psychiatrist (Dancy, 2000, p.125). But cases like this are the exception, not the rule. It is fair, then, to say that for most actions *A*, the fact *F* that is a reason for agent *P* to do *A* is a fact about the circumstances in which *P* finds herself rather than a psychological fact about *P*'s beliefs or desires. Thus, for instance, the fact that someone is in need or in danger is, in certain circumstances, a reason for me to help; the fact that someone is obnoxious is, in many circumstances, a reason to avoid that person and the fact that someone is a human being is a reason to respect her. Therefore, if we are indeed capable of acting in response to the normative reasons we recognize, we are capable of acting in response to facts about the circumstances in which we find ourselves in.

One kind of reaction to this claim is to argue that it is simply impossible to act in response to facts. The kind of action we are interested in is intentional action, the kind of action that is performed with a view to a particular goal. From the fact that an agent performed an action *A* with a view to a goal *G* it follows that she had a pro-attitude towards the goal *G* (she either desired *G*, wanted *G*, prized *G*, was inclined towards *G*, etc.) and believed that performing *A* was either a means to *G* or constituted *G*. In light of this fact, one could argue that we never act in response to facts but are rather moved by psychological states, namely, belief-desire pairs. But that would be a mistake. All that follows from the fact that the performance of an action *A* with a view to a goal *G* *authorizes* the ascription of a particular belief-desire pair to the agent is that one cannot act in response to a particular fact without having a particular belief-desire pair – and that does nothing to upset the claim that, at least on occasion, we act in response to facts. One cannot run in response to the fact that the bus is about to leave if one does not care about getting on the bus or does not believe running to be a means of getting to the bus, but that need not change the fact that in running one is responding to that fact.

Another possible reaction to the claim that we are capable of acting in response to facts about the circumstances in which we find ourselves is to deflate the claim. This reaction is manifested in the view of those who hold that we always decide to act in light of our beliefs and that talk of acting or deciding to act in light of facts should be understood as an elliptical way of talking about deciding to act in light of true beliefs. The main support for this view comes from the argument from error cases, i.e., cases in which the agent decides to perform

action A in light of the consideration that M but her belief that M turns out to be false. In a case of this kind (the argument continues), it is clear that the agent cannot be said to have decided to act in light of the fact that M, since M is not the case. The right thing to say in this case is that the agent decided to act in light of her belief that M. Now, suppose we compare two different cases: a non-error case in which the agent decides to run because she sees that her bus is about to leave and an error case in which the same agent decides to run because she falsely believes that her bus is about to leave. If we restrict our attention to the episodes of practical thinking that resulted in the agent's decision in each case ("the bus is about to leave so I better run"), we will be unable to distinguish between them. Subjectively, the episodes of practical thinking in question are indistinguishable. And that means, the proponent of the argument from error cases would argue, that the agent is mobilizing exactly the same capacities for practical thinking in both cases. Given that in the error case she is mobilizing her capacity to decide how to act in light of her beliefs, the same must be true in the case in which she acts in response to a perceived fact. What should follow is that acting in light of a fact is simply a way in which to act in light of a belief, namely, it is to act in light of a true belief. There is nothing else to being capable of acting in light of facts than being capable of entertaining true beliefs and acting in light of these beliefs.⁷⁰

This conclusion, I believe, should be resisted. The goal of this paper is to argue for the view that we do have a capacity to act in light of facts in a stronger sense – a capacity that is not successfully deployed when we merely act in light of a true belief.

I will start, in section 2, by distinguishing the argument from error cases just described from a similar argument for the conclusion that motivating reasons are always to be identified with psychological facts about the agent. The latter argument is discussed by Jonathan Dancy, but we shall see that his response to it fails to address the argument from error cases that will concern us here. In section 3, I will provide a counterexample to the thesis that to act in light of a fact is simply to act in light of a true belief, thus showing that the argument from error cases is unsound. The kind of case I will present is already familiar from the discussion about motivating reasons. Some philosophers appeal to such cases in order to argue for the view that one can only act for the reason that *p* if one knows that *p*. I mobilize it

⁷⁰ This argument, of course, has exactly the same form as the argument from illusion for the conclusion that the immediate objects of perception are sense-data. See Ayer (1993, Ch. 1) for a classic exposition of the argument. It is not a coincidence that my reply to this argument from error cases will be analogous to McDowell's disjunctivist reply to the argument from illusion. I discuss the structural similarities between McDowell's disjunctivism and my view in section 3.

not to make that point but rather to show that acting in light of a true belief and acting in light of a fact are, in some sense, different activities. In section 4 I explain this difference as the difference between an unsuccessful and a successful deployment of our capacity of practical thinking. The idea to be defended is, roughly, that we have a capacity to (decide to) act *in light of a consideration* and that exercises of this capacity can take two forms: we either decide to act in light of a (true or false) belief or we decide to act in light of a fact. Only in the latter case is our capacity for decision (our capacity to engage in practical thinking) perfectly manifested. When one acts in light of a mere belief (even a true belief) the ensuing action can be perfectly intelligible and reasonable in light of the agent's beliefs, but it is still the product of a (to some extent) defective instance of practical thinking. In section 5 I extend the disjunctive analysis to the very notion of a *consideration*. Considerations themselves are either propositions or facts and that is why to act in light of a consideration can be either to act in light of a believed proposition or to act in light of fact.

2. Motivating Reasons and Acting in Light of Beliefs

It is important to notice that the problem of explaining how we can possibly act in light of a fact given that there are error cases is different from the problem of explaining how, given the existence of error cases, the reason for which we act (our motivating reason) can correspond to a normative reason there is for us to act. It will be helpful then to examine Dancy's account of motivating reasons as it aims at addressing the latter problem.

Dancy argues very persuasively against the view that motivating reasons are always to be identified with psychological states of the agent or with psychological facts about the agent. His argument is roughly the following:

- (a) The motivating reason for which an agent *P* performs an action *A* can be identical to the normative reason that recommends the performance of *A*;
- (b) Normative reasons are usually facts about the circumstances in which the agent acts;
- (c) Therefore, motivating reasons can be identical to facts about the circumstances in which the agent acts.⁷¹

Premise (a) is an expression of the common assumption that we are able to act for good reasons. Good reasons are normative reasons, i.e., reasons that actually favor the

⁷¹ This line of reasoning is developed in Dancy (2000, Ch. 5, especially p.103-8).

performance of the action they recommend. And the reasons for which we act are motivating reasons. Therefore, if the reasons for which we act can be good reasons, our motivating reasons can, at least on occasion, be normative reasons. Premise (b) seems to be beyond dispute for the reasons given above. Conclusion (c) follows from (a) and (b). But (c) is threatened by an argument from error cases:

(1) When an agent falsely believes that *M*, the motivating reason for which she acted cannot be that *M* – it must rather be that *she believed that M*.

(2) The distinction between true and false beliefs should not change the agent's motivating reason. That is, if the agent's motivating reasons in an error case is that *she believed that M*, then, in a case in which the agent's belief that *M* is true, the motivating reason is the same.

(3) Therefore, even when the agent's belief is true, the reason for which she acted must be that she *believed that M*.

Given that our normative reasons for action are psychological facts only in exceptional cases, conclusion (3) entails that only in these exceptional cases can our motivating reasons be identical to the normative reasons there are for us to act. That would mean that our capacity to act for good reasons would be seriously impaired – it could be realized only in exceptional cases. Dancy takes this argument to be valid and is willing to accept premise (2) (Dancy, 2006, p.127).⁷² His reply to it consists in denying premise (1), i.e., in holding that the reason for which one acted can be that *M* even when *M* is not the case. That means that statements of the form “*P* did *A* for the (motivating) reason that *M*” are not *factive* – they do not entail that *M* is the case. According to Dancy, there is nothing wrong in saying something like “his reason for doing this was that it would increase his pension, a matter about which he was sadly mistaken” (Dancy, 2006, p.127). The motivating reason for which *P* does *A* can be *M* even when the agent is mistaken in taking *M* to be the case. In Dancy's words: “[...] a thing believed that is not the case can still explain an action” (Dancy, 2000, p.134). Therefore, we have no reason to claim that in an error case the agent's motivating reason was that she *believed that M* rather than *M* itself. But what is it that we mean when we say that the reason for which *P* did *A* was that *M* but *M* was not the case?

⁷² Dancy formulates premise (2) as the claim that “the distinction between true and false beliefs should not change form of the explanation” because he takes motivating reasons to be explanatory reasons. I believe we should distinguish between motivating and explanatory reasons – a view Dancy came to accept (2014, p.89-90). So I formulate the premise as a claim about motivating reasons. Nothing turns, however, on how exactly we formulate premise (2) as long as it expresses the view that motivating reasons are the same, whether or not the corresponding belief is true or false.

According to Dancy, when we ascribe a motivating reason to someone what we are trying to do is to explain the agent's action by "laying out the considerations in the light of which the agent acted" (Dancy, 2000, p.132). In another passage, Dancy claims that this explanation is addressed to the question "what were the considerations in the light of which the agent chose to do what he did" (Dancy, 2000, p.175). When we ascribe a reason to an agent, we are trying to lay bare part of the agent's practical thinking that issued in action. To say "*P* did *A* for the reason that *M*" is, then, really to say that *P* decided or chose to do *A* in light of the consideration that *M* – a claim that applies equally to cases in which the agent's belief that *M* is true and cases in which that belief is false.

For Dancy, therefore, a reason ascription of the form "*P* did *A* for the reason that *M*" is correct in error cases and non-error cases alike. Reason attributions of the form "*P* did *A* for the reason that she believed that *M*" are correct only in cases in which *P* decided to do *A* in light of the consideration that she believed that *M* – such as the case in which someone decides to go see a psychiatrist in light of the consideration that she believes that there are pink rats living in her boots.

While denying the factivity of motivating reason attributions is enough to defuse the argument (1)-(3), it is plain to see that it is not enough to counter the argument from error cases presented in the previous section. According to Dancy, we can deny the factivity of motivating reason attributions because to enumerate an agent's motivating reasons is simply to enumerate the considerations in light of which the agent decided to act. Thus understood, the claim that an agent did *A* for the reason that *M* where *M* is a normative reason for her to do *A* (and, therefore, a fact about the circumstances of action) is compatible with the claim that the agent decided to act in light of a belief. All it takes for one to have acted for the reason that *M* is for one to have decided in light of the consideration that *M* and deciding in light of the belief that *M* is one way to do that. Dancy's account of motivating reasons is, then, perfectly compatible with the claim that to decide in light of a fact is simply to decide in light of a true belief. It only adds that in every case (both those in which the belief is true and those in which it is false) the agent's motivating reason is to be identified with the content of the belief in light of which the agent decides rather than the belief itself.

In what follows I will assume that Dancy is correct in claiming that an agent's motivating reasons are those considerations in light of which she decided to act and, thus, that motivating reason attributions are non-factive. The conclusion of the argument from error

cases that concerns us here can then be read as the claim that acting in light of a belief and acting in light of a consideration (that is, acting for a reason) are one and the same thing. It follows that acting in light of a fact can only be a special case of acting in light of a belief, since it is a way of acting for a reason

3. A Counterexample

The claim that acting in light of a fact is the same as acting in light of a true belief is false. Consider the following counterexample:

COUNTEREXAMPLE: Sue believes she has a paranormal power: she often has vivid dreams and she believes them to be premonitions. These “premonitions” have proved wrong more often than not but she has not taken heed of this fact and continues to believe in her prophetic powers. One fine day she has a particular vivid dream in which she sees men plotting to rig the federal lottery so as to ensure that the next winning numbers will be 1, 2, 3, 4 and 5. In light of this dream, Sue comes to believe that the lottery has been rigged. She decides, in light of this consideration, to place a bet on those numbers. Unbeknownst to her, the lottery had indeed been rigged so that those would be the winning numbers.⁷³

In this example Sue decides to place a bet in light of the consideration that the lottery had been rigged and she does place the bet. That means that the consideration that the lottery had been rigged is the motivating reason for which she placed the bet and that her motivating reason is a true consideration. But it is extremely implausible to say that Sue’s action was a practical response to the *fact* that the lottery had been rigged. The reason why we cannot see Sue’s action as a practical response to the fact that the lottery had been rigged is that she is simply not aware of that fact. There is no sense in which Sue is *acting in light of the fact* that the lottery had been rigged. Rather Sue is simply *acting in light of her (true) belief* that the lottery had been rigged. That is reflected in the fact that it is inappropriate to say that “Sue placed a bet *because* the lottery had been rigged” – rather we should say that she placed a bet because *she believed* that the lottery had been rigged.

We should be careful in extracting consequences from this example. Some have argued that cases such as Sue’s show that one cannot act for the reason that *M* if one does not know that *M*.⁷⁴ If we identify the reasons for which an agent acted with her motivating

73 This example has exactly the same structure as the example Hornsby puts forward in order to reject the view that “*P* did *A* because *MP*” is a proper explanation of action *A* whenever it is true that *P* did *A* because she believed that *M* and *M* is a reason for one to do *A*. See Hornsby (2008, p.250-1).

74 This is Hornsby view (2008).

reasons, that amounts to the claim that *M* cannot be one's motivating reason if one does not know that *M*. This clashes directly with Dancy's claim that motivating reason attributions are non-factive, for one can only know what is the case. Dancy has reacted to this line of argument by denying that we can infer that "Sue did not place a bet *for the reason that* the lottery had been rigged" from the claim that "Sue did not place a bet *because* the lottery had been rigged" (Dancy, 2008, p.276). His point is that to lay out the motivating reason for which an agent acted is to lay out the considerations in light of which the agent decided to act, and thus that one's motivating reason can be that *M* even if one does not know that *M* and, consequently, the fact that *M* cannot explain one's action in the way marked by the factive "because". That seems to be correct.⁷⁵

The fact remains, however, that it is extremely implausible to describe Sue as acting in light of the fact that the lottery was rigged. Even if we accept Dancy's account of motivating reasons, Sue's case shows that having a true consideration as one's motivating reason is not enough for one to act in light of a fact.⁷⁶ Acting in light of a true belief and acting in light of a fact cannot be the same. That the example shows this much, I believe, should be uncontroversial.

75 However, it does raise some concerns about the relation between motivating reasons and the explanation of actions. Dancy takes this to be the main relevance of examples such as Sue's (Dancy, 2014, p.88-91).

76 Although this claim is compatible with Dancy's non-factive view of motivating reasons, it does point to a much deeper problem for his account of reasons. Dancy's theory is motivated by the idea that if we are capable of acting for good reasons then it should be possible for our motivating reasons to be normative reasons. In his words "motivating reasons should be the right sort of thing to be normative reasons" (2000, p.103). Now, normative reasons are facts. So Dancy is committed to the idea that our motivating reasons can be facts. I agree with this view: when we act in response to a fact our motivating reason should be the fact itself. But the claim that having a true consideration as one's motivating reason is not enough for one to act in light of or in response to a fact seems to create a problem for the view that motivating reasons can be identical to facts. For motivating reasons, as Dancy conceives of them, are the considerations in light of which we decide to act. Sue's case shows, however, that one can decide in light of a true consideration and yet fail to decide in light of a fact. Therefore, the true consideration in light of which Sue decided to act is not a fact. But now suppose she knew that the lottery had been rigged and that she decided to act in light of that information in such a way that we would be justified in claiming that she acted in light of a fact. It seems that the consideration in light of which she would have decided in this case is exactly the same consideration in light of which she decided in the original example, namely, "that the lottery had been rigged". If that consideration was not a fact in the original example, how could it be a fact now? It seems that it cannot be a fact. And then it seems as if the very possibility of acting in light of a true belief while failing to act in light of fact would show that the considerations in light of which we decide to act can never be facts – for we could always decide to act in light of these considerations without deciding to act in light of facts. Given that motivating reasons are considerations in light of which we decide to act, it would follow that motivating reasons are never facts (and, therefore, are never identical to normative reasons). This is a serious problem for Dancy's view. The way out of the problem is to reject the supposition that both when one acts in light of the belief that *M* and when one acts in light of the fact that *M* one acts in light of the same consideration. I argue for the rejection of this supposition in section 4 by appeal to the disjunctive conception of acting in light of a consideration defended below. If my argumentation is correct, then this disjunctive conception is necessary to vindicate Dancy's intuition that motivating reasons can be normative reasons and, therefore, facts.

4. The Disjunctive Model

Sue acted in light of the true belief that the lottery had been rigged, but she did not act in light of that fact. We must distinguish, therefore, between these two activities: acting in light of a (true or false) belief and acting in light of a fact. And that shows that the argument from error cases with which we started our discussion is unsound, for its conclusion collapses the activity of acting in light of a fact into the activity of acting in light of a true belief.

But where is the flaw in that argument? It was composed of only two premises: (i) that in error cases the agent cannot be said to have decided in light of a fact, rather she must have decided in light of a mere belief and (ii) that from the standpoint of the agent engaged in practical thinking, the activities of deciding to act in an error case and deciding to act in a case in which we are willing to say that she acted in light of a fact are indistinguishable. From (ii) it is supposed to follow that (iii) the agent is doing the same thing in both cases. And from (iii), combined with our knowledge, from (i), that in the error case she is deciding to act in light of a belief, it is supposed to follow that she is deciding to act in light of a belief in the case in which she acts in light of a fact. Thus, we are expected to conclude that (iv) to act in light of a fact is to act in light of a (true) belief.

I am willing to concede both premises. Denying (i) is not an option. While denying the factivity of motivating reason attributions (as Dancy does) is a defensible move, denying the factivity of the claim that an agent decided to act in light of a fact is plainly contradictory. And (ii) seems plausible enough. The argument must, therefore, be invalid. One could suggest that the problem with the argument is that the move from (ii) to (iii) is invalid: from the fact that two things are indistinguishable from a particular standpoint it does not follow that they are the same. But I believe there is something to be said for the view that the move from (ii) to (iii) is valid. One can argue that the reason why the activities mentioned in (ii) are indistinguishable is that, at least in a particular sense, the agent is doing the same thing in both cases. Consider how we would answer the question “why are the activities of deciding in an error case and deciding in a ‘success’ case indistinguishable from the subjective standpoint of the agent engaged in practical thinking?” Well, the answer seems to be “because the agent is doing exactly the same thing in both cases, namely, taking such and such to be case and, in light of that consideration, deciding to perform a particular action - or, which is the same, in both cases the agent is exercising her capacity for practical thinking”. That seems beyond dispute. But to accept that is simply to accept (iii), at least given a natural reading of

(iii). According to this reading, we should take “doing the same thing” in (iii) to mean something like “exercising the same capacity” or “engaging in the same activity”.

If we read (iii) in this way, however, the move from (i) and (iii) to (iv) is invalid. The problem is that this move ignores the fact that it is possible to exercise the same capacity or to engage in the same activity in different ways. In particular, it ignores the distinction between successful and unsuccessful deployments of our capacity for practical thinking. The fact is that even though it is true that both in error cases and in cases in which the agent acts in light of a fact she is doing the same thing, namely, deploying her capacity for practical thinking, it does not follow that there is no difference between acting in light of a fact and acting in light of a belief. The difference between the two may be the difference between a successful or perfect deployment of our capacity for practical thinking and an unsuccessful or flawed deployment of that capacity. In order to show the argument from error cases to be invalid, therefore, we must show that cases in which agents act in light of a mere belief, be it true or false, are cases in which their capacities for practical thinking are not successfully deployed. That is what I now turn to.

When one acts in light of a fact it is possible to understand one’s action as a response to that fact. That is, I can see the action as an intelligent reaction to features of the situation in which the agent acts. I cannot understand the action in the same way when the agent is not acting in light of a fact, even if the agent decides to act in light of a consideration that happens to be true – that much is made clear by Sue’s case. But producing actions that are intelligent reactions to facts, I will now argue, is a formal end of practical thinking.

A formal end of an activity is an end that must be ascribed to any agent insofar as she is engaged in that activity.⁷⁷ That is, as long as one does not have the formal end of an activity in view, one cannot be described as engaged in that activity. A formal end provides a standard to evaluate one’s performance of the activity that is internal to the activity itself. In particular, if acting in light of a fact is a formal end of practical thinking, to represent an agent as engaging in practical thinking is to represent her as engaged in the effort of guiding her action in light of the facts, and that is so even when she ends up acting in light of a mere belief. Our capacity for practical thinking is successfully deployed or perfectly manifested, therefore, only when the ensuing action can be seen as an intelligent response to features of the situation we find ourselves in. To act in light of a mere belief is, by the standard set by the

⁷⁷ For a discussion of the notion of formal end see Tenenbaum (2007, p.6-9).

end to which anyone engaged in practical thinking is committed, the outcome of flawed practical thinking.

We can see that acting in light of a fact is a formal end of practical thinking if we consider the form practical thinking usually takes. Consider Sue. She fails to act in light of a fact, but her practical thinking is an effort to regulate her action in light of the facts. Her practical thinking is reasonably reconstructed as follows: “*Given that* the lottery had been rigged, betting on these numbers is a sure way to win, so I will do just that”. This piece of practical thinking starts from a putative fact and ends with a decision to act somehow. It reveals a lot about Sue: it tells us something about what are her beliefs and what she desires and give us a glimpse of her character. But it also presents her as attempting to adjust her behavior in light of what the facts are, even though she ends up acting in light of a mere belief. Cases such as Sue’s are, therefore, cases in which the agent’s practical thinking and the ensuing action are defective by the agent’s own standards – she set out to adjust her behavior in light of the relevant facts so that her action would be an intelligent reaction to those facts, but failed to do so.

Indeed, any instance of practical thinking that hopes to result in action must start from the consideration of a putative fact. It would be preposterous, for instance, to ascribe to Sue a practical thinking of the form “suppose that the lottery had been rigged; if that was the case betting on these numbers would be a sure way to win; so I will bet on these numbers”. One cannot move from a supposition to a decision without endorsing the supposition. Practical thinking, therefore, simply is an attempt at adjusting one’s behavior in light of the facts. One could insist that engaging in practical thinking serves another end (such as maximizing our chances of satisfying our desires). But that does not change the fact that one would not engage in practical thinking if one did not care (instrumentally at least) for adjusting one’s behavior in light of the facts.

So, to engage in practical thinking is to set out to adjust one’s behavior in light of the facts that constitute the situation in which one finds oneself. Our capacity for practical thinking is perfectly manifested only on the condition that that attempt is successful. When one fails in that attempt, one’s practical thinking is defective according to the standards that are internal to the activity itself. That means that, by its own standards, our capacity for practical thinking does not live up to expectations when we end up acting in light of a mere belief – even a true belief. Our capacities for practical thinking are successfully deployed,

perfectly manifested, when we decide to act in light of facts and the ensuing action can be seen as an intelligent, intentional reaction to these facts. In a similar way, one could say that our capacities for theoretical thought are perfectly manifested only when they lead to knowledge, not when they lead to mere true belief.

Now, from the standpoint of the agent, the cases in which her capacities for practical thinking are successfully deployed and the cases in which they do not work properly (such as error cases or Sue's case) are subjectively indistinguishable. In all these cases the agent takes herself to be deciding to act in light of a fact. In particular, Sue's practical thinking would look exactly the same if she was aware of the fact that the lottery had been rigged – she would decide to place a bet in light of the putative fact that the lottery had been rigged. If we take seriously the distinction between acting in light of a fact and acting in light of a belief, it follows, then, that there are two distinct but (potentially) subjectively indistinguishable ways of acting in light of a consideration: one can either act in light of a fact or one can act in light of a belief. The fact that they are subjectively indistinguishable should not prevent one from distinguishing between them and from saying that in one case the action in question is an intelligent response to a fact whereas in the other case it is not.

This disjunctive view of the activity of acting in light of a consideration is, of course, analogous to the disjunctive view of appearances defended by McDowell. McDowell's disjunctivism concerns the epistemic relevance of experiences. He holds that experience comes in two kinds: some experiences are such that they reveal to the subject how things are, and thus provide epistemic warrant to the beliefs about the environment that are based on them, whereas others are misleading, and do not provide epistemic warrant to the beliefs that are based on them. The two kinds of experiences are subjectively indistinguishable, but experiences of the former kind provide opportunities for knowledge, whereas experiences of the latter kind do not.

McDowell's disjunctivism can be presented as the denial of what he calls the "highest common factor view". According to this view, both kinds of experience (those that put us in contact with the environment and those that are illusory) provide exactly the same epistemic warrant to experience-based beliefs on account of their subjective indistinguishability. Given that the "bad" member of the pair does not provide an opportunity to knowledge, the same should be true of the "good" member. McDowell's disjunctivism, in contrast, holds that, despite being indistinguishable, the experiences of the two kinds differ in

their epistemic relevance. Experiences of the “good” kind reveal to the agent how things are in the environment, so that beliefs grounded in these experiences can amount to knowledge, whereas experiences of the “bad” kind seem to do that (McDowell, 2013, p.24). Knowing that an agent came to believe that p on account of having an experience in which it seems to her as if p may be enough to grant intelligibility to her belief and to render it rational. But knowing that much is not yet to know whether the experience in question was of the kind that reveals how things are (in which case the agent’s belief may amount to knowledge) or one that merely seemed to do so (in which case the belief may be rational but the agent is not in a position to know the environment).

The view I am defending here has the same structure. Episodes of practical thinking come in two kinds: those that result in the agent acting in light of a fact and those that result in the agent acting in light of a belief. From the standpoint of the agent, episodes of the two kinds may be indistinguishable, but that is compatible with the claim that in the “successful” case one manages to adjust one’s behavior in light of the facts and, therefore, one’s action may be understood as a reaction to the relevant facts. Whereas in the “unsuccessful” case it only seems to the agent as if she is adjusting her behavior in light of the facts but that is not the case and, therefore, her action cannot be understood as a response to a fact.

Of course, the difference between successful and unsuccessful episodes of practical thinking is not a difference in epistemic relevance. Rather, the difference is that in the successful case the practical thinking connects, so to speak, the action to the relevant fact whereas an unsuccessful but indistinguishable episode of practical thinking fails to do so even if it concerns the same action and the same fact.

5. Disjunctivism about Considerations

We are now in a position to answer an objection that may already have occurred to the reader. The objection is that even though we cannot reduce acting in light of a fact to acting in light of a true belief, we can, for all I have said, reduce acting in light of a fact to acting in light of the belief that M while knowing M to be the case. After all, the reason why Sue fails to act in light of the fact that the lottery had been rigged seems to be that she has no knowledge of that fact.⁷⁸ The objector could hold, then, that for all I have said it may be the

⁷⁸ This is Hornsby’s (2008, p.251) and McDowell’s (2013, p.17) view.

case that we always act in light of beliefs but there are three ways in which to do so – namely, to act in light of the belief that M knowing M to be case, to act in light of the true belief that M without knowing M to be the case and to act in light of the false belief that M.

My reply is this: the objection either rests on a misunderstanding of the expression “to act in light of a belief” or it is innocuous. That becomes clear once we consider how we must understand the objector’s use of “to act in light of a belief” so that we have an objection to the view that we are able to act in light of facts. One option is to understand the claim that we always act in light of a belief as the claim that we always decide to act in light of a mental state, rather than a fact. But that is unintelligible: we always decide to act in light of a consideration, i.e., something we take to be the case. A mental state itself cannot be a consideration in light of which we act. Another option is to understand the claim that we always act in light of a belief as the claim that we always decide to act in light of a psychological fact. But that is plainly false: we usually do not decide to act in light of the consideration that we have such and such beliefs but rather in light of the consideration that the circumstances of action are such and such. Finally, if we take the claim that we always act in light of a belief to mean that we always decide to act in light of a consideration in which we believe, then there is no objection at all. Indeed, the disjunctive model discussed in the previous section holds precisely that we always decide to act in light of a consideration which we hold to be true, but that we can do so successfully (in which case we act in light of a fact) or unsuccessfully (in which case we act in light of a belief).

Perhaps what underlies this objection is a complaint that could be put like this: there is no real, deep difference between acting in light of a fact and acting in light of a belief; in both cases one engages in the psychological process or activity of making a decision in light of a consideration; this process is always fueled by one’s beliefs, and the only relevant difference between acting in light of a fact and acting in light of a belief concerns the epistemological credentials of those beliefs. In a way this is true, for according to the disjunctive model deciding to act in light of a fact and deciding to act in light of a belief are two ways of doing the same thing, namely, deciding in light of a consideration. Nevertheless, no one will deny the relevance of the distinction, as reflection on Sue’s case makes evident. The complaint must be really about the terms in which the distinction is drawn. It could be formulated like this: “To characterize the distinction between the two ways of acting in light of a consideration as a distinction between acting in light of a fact and acting in light of a

belief is misleading. Why? Because this way of drawing the distinction implies that what distinguishes the two ways of acting in light of a consideration is the kind of thing in light of which the agent decides – whether it is a fact or a belief. But what determines whether one acted in light of the fact that M or the mere true belief that M is whether or not one knew M to be the case. So, the real distinction is between acting in light of a consideration one knows to be true and acting in light of a consideration one merely believes in. There are not two kinds of things in light of which one can decide to act (facts and beliefs) but only one (considerations). What distinguishes the two ways of deciding is not that in light of which the agent decides but the agent’s epistemic stand in relation to the consideration in light of which she decides. Therefore, to characterize the distinction between these two ways of deciding as a distinction between acting in light of a fact and acting in light of a belief is misleading”. This concern should dissipate, however, once we have a clearer understanding of what a consideration is and, especially, once we see that the disjunctive model extends all the way down to the very notion of consideration.

If a consideration was a belief or some other mental state, then to draw the distinction between ways in which to decide in light of a consideration in terms of the contrast between facts and beliefs would indeed be misleading. But that is not the case. Considerations are that from which practical thinking proceeds. They are, so to speak, the premises on which the practical thinker decides to act in a particular way. As such, considerations cannot be mental states, for mental states do not figure in practical thinking in this manner. But what are considerations? Here are some truism about considerations: a consideration can be true or false; one can believe in a consideration; given a consideration one can make certain inferences and one can know what follows from a consideration even if one regards it as false; one can know a consideration to be true or to be false. These truisms may suggest that considerations are propositions. And that is at least part of the truth. Consider an error case: she runs because she falsely believes that her bus is about to leave. The consideration in light of which she decides is properly expressed by “my bus is about to leave”. This consideration is the content of the belief in light of which she acted: she believed that my bus was about to leave. The consideration in light of which she decided in this case is, then, the object of her belief. The received view is that the objects of beliefs are propositions. So, in this case, the consideration in light of which the agent decided was a proposition. But now consider a case in which we are willing to say that an agent acted in light of a fact: say, an agent sees that her

bus is about to leave and decides to run in light of that consideration. For this to be a genuine case of acting in light of a fact, the agent must know that her bus is about to leave. So the consideration in light of which she decides, namely, that “my bus is about to leave”, is something she knows. That is to say, the consideration in light of which she decided is a possible object of knowledge. And what can be known? What are the possible objects of knowledge? Well, knowledge is always knowledge of a fact. That means that a known consideration is actually a fact. So when one decides in light of a known consideration one decides in light of a fact.

That is enough to show that talk of acting in light of facts is perfectly reasonable: surely we act in light of considerations, and some considerations are facts. But it may seem to have troublesome consequences. If a fact is just a true proposition, then it should follow that whenever an agent acts in light of a true belief, she acts in light of a fact – and that, we have already established, is not the case. We can avoid that undesirable conclusion, however, by denying that facts are true propositions, and there are good reasons for doing so. I will point out only two. First, facts cannot be true or false as propositions can. So, for instance, one can assume for the sake of argument that a true proposition is false, but there is no such thing as assuming a fact to be false. One can assume for the sake of argument that the facts are different, but facts themselves cannot be said to be true or false. Second, and more in touch with our concerns here, propositions are not suitable objects of knowledge as facts are. If something can be an object of knowledge, then it can be learned or discovered. But when one learns something, one does not learn a proposition and one that is true (what could that even mean?). Rather one learns or discovers (the fact) that a particular proposition is true.⁷⁹

We should, then, distinguish between propositions (either true or false) and facts. The former are the objects of belief, the latter the objects of knowledge. But the considerations in light of which we decide to act can be either things in which we believe or things we know. What that shows is that we should adopt a disjunctive conception of the notion of a consideration. A consideration is either a mere believed consideration or a fact. And that explains why acting in light of a consideration can take the form either of acting in light of a belief (or, as one may put it, in light of something believed) or acting in light of a fact.

⁷⁹ See Hyman (2015, p. 163).

Before concluding, let me note a relevant consequence of this disjunctive conception of considerations. The motivating reasons for which one acted are the considerations in light of which one decided to act. But if motivating reasons are considerations, and considerations are either propositions or facts, then motivating reasons also are either propositions or facts. This allow us to do away with a potential objection to the view that our motivating reasons can be facts. The observation that one can fail to act in light of a fact even when the motivating reason for which one acted is a true consideration may have suggested that even when we act in light of a fact our motivating reasons themselves are not facts. For it seems that an agent can decide in light of the same consideration, say M, both in a case in which she decides to act in light of a true belief and in a case in which she decides to act in light of a fact. If that is the case and motivating reasons are the considerations in light of which we decide, then the agent's motivating reason should be the same in both cases. Since in the first case the agent does not decide in light of a fact, her motivating reason cannot be a fact. If the motivating reason is the same in both cases, then it is also not a fact in the case in which the agent decided to act in light of a fact.⁸⁰ The way out of this problem is to deny that an agent that acts in light of the true belief that M and an agent that acts in light of the fact that M decide to act in light of the same consideration. And the disjunctive conception of considerations allows us to do so. Even if there is a sense in which both agents decided to act in light of the same consideration (in that their motivating reasons are expressed in exactly the same way), in another sense the considerations in light of which they decided are different in that one is a fact and the other a mere true proposition.

This leads to another interesting consequence. When one acts in light of a fact, one's motivating reason (the consideration in light of which the agent decided) is itself a fact. To the extent that we are motivated, moved or led to act by our motivating reasons, when an agent acts in light of fact what moves her or what leads her to act is the fact itself. Facts about the circumstances of action can be motives in this sense. And then, when one acts in light of a fact, it is possible for the motive for which one acted to be identical to a normative reason to act in that manner.

The starting point of our discussion was our self-conception as beings capable of acting in response to normative reasons. This self-conception entails that we are capable of

⁸⁰ This is the same problem presented in note 9.

acting in response to facts. Arguments from error cases might seem to force us to deflate that self-conception for they seem to show that to act in light of a fact is simply one way of acting in light of a belief. I have argued against this deflationary view. I offered a counterexample to it and argued that in order to reject the argument from error cases on which it is grounded we should adopt a disjunctive conception of acting in light of a consideration. This disjunctive view, I hope to have shown, allows us to take seriously the idea that we are capable of acting in response to facts and thus to take seriously our self-conception as beings that respond to and are capable of being moved by normative reasons for action.

Conclusion

I have argued for the view that the will is the source of our actions, at least when we act for a reason. According to the *volitionalist* model I have defended, we are endowed with a will, a capacity to make decisions. When we act for a reason, the activity of the will is part of the etiology of the action. That is, part of what explains an action is the fact that the agent has exercised her will so as to decide to act in light of a particular consideration. According to the volitionalist model, the activity of the will leads to action in the following way: in deciding to act in a particular way, the agent forms an intention; an intention is best understood as a plan of action; so, in deciding to act, the agent settles on a plan of action; if that plan remains in place, if it is not revised or forgotten, then it will lead to action when the time to execute it comes.

Furthermore, I argued that the activity of the will cannot be reduced to the operation of desires or normative judgments. Consider desires first. If by “desires” we mean substantive desires, then one can act in a particular way, even if one has no desire to so act and actually desires not to perform the action. Nevertheless, in one such case, one acts because one decides to act in light of a particular consideration. That is, one’s action is the product of the activity of the will. The activity of the will, therefore, does not depend nor can it be reduced to the operation of substantive desires. Philosophers usually use “desires” in a broader sense that includes all pro-attitudes. But the activity of the will cannot be reduced to the operations of desires in the broad sense either. I have argued that desires in this sense can only be understood as dispositions to decide in light of certain considerations. Since the activity of the will consists in deciding to act in light of certain considerations, desires in the broad sense are nothing but dispositions to engage in the activity of the will in a particular way. These dispositions can only manifest in the agent exercising her will in a particular way. Therefore, the activity of the will cannot be reduced to the operation of desires in this sense. Quite the opposite is true: the motivational effects of desires in the broad sense can only be understood by reference to the activity of the will.

The role the will plays in the production of action also cannot be reduced to the operation of normative judgments or what I have called judicative reason. Although I argued that we should identify the will with practical reason, that is not to say that we should conceive of it as a capacity to issue normative judgments regarding what our reasons are or what we

should do. Rather, we should distinguish between judicative reason, our capacity to engage in pieces of reasoning that conclude in normative judgments about our reasons for action, and practical reason, a capacity to engage in pieces of reasoning that conclude in an intention. The will is identified with the latter. The course a piece of practical reasoning takes tends to follow the agent's normative judgments. To the extent an agent is rational, if she judges that R gives her a decisive reason to do A, then she will decide to do A in light of the consideration that R. Nevertheless, the activity of the will is independent of our normative judgments. In cases of akrasia, one decides in a way that conflicts with one's normative judgments. And in cases of normatively underdetermined decisions, one decides to act in light of a particular consideration even though one cannot reach a normative judgment about what one should do.

Because it holds that the will has a central role to play in the production of action and that its activity cannot be reduced to the operation of desires or normative judgments, the volitionalist model qualifies as an anti-reductionist theory of action. It holds that the agent has an irreducible role to play in the production of action. When an agent acts for a reason, her action is explained, at least in part, by the fact that she exercised her will so as to decide to act in light of a particular consideration. To exercise one's will is to engage in practical reasoning. So, when an agent acts for a reason, she plays a role in the production of her action, namely, the role of the reasoner. Since that exercise of the will cannot be reduced to operation of other mental states, agents have an irreducible role to play in the production of action.

That is the case even if we admit that dispositions cause their manifestations and, therefore, that desires in the broad sense cause the agent to engage in practical reasoning and cause that practical reasoning to take a certain course. If desires are dispositions to decide in light of a certain consideration and dispositions cause their manifestations, then they can cause an agent to engage in the activity of the will in a particular way. But in that case, what they cause is indeed an activity, of which the agent is the subject. Even if desires cause the activity of the will, the practical reasoning in which the exercise of the will consists is something the agent does. Given that the episode of practical reasoning is an indispensable link in the causal chain leading to action, the agent still has an irreducible role to play in the production of action. Therefore, the claim that desires, in the broad sense, cause the activity of the will is perfectly compatible with an anti-reductionist theory of action.

Since the will, conceived of as practical reason, is the source of our actions, at least when we act for a reason, we can say that it is a source of motivation. But we should be

careful with what we mean by that. The starting point of my argument in favor of the volitionalist model was the fact that there are multiple-incentives cases. I have argued that these cases show that we are not moved by motivational forces (regardless of whether the source of these forces are our desire, our normative judgments, the taking of a consideration as a reason, etc.). The reason for that is that, if multiple-incentives cases are true, then we are capable not only of choosing what we will do, but also choosing the goal with a view to which we will act. And the idea that we are moved by motivational forces is incompatible with this claim. In saying that the will is a source of motivation, therefore, we are not saying that it is a source of motivational forces that dispute the determination of our behavior with motivational forces that issue from other sources. Rather, all we mean is that the activity of the will can move us.

If the volitionalist model is correct, we are not dragged into action by our desires, nor is our behavior determined by the tug of war between desires and reason. Rather, certain courses of action are suggested to us by features of the situation we find ourselves in, we ask ourselves whether to pursue these courses of action and settle that question in light of certain considerations. In settling that question, we settle on a plan of action, which, if all goes well, we eventually execute in action. The way in which we settle the question of whether to act determines how we will act. Desires and normative judgments can affect that process. But, ultimately, it is through the activity of the will that we settle that question and, thereby, determine our own behavior.

References

- Allison, H.E. *Kant's Groundwork for the Metaphysics of Morals*, New York: Oxford University Press, 2011.
- Alvarez, M.. Reasons for action, acting for reasons, and rationality. *Synthese*, 195(8), 2016, pp. 3293–3310. <https://doi.org/10.1007/s11229-015-1005-9>
- Anscombe, E. *Intention*. Harvard University Press, 1963.
- Arpaly, N. Moral Worth. *The Journal of Philosophy*, 99(5), 2002, pp.223-245.
- Ashwell, L. The metaphysics of desire and dispositions. *Philosophy Compass*, 9(7), 2014 , pp.469–477. <https://doi.org/10.1111/phc3.12143>
- Audi, R. Intending. *Journal of Philosophy*, 70(13), 1973, pp.387–403.
- Austin, J.L.. A Plea for Excuses: The Presidential Address. *Proceedings of the Aristotelian Society*, 57(1), 1957, p.1–30.
- Ayer, A. J. *The Foundations of Empirical Knowledge*. London: Macmillan, 1963.
- Bratman, M.E. *Intention, Plans, and Practical Reason*. Cambridge: Harvard University Press, 1999.
- Broome, J. Reasons and Motivation. *Aristotelian Society*, Suppl. Vol. 77, 1997, pp.131–46.
- Buss, S. Weakness of Will. *Pacific Philosophical Quarterly*, 78(1), 1997, pp.13–44.
- Carey, J. Reason, Desire, and the Will: In Defense of a Tripartite Moral Psychology. 2017. <https://escholarship.org/uc/item/2hf590r3>
- Chang, R. Voluntarist reasons and the sources of normativity. In D. Sobel & S. Wall, eds. *Reasons for Action*. New York: Cambridge University Press, 2009, pp. 243–271.
- Chang, R. Grounding practical normativity: Going hybrid. *Philosophical Studies*, 164(1), 2013, pp.163–187.
- Dancy, J. *Practical Reality*. New York: Oxford University Press, 2000.
- Dancy, J. Acting in the Light of the Appearances. In Macdonald, C. and Macdonald, G. (eds) *McDowell and His Critics*. Oxford: Blackwell Publishing, 2006, pp. 121–134.
- Dancy, J. On How to Act-Disjunctively. In Haddock, A. and MacPherson, F. (eds) *Disjunctivism: Perception, Action, Knowledge*. New York: Oxford University Press, 2008, pp. 262–279.
- Davidson, D. *Essays on Actions and Events*. New York: Oxford University Press, 1981.
- Dickenson, J. Reasons, causes, and contrasts. *Pacific Philosophical Quarterly*, 88, 2007, pp.1–23.
- Frankfurt, H. Freedom of the Will and the Concept of a Person. *The Journal of Philosophy*, 68(1), 1971, pp.5-20.

- Frankfurt, H. Identification and Wholeheartedness. In: Frankfurt, H. *The Importance of What We Care About*. New York: Cambridge University Press, 1988, pp.159-176.
- Ginet, C. Freedom, responsibility, and agency. *The Journal of Ethics*, 1(1), 1997, pp. 85-98.
- Hare, R. M. *Freedom and Reason*, Oxford: Clarendon Press, 1963.
- Herman, B. On the Value of Acting from the Motive of Duty. In: Herman, B. *The Practice of Moral Judgment*. Cambridge: Harvard University Press, 1993, pp.1-22.
- Hieronymi, P. The will as reason. *Nous - Supplement: Philosophical Perspectives*, 23(1), 2009, pp.201–220.
- Hieronymi, P. Reasons for Action. *Proceedings of the Aristotelean Society*, vol. 111, 2011, pp. 407–427. <https://doi.org/10.1111/j.1467-9264.2011.00316.x>
- Holton, R. Intention and Weakness of Will. *The Journal of Philosophy*, 96(5), 1999, pp.241–262.
- Holton, R. *Willing, Wanting, Waiting*, New York: Oxford University Press, 2009.
- Hornsby, J. A Disjunctive Conception of Acting for Reasons. In Haddock, A. and MacPherson, F. (eds), *Disjunctivism: Perception, Action, Knowledge*. New York: Oxford University Press, 2008, pp. 244–261.
- Hume, D. *A Treatise of Human Nature*. Oxford: Clarendon Press, 1978.
- Hursthouse, R. The virtuous agent's reasons: a reply to Bernard Williams. In: Heinaman, R. *Aristotle and moral realism*, San Francisco: Westview Press, 1995, pp.24-34.
- Hyman, J. *Action, Knowledge, and Will*. Oxford: Oxford University Press, 2015.
- Kant, I. *Groundwork of the Metaphysics of Morals*. Translated by Mary Gregor. Cambridge: Cambridge University Press, 1998.
- Korsgaard, C. *Self-Constitution*, New York: Oxford University Press, 2009.
- Lowe, E. J. *Personal agency: The metaphysics of mind and action*. New York: Oxford University Press, 2008.
- McDowell, J. Are Moral Requirements Hypothetical Imperatives? *Proceedings of the Aristotelian Society*, 52, 1978, pp.13-29.
- McDowell, J. Virtue and Reason. *The Monist*, Vol. 62, No. 3, 1979, pp. 331-350.
- McDowell, J. Non-Cognitivism and Rule-Following. In: McDowell. *Mind, Value and Reality*. Cambridge, Harvard University Press, 2002, pp.198-218.
- McDowell, J. Acting in the Light of a Fact. In Bakhurst, D., Hooker, B. and Little, M.O. (eds), *Thinking About Reasons: Themes From the Philosophy of Jonathan Dancy*, Oxford University Press, 2013, pp. 13-28.
- Mele, A.R. *Springs of Action*, New York: Oxford University Press, 1992.
- Nagel, T. *The Possibility of Altruism*. Princeton: Princeton University Press, 1978.

- Oddie, G. *Value, Reality, and Desire*. New York: Oxford University Press, 2005.
- Raz, J. *Engaging Reason*, New York: Oxford University Press, 1999.
- Ridge, M. Humean Intentions. *American Philosophical Quarterly*, 35(2), 1998, pp.157–178.
- Scanlon, T. M. *What We Owe to Each Other*. Harvard University Press, 1998.
- Scanlon, T.M. Reasons: A Puzzling Duality? In: Wallace, R.J. et al. (eds), *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*. New York: Oxford University Press, 2006.
- Scanlon, T. M. *Being Realistic about Reasons*. New York: Oxford University Press, 2014.
- Schroeder, M. *Slaves of the Passions*. New York: Oxford University Press, 2007.
- Schafer, K. Perception and the Rational Force of Desire. *The Journal of Philosophy*, 3, 2013, pp. 258–281.
- Schapiro, T. The nature of inclination. *Ethics*, 119(2), 2009, pp. 229–256. <https://doi.org/10.1086/596469>
- Schueler, G. F. *Desire: Its role in practical reason and the explanation of action*. Cambridge: The MIT Press, 1995.
- Setiya, K. *Reasons Without Rationalism*. Princeton: Princeton University Press, 2007.
- Silverstein, M. *Ethics and Practical Reasoning*. *Ethics*, 127(2), 2017. <https://doi.org/10.1086/688741>
- Smith, M. *The Moral Problem*. Oxford: Blackwell Publishing, 1994.
- Smith, M. Rational Capacities, or: How to Distinguish Recklessness, Weakness, and Compulsion. In: Stroud, S. & Tappolet, C. (eds). *Weakness of Will and Practical Irrationality*, New York: Oxford University Press, 2003, pp.17-38.
- Smith, M. Four Objections to the Standard Story of Action (and Four replies). *Philosophical Issues*, 22, 2012.
- Smith, M. A constitutivist theory of reasons: its promise and parts. *Leap*, 1(24), 2013, pp.9–30.
- Stampe, D. The Authority of Desire, *Philosophical Review*, xcvi, 3, 1987.
- Tenenbaum, S. *Appearances of the Good*. New York: Cambridge University Press, 2007.
- Velleman, J.D. *Practical Reflection*. Princeton: Princeton University Press, 1989.
- Velleman, J.D. What Happens When Someone Acts? *Mind*, 101(403), 1992, pp.461–481.
- Velleman, J.D. The Possibility of Practical Reason. *Ethics*, 106(4), 1996, pp.694–726.
- Wallace, J. *Normativity and the Will*. New York: Oxford University Press, 2006.
- Watson, G. Free Agency. *The Journal of Philosophy*, 72(8), 1975, pp.205–220.
- Watson, G. Skepticism about Weakness of Will. *The Philosophical Review*, 86(3), 1977, pp.316–339.

Watson, G. The Work of the Will. In: Stroud, S. & Tappolet, C. (eds), *Weakness of Will and Practical Irrationality*, New York: Oxford University Press, 2003, pp.172-200.

Way, J. Reasons as Premises of Good Reasoning. *Pacific Philosophical Quarterly*, 98(2), 2017, pp. 251–270. <https://doi.org/10.1111/papq.12135>

Williams, B. Internal and External Reasons. In: Williams, B. *Moral Luck: Philosophical Paper 1973-1980*. Cambridge: Cambridge University Press, 1981, pp.101-113.