

Abstract BA Thesis 'How to Think About Not Thinking: Re-examining Deliberation Through the Lens of Inverse Akrasia'

Student: Guido Freriks (3714403)

Supervisor: Annemarie Kalis

Submitted 12 June 2017

In our thinking on practical reasoning, we generally assume that to deliberate means to ponder one's situation in order to select the most rational course of action, and importantly, to take the time to do so. It is an *intellectualist* exercise. In this thesis, it is argued that this might not always be so; that we can imagine agents who act rationally on the basis of reasons that were discovered through a deliberative process, but who at the same time may not be fully aware of this deliberative process going on.

This conclusion is reached through a study of 'inverse akrasia' cases as brought forth by Alison McIntyre and Nomy Arpaly: cases of weakness of will in which the 'weak-willed' course of action appears to be *more rational* than the course of action that would have aligned with an agent's best judgment, or so they conclude.

Their conclusion is tested via the work of Bernard Williams on *internal* and *external* reasons. McIntyre and Arpaly follow Williams in his conclusion that internal reasons are the only intelligible reasons for action. However, Williams is also found to put great emphasis on the presence of a deliberative process in internal reasons generation. This deliberative aspect seems, at first glance, not to be present in inverse akrasia cases.

The final part suggests that, in order to be able to include inverse akrasia cases in the scope of what we may call 'rational action', we may have to accommodate for a kind of deliberation that happens despite the agents involved not being explicitly aware of it. There is substantial evidence for this happening, but moving deliberation outside the realm of 'intellectualism' brings with it problems of its own. In further discussion on practical reasoning, the constraints of intellectualist deliberation need to be acknowledged and addressed.

How to Think About Not Thinking

Re-examining Deliberation Through the Lens of Inverse Akrasia



A bachelor thesis for the department of Philosophy & Religious Studies,
Utrecht University

Student:	Guido Johannes Freriks
Student Number:	3714403
Supervisor:	Dr. Annemarie Kalis
Date:	June 2017
Word Count:	13142
Including Notes:	14757

Preface

As the supervisor of this thesis is well aware, there is a certain irony in taking on a large project centred around the topic of ‘weakness of will’. It starts out frivolously: you tell yourself that sleeping in this morning was a kind of ‘fieldwork’, and that your own rationalizations on everything you would *rather* do than work on your project provide you invaluable insight into the akratic mind. It is a fool-proof party joke. Personally, I feel the humour of it has worn out.

My path to graduation has not been that of a straight line – twists, turns, and even the occasional bump in the road have turned me into, euphemistically put, an *experienced* student. With that in mind, I could not be happier to finally be at this point. I wish to extend my joyous gratitude to the group of friends who have joined me in the library for the past year; peer pressure is a thing of beauty. My family has also been a source of great support. Being able to vent on the phone or over a dinner table is a privilege, and one for which I am immensely thankful.

Finally, Annemarie. I don’t quite know how to thank you enough, so these words will have to do. Under your supervision, I felt a joy in doing philosophy that I never expected. And philosophize we did: the incredible depth of our discussions made it feel like I was working *with* you, and never *under* you. The fact that sometimes we both fell silent, trying to retrace the many steps that led us to the conclusions we were debating, motivated me to keep refining the points presented herein. Thank you.

Contents

Preface	[2]
Contents	[3]
Introduction	[4]
1. Why do we care about akrasia?	[6]
<i>Classical history</i>	[7]
<i>Davidson's 'unification attempt'</i>	[8]
<i>Flipping akrasia: the case of Huckleberry Finn</i>	[10]
2. The case for inverse akrasia	[12]
<i>Deliberation in McIntyre and Arpaly</i>	[14]
3. McIntyre and Arpaly vs. Williams	[17]
<i>Williams' internal reasons thesis</i>	[17]
<i>McIntyre's claim to internal reasons</i>	[20]
<i>Arpaly's claim to internal reasons</i>	[22]
<i>McIntyre, Arpaly, and Williams' proposition (iii(b))</i>	[23]
4. Döring criticism, the advent of 'intellectualism'	[27]
5. Making 'sense' of deliberation: against intellectualism	[29]
<i>Intellectualist deliberation</i>	[29]
<i>How does the intellectualist conception show itself?</i>	[30]
<i>Looking outside intellectualism: the evidence</i>	[31]
6. Williams, and the 'rational relation'	[34]
<i>The necessity of a 'rational relation'</i>	[35]
<i>The possible nature of a 'rational relation'</i>	[36]
7. Revisiting inverse akrasia: solutions and problems	[38]
<i>Closing remarks</i>	[39]
Bibliography	[41]

Introduction

Imagine landing the job you always wanted. You worked hard, spent your savings on extra training, and endured the digs from your family as they grudgingly supported you in your pursuit of what looked like a pipedream. Arriving at your new job, however, you quickly find things to be somewhat different from what you had anticipated. Perhaps it is the demands made at work, the behaviour of your co-workers, or even just the office lighting; you find it hard to put your finger on, but a certain disillusionment creeps in regardless. Still, you keep on. After all, this job is your dream!

Some time passes, and things do not change much. You continue to be supremely convinced of the fact that yes, *this is the place for you*. ‘I worked so hard to get here’, you tell yourself. On the flipside, you continue to be annoyed by the things that bothered you from the start. Imagine now, that at one point, the character of your co-workers simply gets too much for you. After what feels like the millionth inane request for help (‘he is clearly capable of doing it himself!’), those two magical words escape your lips: ‘I quit!’.

For a moment, or maybe even for a week, you are panicked. This was *not* the plan. Despite the panic at the rashness of your decision, a feeling of relief and renewed vigour are present, too. In fact, no longer hampered by your previous tunnel vision, you manage to find a more enjoyable vocation with surprising ease. How did this happen? Did you, despite having solemnly resolved to keep the job, somehow make the right call by impulsively quitting?

The mystery of what happened in the above scenario is central to this paper, as we will ponder the question of whether or not we can *act rationally against our best judgment*. In order to explore this, we will first focus on what it means to act against your best judgment in the first place, to display what the Greeks called *akrasia*, or weakness of will. We will discuss its history via the likes of Plato and Aristotle, as well as its lasting relevance through the work of Donald Davidson.

Following that, we will introduce Jonathan Bennett, who first explored the notion of *inverse akrasia*: an act of akrasia in which an agent *does the right thing*. The theories of Alison McIntyre and Nomy Arpaly will provide the bulk of discussion in this paper, as they argue that it might sometimes be more rational to act *against* your best judgment than to *follow* it. From this point, Bernard Williams’ ground-breaking work on ‘internal reasons’ will be central to our discussion. Finding that McIntyre and Arpaly make an unjustified appeal to these reasons in their claims on the rationality of akrasia, we

eventually turn to the concept of *deliberation*. We will make partial use of a critique by Sabine Döring in exploring the idea that in order to properly call ‘rational’ the actions of agents such as the one in the example above, we need to consider the exact nature of deliberation in our everyday reasoning. Is it an intellectual, explicit effort in which we ponder our situation, or can it happen in more implicit ways, too? Finally, we note that choosing either way of thinking about deliberation over the other has far-reaching consequences. The last will not have been said on the (ir)rationality of weakness of will, but if one wishes to say anything more, they should first consider how we, as humans, deliberate.

1. Why do we care about akrasia?

Akrasia, or weakness of will, is the phenomenon where one acts against one's own best judgment.¹ It points to a certain personal contradiction, where an agent judges something to be best, but somehow fails to follow up that judgment with a corresponding action. As an example, you might be able to imagine yourself sitting down in front of the television after dinner, basking in the afterglow of a satisfying meal. Yet, for all your satisfaction, you find yourself reaching for the bag of crisps that, just moments earlier, you had explicitly resolved to forbid yourself from eating. The temptation proved too much, however, and while still holding that you should not be snacking at this hour, you dig in. Formulated properly, it is “free and intentional action against one’s own judgement regarding what it would be best to do.”² It can, in fact, be formulated not only as personal contradiction (as in our after-dinner-snack example), but also as a logical one: if we genuinely think A is better than B, it follows that we should be more motivated to do A than to do B. Why would we ever do B?³

Despite the above being formulated theoretically, it is a very real phenomenon. Think back to the many times in your life when you made a promise to yourself and failed to carry it out. Promises do not even need to be involved: oftentimes we are simply overwhelmed by emotion and act imprudently. We stuff our faces with chocolate, drown our sorrows in alcohol and go on a shopping spree to combat existential dread. Still, these are all quite mundane matters. At most, they cause a certain personal discord where we feel we should have been stronger, but were not. Why care so much?

The answer to that is layered, and will depend largely on who you ask. The problem of akrasia is intimately tied to notions such as agency, rationality and morality. In terms of agency: what can we say of the authenticity (and the connected blameworthiness) of an agent acting against their own best judgment? In terms of rationality: how can we take humans to be generally rational beings when they are capable of intentionally forgoing the one tool they have to guide their actions, namely the reasons they form? And in terms of morality: if an agent’s actions are not reliably a product of their reasoning, how can we

¹ For the purposes of this paper, it suffices to treat ‘weakness of will’ and ‘akrasia’ as synonymous. As such, these terms will be used interchangeably throughout this paper, despite the presence of literature discussing possible differences between the two, such as Richard Holton’s 1999 paper “Intention and Weakness of Will”, *The Journal of Philosophy*, 96, 241-262, accessed March 2017, doi: 10.2307/2564667).

² Annemarie Kalis, *Failures of Agency*, vol. 57 in series *Quaestiones Infnitae* (Utrecht: Zeno Institute of Philosophy Publications, 2009), 2.

³ Provided one accepts Davidson’s assumptions on the nature of akrasia, as they will be discussed later on in this chapter.

judge the agent's morality and values by looking at their actions (as we so often do)?⁴ For us, the connection to rationality will be the most interesting.

As you can see, we are right to care about akrasia. One does not need to (but one still can) mention the possibility of bloody, rage-driven *vendettas* to make the argument that it is important to be 'somewhat dramatic' about agents going against their best judgment.⁵

In this chapter, we will explore some origins and interpretations of akrasia. We will not be able to reflect on the entire history of akrasia as a phenomenon in philosophy; the parts we do discuss will serve to help answer the question of the (ir)rationality of inverse akrasia by introducing it as a problem.

Classical history

As a philosophical phenomenon, akrasia has most venerable roots. From being first mentioned by Plato's Socrates in *Protagoras* to the ways in which Aristotle broadened the definition, we will come to see that even the modern debate on akrasia still broadly follows the lines set by these two philosophers.⁶

For Plato, the problem of akrasia was as obvious as it is for anyone, and yet his view on it was simple. We, as humans, strive for the good. We, as humans, would never *knowingly* choose the bad over the good. We, as humans, still seem to do so. Plato's solution is then to say that the key lies in the part about *knowingly* choosing the bad. As knowledge is always master over passion, akratic action is really just ignorant action: had the agent *known* the good, he would have been motivated appropriately. Akrasia, according to Plato, does not exist. No one would *willingly* follow a worse course of action over a better one.⁷

Aristotle's views are slightly more complex. He appears to attempt to allow for the possibility of akrasia, based on a simple look around: it is 'patently' clear that cases of weakness of will are commonplace: we often choose the lesser good over the greater.⁸ However, it also looks like he has to grant to Plato that akrasia is a problem of incomplete knowledge. Aristotle says that an akratic agent knows the *universal premises* (say, a value

⁴ Of course, this presumes a connection between the content an agent's reasoning and their moral character, which is on its own not entirely an uncontested statement either.

⁵ "It is important to be somewhat dramatic about the general phenomenon of going against your judgement", Kalis, *Failures of Agency*, 11.

⁶ With Plato's Socrates being generally seen as a *dramatis personae*, we will ascribe the views to just 'Plato' in this paper (unless otherwise specified).

⁷ Plato, *Protagoras*, cited in Sarah Stroud, "Weakness of Will", *The Stanford Encyclopedia of Philosophy* (Spring 2014 Edition), ed. Edward N. Zalta, <https://plato.stanford.edu/archives/spr2014/entries/weakness-will/>, accessed April 2017.

⁸ Aristotle, *Nicomachean Ethics*, cited in Kalis, *Failures of Agency*, 21.

such as courage) on which he wants to act, but fails to correctly apply the situational, *particular premises* (or, how this value relates to the situation). Akrasia then happens when the particular premises are not fully known to the agent.⁹ Again, if it is simply a problem of knowledge, can akrasia properly exist?

Here, Aristotle's somewhat nebulous notion of 'inactive knowledge' becomes central. Generally, an agent's knowledge of the good will be 'active', in which case it guides the agent towards what is best to do. In akratic cases, however, this knowledge is inactive, meaning that "[the agent] is not ready to *use* that knowledge."¹⁰ The reason for its inactivity is, as often seems to be the case with akrasia, *desire*. A certain 'weakness to resist' is to blame for akratic action, making an agent chase the physical appetites over the higher virtues.

Davidson's 'unification attempt'

As we fast forward some two thousand years, we find Donald Davidson attempting to 'unify' the positions of Plato and Aristotle by focusing the debate on the link between 'judging' and 'doing'. In his 1970 paper *How is Weakness of the Will Possible?* he lays out a theory that allows for the possibility of akrasia while avoiding the apparent logical contradiction between judging X to be better (and thus wanting it more) and choosing to do Y.

First however, let us see how this logical contradiction comes to pass. In laying out his theory, Davidson makes two assumptions that, in his words, follow "from a very persuasive view of the nature of intentional action and practical reasoning."¹¹ They are, as follows:

P1: 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.

P2: If an agent judges that it would be better to do X than to do Y, then he wants to do X more than he wants to do Y.¹²

⁹ Kalis, *Failures of Agency*, 21.

¹⁰ Ibid., 24, drawing from Sarah Broadie's commentary on *Nicomachean Ethics* (2002).

¹¹ Donald Davidson, "How Is Weakness of the Will Possible?", *Essays on Actions and Events* (Oxford: Clarendon Press, 2001), 31.

¹² Ibid., 23.

The contradiction emerges clearly: the akratic agent intentionally doing Y seems impossible while following these two premises.

The key to Davidson's approach to akrasia lies in his distinction between 'unconditional' (or absolute) judgments, and 'conditional' (or relative) judgments. According to Davidson, what we do in practical reasoning is form 'all-things-considered' judgments. Judgments of this sort are always conditional, in that they find their expression in sentences such as '*in light of* reasons r1-rN I should do X'.

As such, they stand in stark contrast to unconditional judgments of the form 'It is better to X'. It is in this contrast that Davidson's theory distinguishes itself: since these judgments are of such different logical form, there is no contradiction between holding a conditional judgment (all things considered, I should X) and an unconditional judgment (it is better to Y) simultaneously.

Having established the logical difference between conditional and unconditional judgments, Davidson makes the claim that it can never be *just* our conditional judgments that bring us to action. A conditional judgment may simply be of the form 'X is more exhausting than Y' and not take anything else into account. We engage in a process of practical deliberation, in which our all-things-considered judgment gives rise to an unconditional judgment. The deliberative product 'it is best to X' carries in itself our motivation to X.

This leaves the question of what happens in the case of akrasia. According to Davidson, the akratic agent is psychologically defective. Their rationality has failed them, since his 'principle of continence' would prescribe that one should follow their all-things-considered judgment.¹³ An unconditional judgment to X should have risen from the agents' all-things-considered judgment, but somehow did not, leading to akrasia.

Psychologically, what happens in cases of Davidson's akrasia has been characterized by Alison McIntyre as having to do with what she calls 'renegade reasons'¹⁴. In this case, one of the reasons contained in the set of considerations that was weighed in the all-things-considered judgment managed to 'jump out' of this set and give rise to an unconditional judgment all by itself. Clearly this is irrational action, since following a single reason

¹³Meaning one should "perform the action judged best on the basis of all available relevant reasons", Davidson, "How Is Weakness of the Will Possible?", 41.

¹⁴ Alison McIntyre, "Is Akratic Action Always Irrational?", in *Identity, Character, and Morality: Essays in Moral Psychology*, eds. O. Flanagan & A. Oksenberg Rorty (Cambridge, Mass: MIT Press, 1990), 388.

when one could also follow the set of considerations that *already contained* that reason is necessarily the less rationally inclusive option.

What can we take from this? We can take from it that akratic action is something for which an agent can never provide a reason. It is surely performed for one or more reasons, but when asked, the agent (by demand of the earlier made best judgment) is forced to admit something like ‘I don’t know why I did it’. Does that, however, necessarily make it an irrational move? Not in cases of *inverse akrasia*.

Flipping akrasia: the case of Huckleberry Finn

The Conscience of Huckleberry Finn is a 1974 paper by Jonathan Bennett in which, for the first time, something resembling ‘inverse akrasia’ is mentioned.¹⁵ It is centred around the case of Huckleberry Finn, the main character from the homonymous Mark Twain novel. The main point of Bennett’s essay is the following: Our emotion is a powerful motivating force that can compete with, alter, or even do away with our moral judgments altogether – and this might not even be a bad thing.

We are presented with the character of Huck. He is a simple white boy from the American south in the times of diminishing, but still very real slavery. During the novel, he gets to know Jim, a rather sympathetic slave. At some point Jim becomes a runaway, and it is when Huck and Jim find themselves on a boat together that Huck’s internal struggle about whether to turn Jim in or not is highlighted. In the end, Huck does not turn Jim in. The reason he does not do so and the reason that this is an interesting case are the same: it is his *sympathy*, not his *moral judgment*, that eventually brings him to help Jim escape.

According to Bennett, Huck has a ‘bad morality’.¹⁶ Through his ‘bad morality’, it makes complete sense for Huck to return Jim to his owner, and in fact he would see it as virtuous behaviour to do so. After all, slaves are a good to be possessed, and not returning stolen goods is a crime. Consequently, he suffers the pangs of sympathy, brought forth by his growing feelings of friendship for Jim, as it comes to blows with his resolution to do the ‘right thing’. In the end, he *lacks the strength of will* to follow his moral judgment. He *sincerely* believes his best judgment to be to turn Jim in, but fails to do so.

¹⁵ The term itself is not used until Nomy Arpaly and Timothy Schroeder’s ‘Praise, Blame and the Whole Self’, *Philosophical Studies*, Vol. 93, No. 2 (1999). It has since become so widespread that we will retroactively apply it when talking about the phenomenon of ‘doing the right thing against your best judgment’.

¹⁶ Jonathan Bennett, ‘The Conscience of Huckleberry Finn’, in *Philosophy*, 49 (1974), 123–24. Bennett makes no further attempt to elaborate on this term, other than that we, as readers, would probably agree with him in so characterising Huck.

We can thus clearly conclude that Huck acts akratically. Recalling the words of Davidson: he made an all-things-considered judgment to turn Jim in, but showed weakness of will in carrying that judgment out. Still; he helped free a slave, and did so out of sympathy for his plight. Something seems weirdly off in putting the case to rest by calling Huck's action patently irrational – did he not do the right thing? We have here the first described case of inverse akrasia: *doing the right thing against your best judgment*.

2. The Case for Inverse Akrasia

Understandably, if the concept of akrasia is disputed, *inverse* akrasia must be doubly so. In the coming chapter, we will feature Alison McIntyre and Nomy Arpaly. These are two authors who have, since the advent of Bennett and his argument for sympathy surrounding ‘Huck’, made a case for the existence and legitimacy of inverse akrasia as a phenomenon in our practical reasoning. Importantly, however, McIntyre and Arpaly do not explicitly argue for the idea that cases of inverse akrasia showcase an agent taking the ‘right’ path. Instead, they both argue the following: there are cases in which an agent, in acting akratically, take the *more rational* path of action, over following their best judgment.¹⁷¹⁸ We will here bring to attention how their views on the role of deliberation in practical reasoning lead them to the conclusion that inverse akrasia must exist.

One way to clearly bring to light the position of these inverse akrasia theorists on practical reasoning is to put them opposite Davidson on the topic of the ‘special’ link between one’s best judgment and one’s path of action. The theory Davidson puts forth surrounding one’s best judgment holds an agent’s judgment on what to do to have a special place in the motivational process leading to action – otherwise the existence of akrasia would not be contested at all.

For McIntyre and Arpaly, this is not as much the case. In line with modern empirical research in psychology on practical agency, they sketch a world wherein an agent is subject to many more influences than just the ones that star in their explicit weighing of what to do. According to them, these influences can not only have the kind of motivational force that could steer us away from following our best judgment, but can even contain legitimate, *good* reasons for action. Defining rational action as *acting for the best reasons*, inverse akrasia must then be possible.¹⁹

¹⁷ While McIntyre & Arpaly claim to limit themselves to speaking of ‘rationality’, and avoiding talk of the ‘good’ or the ‘right’, there is of course the threat of conflation between these terms. Unfortunately, this is a topic beyond the scope of this paper.

¹⁸ We should strongly note the use of ‘more rational’ instead of flat-out ‘rational’. This is because both McIntyre and Arpaly are well aware that, even if the akratic action turned out to be ‘doing the right thing’, there still lies a procedural rational defect in the fact that the agent’s best judgment was abandoned. Despite their use, we will in this paper sometimes use ‘rational’ to refer to the inverse akratic path of action regardless, since we are mostly investigating what it is that *makes something more rational* than the best judgment, and as such are not so concerned with perfect rationality that this distinction is of great importance at this point.

¹⁹ ‘Good reasons’ should in this context be taken to mean ‘the kind of reasons that could make an action *more rational* than the best judgment’, since inverse akrasia is about akratically taking the *more rational* path of action. This definition, despite being limited, will serve its purpose for our discussion in this paper. Of course, it leaves a lot of room for further interpretation of (for example) the Humean vs. Aristotelean sort: is ‘good’ supposed to point to desire-satisfaction, or some kind of *eudaimonia*? We cannot go into this

Still, how is it possible that these influences can be called ‘reasons’? One would think that we humans ponder our situation, assess whatever we feel is influencing us, and at the end of the road produce a rational judgment, guided by what we may (through this process) call ‘reasons’.²⁰ This is what seems intuitively to be the essence of rational action.²¹ Yet, the influences mentioned above never seem to enter into a judging process, or anything even related to the kind of deliberative process mentioned here. They seem to miss something that could transform them from mere ‘influences’ into ‘good reasons’.

Could these influences in fact be what are commonly called ‘external reasons’? That is to say, could they be reasons that, independent of any agent, can be called upon to exert their motivational force on an unsuspecting subject? It would solve our issue of them not entering our deliberation. However, McIntyre and Arpaly themselves say that they *do not make any claims* that presume the externality of these reasons. Instead, they take the reasons motivating inverse akrasia to be *internal*, meaning in the case of McIntyre and Arpaly that they are intimately tied to the agent and the agent’s belief and desire set. We will very briefly go into the arguments they offer for why this is so. In later chapters, these arguments will be treated in far greater detail, and contrasted with Bernard Williams, one of the ‘fathers’ of the internal-external reasons debate.²²

For McIntyre, the reasons that make the akratic path of action rational are internal because – simply put – *they potentially are*. Had the agent had the time and the opportunity to further deliberate, they would have come to see, in cases of inverse akrasia, that the akratic path was, in fact, the more rational path of action. Following Williams, then, the agent would have *internalized* the reasons. The agent, showing through their actions to be sensitive to reasons that would have prompted a revision of judgment if the opportunity to revise had been there, chooses the more rational path by straying from their best judgment.

For Arpaly, the reasons motivating inverse akrasia are internal because they are produced by the agents themselves; by their total collection of beliefs and desires. In her view, cases of inverse akrasia can come to pass when an agent makes a best judgment

further, but will instead continue to use ‘good reasons’ throughout this paper as mentioned and explained above.

²⁰ The importance of this ‘reason-guidedness’ is close to ubiquitous in literature on this subject. Examples can be found in the likes of R.M. Hare, Donald Davidson, but also Sabine Döring (whose paper will be treated in chapter 4).

²¹ Much like with ‘good reasons’, developing a further fleshed out concept of ‘rationality’ is beyond the scope of this paper.

²² Like the further argument, exact citations for these claims in their papers will be provided in chapter 3.

that fails to take into account certain of their own beliefs and desires. If these ‘undeliberated upon’ beliefs and desires are ones that the agent cares more strongly about than the ones figured in the original judgment, they will bring the agent to an akratic action that is in favour of these ‘other’ beliefs and desires. In the end, Arpaly states that the agent here chooses the path of action most in line with the complete, coherent picture of themselves and their beliefs and desires; the agent, despite acting akratically, chooses the *rational* path, for *good reasons*.

As we can see, both McIntyre and Arpaly hold our best judgment to be, at best, a well-intentioned (yet fallible) guide in our practical reasoning. They both believe practical reasoning to be a somewhat fuzzy, indeterminate process, in which we as agents are not always fully aware of the forces that steer us. Not all is similar in their view on agents and their relation to practical reasoning though: throughout their writing it becomes clear that they find differences in how strongly they envision the link between rationality and deliberation. This will be discussed below, as well as in further detail in chapter 3.

Deliberation in McIntyre and Arpaly

When we think of deliberation, vivid images come to mind. Most will probably have something to do with calm, reflective thinking – say, a man in a comfortable chair, smoking a pipe. We presume that this type of deliberation will lead to some exquisitely rational outcomes. How could it not? However, when it comes to the McIntyre and Arpaly, only the former would consider there to be anything going on here that is connected to rationality.

This is what paints the difference between McIntyre and Arpaly on deliberation. For McIntyre, deliberation is the primary process through which we attain rationality in our practical reasoning. For Arpaly, not so much: while deliberation is the method through which we come to ‘best judgments’, these judgments are in Arpaly’s view mere singular beliefs, as (mis)guided as any other belief could be. Let us examine how, despite the similar views on practical reasoning mentioned above, these two authors can diverge on such a key notion as deliberation.

Part of the more central role McIntyre ascribes to deliberation was already briefly touched upon: one of the conditions that must be met for inverse akrasia to be possible, is that the agent must (potentially) be able to come to discover the rationality of the akratic course of action *through deliberation*. In fact, the requirement is even stricter: in order to depart from an all-things-considered judgment, the agent must have “almost attained the

deliberative perspective from which the more rational judgment could be made.”²³ McIntyre thus paints a picture of the agent as a deliberative being who can *generally* trust their best judgment, as evidenced by her argument for ‘resolute continence’.²⁴ Still, our deliberative capacities will sometimes find themselves unable to keep up. It is in those moments that inverse akrasia can happen, and where ‘further reflection’²⁵ will be necessary to make sense of what *really* motivated us.

Such faith in reflection (and similar deliberative capacities) would not fit well with Arpaly’s story. While naturally, she does not deny that we humans are deliberative beings, she places very little stock in the capacity of single-agent deliberation to tell right from wrong, clear from unclear, or rational from irrational. In fact, though we may deliberate our way towards our ‘best judgment’, these best judgments, taken at face value, are just ‘one belief among many’.²⁶ As evidenced above, it may equally be any of our other beliefs and desires that rationalise (and thus, justify) action as opposed to our best judgment, as long as it turns out that it was the ‘other belief/desire’ that we cared more about. Arpaly also proceeds to provide examples of action that we would call ‘rational’ in which deliberation has no place. She mentions the fact that we judge sports players on the (ir)rationality of their plays while they clearly do not take the time to deliberate upon their moves. She also makes the case for ‘dawning’: agents who over time, in light of contrary evidence, slowly come to abandon a firmly held belief, but never find themselves deliberating on the matter. Instead, they at one point have a realisation – they no longer hold the belief.²⁷ From this, we can safely conclude that Arpaly does not hold our individual deliberative capacities in high esteem.

Two things can be concluded from what was said in this chapter. One is that according to McIntyre and Arpaly, inverse akrasia is a real thing, and a thing in which we are motivated by *internal reasons*. Our second conclusion is that while for both, deliberation obviously plays a central role in practical reasoning, McIntyre and Arpaly disagree over how much it has to do with making rational decisions. We will next dive into the work of Bernard Williams, a man for whom deliberation is key. His work on his so-called ‘internal

²³ McIntyre, “Is Akratic Action Always Irrational?”, 390.

²⁴ Briefly put, this states that while stubbornness should be avoided when considering one’s own best judgment, we would do well to act in a generally continent fashion, i.e. in accordance with our best judgment.

²⁵ McIntyre, “Is Akratic Action Always Irrational?”, 399.

²⁶ Nomy Arpaly, “On Acting Rationally Against One’s Best Judgment”, *Ethics*, 110 (Chicago: The University of Chicago Press, 2000), 501.

²⁷ These examples will be elaborated on in later chapters of this paper.

reasons thesis' will be crucial in examining the following question: how valid are the claims to internal reasons that McIntyre and Arpaly make?

3. McIntyre and Arpaly vs. Williams

In the previous chapter, brief mention has been made of internal and external reasons. We have, however, done so without giving proper credit to the man who so influenced the debate on this part of metaethics, and to whom McIntyre and Arpaly are (implicitly or explicitly) continually deferring in their articles: Bernard Williams. To be clear: McIntyre and Arpaly claim that the reasons motivating agents in inverse akrasia cases are internal. This ‘internal’ status of their reasons is of crucial importance to the claim that inverse akrasia can be rational to begin with, since their entire discussion on internal- or external status of reasons is based on the work of Williams; work in which he states flat-out that, in his view, external reasons do not exist. The fact that our inverse akrasia theorists follow him in this conclusion, means that they must make good on their claim that the reasons motivating cases of inverse akrasia are internal, or not only are inverse akrasia cases no longer rational – they would be nonsensical, or not ‘reasons’ at all.

I will expand on the above by first going into Williams’ ‘internal reasons thesis’, showing that he believes internal reasons for action to be the only reasons for action, and what the nature and the requirements of these internal reasons are. What follows will be the test: can McIntyre and Arpaly legitimately say that the agents in their cases act on the basis of internal reasons? I will argue that they cannot, at least not entirely. Reading Williams closely reveals that the role deliberation plays in his work is too important for McIntyre and Arpaly’s cases to bear. We will briefly observe what this means for the claims to rationality that their agents make, and in later chapters see if we can, in some way, ameliorate this ‘break’ between these two authors and Williams.

Williams’ internal reasons thesis

In his seminal 1981 paper *Internal and External Reasons*, Williams claims that internal reasons are the only way we can intelligibly interpret agent motivation. In fact, Williams considers the entire idea of external reasons to be ‘incoherent’.²⁸ One would intuitively be tempted to think of internal and external reasons along something like the following lines: a sentence expressing an internal reason would be of the form ‘A has reason to X’, while an external reasons expression would look more like ‘there is a reason for A to X’. The definition could also be framed to say that internal reason statements are claims on what the agent has *reason* to do, while external reason statements are claims on what would

²⁸ Bernard Williams, “Internal and External Reasons”, in *Moral Luck: Philosophical Papers 1973-1980* (Cambridge: Cambridge University Press, 1981), 111.

be *good* for the agent to do, in the widest sense of the term.²⁹ The point, however, is that Williams would say that something like the latter definition ('good' for the agent) can never be said without appealing to the agent having a 'reason' for action, in an internal sense. The internal sense is the only way an 'external reason' can be understood, meaning that only internal reasons truly exist.³⁰

Next, we will look at the nature of internal reasons, and what is required to speak of something as being an 'internal reason'. Williams sets forth four propositions that he believes to be true of internal reason statements. They are the following:

- i. An internal reason statement is falsified by the absence of some appropriate element from *S* (the full subjective motivational set).
- ii. A member of *S*, *D*, will not give agent *A* a reason for *X*'ing if either the existence of *D* is dependent on false belief, or agent *A*'s belief in the relevance of *X*'ing to the satisfaction of *D* is false.
- iii.
 - a. Agent *A* may falsely believe an internal reason statement about himself, and (we can add)
 - b. agent *A* may not know some true internal reason statement about himself.
- iv. Internal reason statements can be discovered in deliberative reasoning.^{31,32}

Let us first briefly explore the term 'subjective motivational set', hereafter referred to as *S*. As we can see by surveying the four statements, in order to count as an internal reason, the belief in question has to first fall under *S*. In *S*, we find the full collection of an agent's desires as connected to their motivations. Reserving the term 'desires' for the contents of *S* would be limiting, however, as it can contain "such things as dispositions of evaluation, patterns of emotional reaction, [...] as they may be abstractly called, embodying commitments of the agent."³³ One's *S*, in short, can be called 'anything that, when relevant, could trigger (some level) of motivation in an agent.' Proposition (i) then means

²⁹ Philip Pettit and Michael Smith, "External Reasons", in *McDowell and his Critics* (Blackwell Publishing, 2006), eds. C. Macdonald and G. Macdonald, DOI: 10.1002/9780470776254.ch7, accessed March 2017, 142.

³⁰ Williams' method in *Internal and External Reasons* for proving his point is to put an impressive array of different attempts to understand the workings of purportedly 'external' reasons against his conception of internal reasons, and to see if there is any way that he can see the external reason working on its own. In the end, he cannot see this happen. For this paper, we have room only to present the most basic level of his point: the conclusion on internal reasons, which was taken over by McIntyre and Arpaly.

³¹ Paraphrased from Williams, "Internal and External Reasons", 102-104.

³² While the original quotation reads 'discovered', in this paper we will also be using the term 'produced' to refer to the process in which we speak of something becoming an 'internal reason'. I do not believe Williams would take notable offense to this.

³³ *Ibid.*, 105.

that one cannot have internal reasons to do things that one cannot be motivated (on any level) to do.

In proposition (ii) we can read Williams' firm resolution to aligning his theory with the notion of rationalisation, and not just explanation. Or, as Pettit & Smith put it: internal reasons must be able to justify an agent's conduct, not just explain it.³⁴ To sketch the difference: were we (third-personally) aware of an agent's false beliefs, we would be able to provide a perfectly coherent explanation of their actions based on those beliefs. We would, however, be unable to rationally *justify* them.

Proposition (iii) is of great importance to us. Part (a) is self-explanatory, while part (b) may appear puzzling. It will come back to us in force once we come to find whether or not our inverse akrasia theorists can claim that their reasons are internal, but also in chapter 6, when we research the 'rational relation' under (2) mentioned below. Williams actually proposes two different sources of (b). In short, they are that

1. An agent may be ignorant of a *fact* that should be able to potentially figure in an explanation of what the agent does while ignorant of that fact, such that on third-personal evaluation we would find the agent's action to be queer *if we did not* apply the missing fact to an explanation of their practical judgment.
2. An agent may be ignorant of an *element in S*. This element can become a reason for action X only if it is rationally related to X. If this element is in the unconscious, it may well not satisfy the conditions for internal reasons generation.³⁵

The above may be hard to follow. We will come back to it more than once over, but the broader point is that in this proposition, Williams notes the 'epistemic consequence' of (ii): if we create a theory focused on *justifying* and not just *explaining* beliefs, we must also leave room for human error and ignorance. After all, reading only (ii) clearly does not encapsulate all human action, meaning (iii) is focused first and foremost on filling these gaps. For this paper, the second source of (iii(b)) will be of special importance when we come to research the boundaries of internal reasons generation.

In (iv) the centrality of Williams' notion of 'deliberative reasoning' must be appreciated. This is because when speaking of the process of practical reasoning in a broader sense, Williams envisions something as follows: an agent deliberates on their *S*, adding,

³⁴ Paraphrased from Pettit & Smith, "External Reasons", 2. While they use the terms 'normative' versus 'motivating' reasons, incorporating these terms here would cause unnecessary confusion.

³⁵ Paraphrased from Williams, "Internal and External Reasons", 103. The full text of the sources will be presented later on in this chapter.

subtracting, and generally ordering certain elements of *S*. Through this process, we come to see what we have reason to do – in an internal sense.³⁶ Pettit & Smith’s refer to it as Williams’ ‘famous formula’, here paraphrased: ‘an agent, *A*, has an internal reason to *X* *only if she is motivated to X when she engages in deliberative reasoning*’³⁷ Keeping this in mind, we should take great care to note that while the proposition reads ‘*can* be discovered in deliberative reasoning’ and not ‘*must*’, this is not just one avenue of arriving at knowledge of our internal reasons. Internal reasons are the sole vocation of deliberative reasoning, and must be a product of that activity in order to count as such.

Before moving on to McIntyre and Arpaly, we should do well to realize that having an internal reason for a certain action is not tantamount to being all-out motivated to carry out that action. While Williams holds that generally, believing that a consideration is a reason for an action constitutes being motivated to carry out that action, the statement ‘*A* has reason to *X*’ *does not* by itself mean that *X* is the all-in preferred result of the deliberative process.³⁸ Agents have (internal) reasons to do many things, some of which more strongly reason-supported than others.³⁹

As has been noted in chapter 2, both McIntyre and Arpaly claim that the reasons motivating the agents in their inverse akrasia cases are internal reasons. We will discuss McIntyre and Arpaly separately on their own arguments for internal reasons, after which we will bring the content of the above proposition (iii(b)) to bear on their theories. Only then will we be able to see how McIntyre and Arpaly fare in claiming their agents are guided by internal reasons.

McIntyre’s claim to internal reasons

McIntyre makes a claim to internal reasons by pointing out that it would be relatively easy to describe her akratic cases using external reasons, but that instead, following Williams, she will only appeal to internal reasons in supporting their claim to rationality. There are two main steps that she takes to support her claim that Williams’ thesis would accommodate for undeliberated-upon reasons to still count as internal. Interestingly, despite directly referencing Williams multiple times, she makes no reference to the above proposition (iii(b)) in these two steps.

³⁶ *Ibid.*, 104, 109-110.

³⁷ Emphasis mine, Pettit & Smith, “External Reasons”, 143.

³⁸ Williams, “Internal and External Reasons”, 107.

³⁹ *Ibid.*, 104.

First, she follows Williams in that an unknown fact might motivate the agent, if it were brought to their attention. McIntyre then takes her first step by saying that a *known fact previously not seen as a reason* may become one by deliberating further. This is still allowed by Williams, who himself states that “an agent can come to see that he has reason to do something which he did not see he had reason to do at all”.⁴⁰

Carrying on from this, McIntyre states as a second step that “a reason [for an agent to act akratically] might exist without being so recognized”.⁴¹ She considers this to be possible, because Williams himself wishes to allow for a certain fluidity in a theory of practical reasoning, claiming it is “an imaginative [process, without] fixed boundaries on the continuum from rational thought to inspiration and conversion”.⁴² She concludes from this that, under Williams’ conception of internal reasons, it would be quite possible for an agent to have deliberated from a position of incomplete knowledge. This would then create a situation in which the agent’s practical judgment reflects their *belief* on what there is most reason to do, but does not lead to the action the agent *has* most reason to do.

Let us examine specifically the reasons that prompt the akratic action in McIntyre’s cases. Could these reasons be part of the agent’s *S*, and if so, does that mean they could count as internal reasons despite having not been deliberated on? In response to the first part of the question: possibly, yes. As was noted when discussing the nature of *S*, its contents can be anything ranging from simple desires to ‘patterns of emotional reaction’ or ‘various projects’.

In that line, McIntyre offers the example of Mary, an impatient employer, who, as part of a greater project to communicate effectively with her employees, resolves to always first give an employee a compliment before criticizing their conduct. One time, in a bout of akrasia, Mary launches into her criticism without prefacing it with a compliment. Had she attempted to give a compliment, however, it would have come out insincere and would have only further soured the employer-employee relationship. Seeing as what Mary *really* cared about was effective communication, McIntyre argues that this was still the *more rational* outcome, over the option that included the insincere, insulting compliment.⁴³

The astute reader may have noticed that the ‘patterns’ and ‘projects’ mentioned by Williams above are exactly the kind of thing that influenced Mary in her eventual akratic

⁴⁰ Ibid.

⁴¹ McIntyre, “Is Akratic Action Always Irrational?”, 386.

⁴² Williams, “Internal and External Reasons”, 110.

⁴³ McIntyre, “Is Akratic Action Always Irrational?”, 393-394.

move in the example given by McIntyre. As such, we can affirm that the reasons motivating inverse akrasia in McIntyre's case are possibly part of an agent's *S*.

Whether this qualifies these reasons as actual internal reasons is another thing entirely, since, as mentioned in our explanation of proposition (iv) on internal reasons, Williams' theory of practical reasoning is centred on the interplay between deliberative reasoning and an agent's *S*. In McIntyre's cases, this process ends on the agent's best judgment, forcing her to state that the *internal reasons* an agent had brought them to a best judgment, at which point *something* happened to sway the agent's motivation in the direction of an akratic action. This 'something' cannot be an external reason, as those do not exist. It appears we are in need of an avenue that allows us to claim that whatever influence swayed the agent's motivation in an inverse akrasia case could still have constituted an internal reason, but Williams' 'formula' (as mentioned under (iv)) is in the way: the presence of deliberation is, after all, key. It seems that McIntyre's claim to internal reasons fails.

Arpaly's claim to internal reasons

What is interesting about Arpaly's claim to internal reasons is that she never literally makes it. All that she mentions in her paper is that her argument "does not depend on any commitment to the existence of external reasons for action."⁴⁴ However, since we are here involved in an enquiry concerning the nature of the reasons that feature in cases of inverse akrasia, we are forced to draw conclusions from the statements Arpaly makes regardless. As such, actively distancing oneself from external reasons for action will be interpreted as equivalent to committing oneself to internal reasons instead – after all, reasons need to be of a certain nature. Having said that, we shall proceed much like we did above with McIntyre's account. First, we shall name what can be said in favour of Arpaly's claim, and then we shall see how far that takes her.

On the topic of rationality, and character of the connected reasons, all we explicitly have to base our judgment of Arpaly's account on is the following: "I will assume throughout this article that one only has a reason to act in a certain way to the extent that the relevant course of action satisfies one's desires given one's beliefs and that acting rationally, whatever it turns out to be, involves doing what one has overwhelming reasons to do, for these reasons."⁴⁵ One thing we can say in favour of Arpaly's claim to internal reasons is

⁴⁴ Arpaly, "On Acting Rationally Against One's Best Judgment", 492.

⁴⁵ Ibid.

that she never speaks of anything outside the agent when discussing the reasons they have for action. In fact, she explicitly values taking into account the *complete picture* of an agent's beliefs and desires. There even are marked similarities between how Arpaly describes the complete picture of her agents, and the way Williams speaks of an agent's *S*, such as when she describes an agent's 'overt psychology', that is to say his 'overall beliefs and desires' as the 'overwhelming reasons' to act akratically.⁴⁶ From this, we can say that the reasons Arpaly brings forth for her agents to act akratically can definitely fall under an agent's *S*.

However, contra Williams, Arpaly's entire project seems concerned with divorcing any necessary link between deliberation and rationality. This is showcased in the following sentence: "Sometimes, an agent who acts against his best judgment is acting on a desire which does not cohere with the rest of his beliefs and desires; other times, he is acting on a desire which wonderfully coheres with all his desires and beliefs— except one."⁴⁷ Note in this citation that the 'best judgment' Arpaly refers to is the deliberative product, and that it does not seem to hold any kind of privileged position whatsoever in comparison with whatever undeliberated 'desires and beliefs' we might otherwise have.

Other ways in which Arpaly attempts to marginalize the role of deliberation were mentioned in our chapter 2: sports players who act rationally but do so without deliberating, or 'dawning' cases in which we suddenly find ourselves to have shifted our judgments over time, without deliberation having been involved.

As said on (iv), the role of deliberation is central to Williams' thesis, and internal reasons for action are formed in an interplay between the agent's *S* and their deliberative capacities. Under a concept of deliberation such as the one Arpaly holds, where it has no special function in generating or discovering (internal) reasons for action, it seems impossible to maintain that the reasons she speaks of are internal in Williams' sense. We are thus forced to conclude that there is not much to be said for Arpaly's claim to internal reasons, based on her own words.

McIntyre, Arpaly, and Williams' proposition (iii(b))

One thing has so far not been granted the attention it deserves: Williams' proposition (iii(b)) about internal reason statements: "[Agent] *A may not know some true internal reason statement about himself*". We have so far seen both McIntyre and Arpaly fall short

⁴⁶ Ibid., 503.

⁴⁷ Ibid., 501.

of providing an adequate internal reasons account. For both, there was a best judgment that was apparently not the most rational option, which was subsequently deviated from on the basis of supplementary influences that were apparently *good reasons* but, importantly, *not* deliberated on.

However, depending on how far we can take this proposition by Williams, McIntyre and Arpaly may have found their redemption. This is for the following reason: internal reason statements unknown to the agent seem like they would necessarily be undeliberated upon; if they are, they could be the (internal) reasons for inverse akrasia that we have been looking for. How far, exactly, can we take this proposition? We noted the two possible sources of (iii(b)): an unknown *fact*, or an unregistered *element in S*. As we are now looking at (iii(b)) in more detail, we will need to look at the full text for both sources that Williams provides in his paper.

[‘[Agent] A may not know some true internal reason statement about himself’] comes from two different sources. One is that A may be ignorant of some fact such that if he did know it he would, in virtue of some element in S, be disposed to X: we can say that he has a reason to X, though he does not know it. For it to be the case that he actually has such a reason, however, it seems that the relevance of the unknown fact to his actions has to be fairly close and immediate; otherwise one merely says that A would have a reason to X if he knew the fact. I shall not pursue the question of the conditions for saying the one thing or the other, but it must be closely connected with the question of when the ignorance forms part of the explanation of what A actually does.

The second source of (iii) is that A may be ignorant of some element in S. But we should notice that an unknown element in S, D, will provide a reason for A to X only if X-ing is rationally related to D; that is to say, roughly, a project to X could be the answer to a deliberative question formed in part by D. If D is unknown to A because it is in the unconscious, it may well not satisfy this condition, although of course it may provide the reason why he X's, that is, may explain or help to explain his X-ing. In such cases, the X-ing may be related to D only symbolically.⁴⁸

In testing both sources, we will once more call upon the case of Huck. Let us first explore the first source: the fact-case. We will assume that some of the reasons we project on Huck for eventually freeing Jim were a part of his *S*: he had a soft heart, and appreciated

⁴⁸ Williams, “Internal and External Reasons”, 103.

friendship. Can we find a fact unknown to Huck that in interacting with his *S* may have provided him with a reason to free Jim? We might be able to imagine that if Huck had known that the abolitionists were going to win, he might not have felt so bad about helping a slave to his freedom. The problem here is that Williams requires this fact to be ‘fairly close and immediate’, while our example clearly is not.⁴⁹ I cannot personally imagine any ‘close and immediate’ fact that Huck overlooked in his practical reasoning that would have made any of his akratic reasons truly internal.

The second source is more puzzling: is it possible that there was some unregistered element of an agent’s *S*? Interestingly, both McIntyre and Arpaly propose something along these lines, albeit worded differently. For McIntyre, there is an agent who acts akratically in service of a broader project, while Arpaly proposes multiple agents of whom it is clear that they had desires and values beyond those that featured in their deliberation on what to do. However, the explanation also states that this element of *S* cannot be in the unconscious. What do McIntyre and Arpaly have to say to that?

It is hard to imagine Arpaly never having heard of this clause of Williams, since in her treatment of the akratic agent’s reasons she straddles exactly that fine line between the agents being *unaware*, yet not *unconscious* of the reasons that guide their akratic action. We can see her conundrum: should the agent be aware of the reason, we would reasonably expect them to act in accordance to it – making the action align with their practical judgment, and not akratic at all. However, complete unconsciousness looks like it doesn’t satisfy Williams’ conditions for internal reasons generation.

She makes some confusing, contradictory sounding remarks on the matter: speaking of Sam, a stressed out student who resolved to become a hermit, but then ended up akratically acting against that judgment, she says of his reasons that “they were not hidden from him in some black box of the unconscious, but were simply overlooked in his deliberations at the time.”⁵⁰ Later she says of the agents in her examples that they are “are moved to action rationally, for good reasons, as a legitimate response to good evidence

⁴⁹ As an example for when the fact-case may work, I imagine something along the lines of the following: a boy walking in a strange town towards a certain destination is determined to get there as quickly and efficiently as possible. He knows the general direction in which he should go, and at one point looks towards a small ledge that he might be able to jump in order to gain a significant shortcut. He decides, however, not to jump the ledge, despite the fact that he could have easily made it. The fact obscured from him that would have influenced his *S* (in this case: his disposition to get there as quickly as possible) such that he would have had reason to jump the ledge, was that he was looking at the ledge from a funny angle, making it appear higher than it actually was. We, third-personally knowing the fact obscured from our boy, can now properly make sense of the action that would have otherwise looked queer: why would he take the long route?

⁵⁰ Arpaly, “On Acting Rationally Against One’s Best Judgment”, 503.

of which they are aware, even though they do not deliberate their way into their actions”.⁵¹ I would like to ask of Arpaly to lay out the difference between ‘overlooking something’ and being ‘unaware of it’, and then going into how unawareness can still imply consciousness, but unfortunately we do not have the time for that. Instead, we must conclude that Arpaly does not manage to convince in her attempt to place her motivating reasons outside the unconscious.

McIntyre’s defence against unconsciousness is simpler: as said in our preliminary treatment of inverse akrasia cases, McIntyre believes her agents must have “almost attained the deliberative perspective from which the more rational judgment could be made.”⁵² Still, to her “the kind of rationality in question here is supposed to be an *internal matter*: what matters is what the agent was able to justify given what he knew, believed, or wanted then, not what could have been justified from some other perspective that he did not at that time have.”⁵³ We can draw from this that, despite the closeness of the agent in deliberatively finding the rational outcome, the only option for the kinds of motivating reasons that prompt her inverse akrasia cases is that they are things the agent wanted, but did not fully know at the time. Similar to Arpaly: if they did know, the agent would have incorporated the desire in their original deliberative process, altering the best judgment and eliminating the need for akrasia altogether. Lacking more refined linguistic tools, we will here, too, have to treat ‘unknown’ as equivalent to ‘unconscious’. This means that McIntyre’s case does not satisfy Williams’ requirements to (iii(b)) either.

It seems we are not left with much. Is there a way that we can save the reasons that prompt the inverse akratic agents from the asphyxiating realm of externalism? There must be, since Williams himself says that any external reason statement must really be an internal one, just ‘misleadingly expressed.’⁵⁴ We can also imagine the reasons prompting inverse akrasia to not be reasons at all, but in fact mere ‘influences’. However, that would entail dropping the entire notion of inverse akrasia as (more) rational, since we would intuitively take rational action to be in some way reason-guided. In the interest of keeping our project alive, we will continue to look for ways to mend the break between Williams’ conception of internal reasons and McIntyre and Arpaly’s inverse akrasia.⁵⁵

⁵¹ Ibid., 510.

⁵² McIntyre, “Is Akratic Action Always Irrational?”, 390.

⁵³ Ibid., 391.

⁵⁴ Williams, “Internal and External Reasons”, 111.

⁵⁵ In fact, drawing the conclusion (as we have here) that McIntyre and Arpaly’s accounts of reasons do not match with that of Williams would actually lead us to a point where we have to say that theirs are not reasons at all, since they appear to be neither internal nor external. For legibility’s sake, however, we will

4: Döring criticism, the advent of ‘intellectualism’

One author that helps us plot the course for making further sense of McIntyre and Arpaly’s apparently incompatibility with Williams’ conception of internal reasons is Sabine Döring. In her 2013 paper *Emotion, Autonomy and Weakness of Will* she argues *against* the rationality of (emotionally motivated) inverse akrasia, but in the course of doing so she inadvertently provides us with a fresh look at what might be the root cause of the abovementioned incompatibility. We will very briefly mention her main argument, but quickly turn to her conclusion on ‘intellectualism’, which is far more interesting for the project of this paper.

Her main argument goes along the following lines: emotional akrasia cannot be rational, because while emotions can bring us to act for good reasons and can even have us act rationally (such as when we spot an open gorilla cage in the zoo, and our fear prompts us to run for safety), Döring’s conception of practical reasoning is such that when there is *conflict* between our emotions and our (best) judgment (such as in cases of emotional akrasia) it is our judgment that provides us with our *best* reason.⁵⁶ On the whole, however, any rational action must be associated with an explicit judgment. The way that Döring then still allows for our emotions to prompt us to act for ‘good reasons’ is by positing that we often (such as in the gorilla cage example) make what she calls ‘non-inferential judgments’ – judgments that take place without deliberation being involved.

Here things get interesting. Her non-inferential judgments are an answer to a problem that she calls ‘intellectualism’: the notion that “all rationalisable action, including fast and habitual action, must consciously and explicitly be authorised by inference and deliberative reflection.”⁵⁷ Non-inferential judgments would in her view escape this trap, because, well, they are *non-inferential*.

I posit that they do not escape this trap, however, because Döring states that when our emotions seem to be deceiving us, we call upon a ‘tribunal of deliberation’ to determine whether or not an emotional impulse warrants our cognitive endorsement.⁵⁸ The problem with an idea like this ‘tribunal’ is that Döring also implies we have to be on watch as to whether our non-inferential judgments are worthy of endorsement or not. But this seems

continue to use the term ‘reasons’ to describe the influences that motivate the agents in the inverse akrasia cases described by McIntyre and Arpaly, and qualify the use of this term when necessary.

⁵⁶ Sabine Döring, “Emotion, Autonomy and Weakness of Will”, in *Autonomy and the Self*, Philosophical Studies series 118, eds. M. Kühler and N. Jelinek (Dordrecht: Springer, 2013), DOI 10.1007/978-94-007-4789-0_8, 186.

⁵⁷ *Ibid.*, 188.

⁵⁸ *Ibid.*, 185, 186.

to mean that we constantly checking whether or not our emotions can be trusted, or put differently; *in a perpetual state of deliberation anyway*. Döring warns us of intellectualism, but her theory may have difficulty outrunning it. Even for opponents of inverse akrasia, the place and importance of deliberation in practical reasoning seem to cause trouble.

5. Making ‘sense’ of deliberation: against intellectualism

Still, despite Döring’s theory apparently falling victim to it regardless, is there something her notion of ‘intellectualism’ can do for us in aiding our thinking on inverse akrasia and its underlying reasons? As it turns out: very much so. As this paper has attempted to bring into view, many of the problems that the inverse akrasia theorists run into when they speak of the potential rationality of akrasia stem from their conception of deliberation and its role in practical reasoning. In a twist on Döring’s notion of intellectualism, we will characterize McIntyre’s and Arpaly’s conception of deliberation as ‘intellectualist’: as something that takes time and effort in an explicit, intellectual way.

In order to properly do that, we will first briefly set out what is meant by an intellectualist conception of deliberation. We shall then show in which ways the discussed authors make their subscription to an intellectualist conception clear, and more importantly, in which ways this conception is to blame for the apparent stalemate in the inverse akrasia debate.

We will next explore other options: do we truly need to speak of deliberation in this way, or can we think of different ways to treat the concept? We will find that there is evidence for deliberation happening in less rigid ways, such as the fact that many agents, while not explicitly deliberating, seem to have a ‘sense’ that they should act in opposition to their best judgment. In chapter 6 we will once again compare these cases with Williams’ conception of internal reasons, where more specifically we will consider the nature of what Williams in his proposition (iii(b)) called a ‘rational relation’ between the action to be carried out, and the element in S motivating this action. Finally, we will address a number of issues that arise when ‘widening’ the concept of deliberation as a requirement for rational action.

Intellectualist deliberation

In our previous chapter on Döring, she mentions ‘intellectualism’ as the kind of view under which “all rationalisable action, including fast and habitual action, must consciously and explicitly be authorised by inference and deliberative reflection”.⁵⁹ She rejects this idea, and attempts to construct a theory of practical reasoning that works around it. Here we go take one step further back and say that what she is actually being blocked by is, as I would term it, *an intellectualist conception of deliberation*. As we can infer from the above

⁵⁹ Ibid., 188.

citation, Döring treats deliberation as a ‘conscious’ process, involving ‘explicit authorisation’ of action.

This, I would propose, can be seen as an intellectualist conception of deliberation: the idea that to deliberate on anything is to truly, consciously, intentionally and explicitly be pondering a certain belief or situation – preferably with a clear mind. Throughout the rest of this paper, we will use ‘explicit’ and ‘intellectualist’ deliberation interchangeably, as both terms carry the meaning I intend to convey.

Next, we will explore the relation that our notion of intellectualist deliberation has to McIntyre and Arpaly’s theories on inverse akrasia.

How does the intellectualist conception show itself?

First: how do McIntyre and Arpaly align themselves with an intellectualist conception of deliberation? For McIntyre, we can speak of an intellectualist conception, since in the examples she brings forward, while the agents have not deliberated *all the way* towards their eventual path of action, it was through *further deliberation* that they would be able to see the rationality of their akratic action. She says “there must be time [...] and even where there is time, there may not also be the opportunity”.⁶⁰ To McIntyre, deliberation is an explicit, intellectual effort: it is something requiring both *time* and *opportunity*. An agent needs to sit down and take a moment to reflect, in order to truly deliberate on their actions. Sadly, these conditions are not always met.

Arpaly makes her intellectualist conception of deliberation clear in other ways, most notably through her example of ‘fast action’ cases. In chapter 2 we briefly mentioned that she uses the example of sports players (say, tennis) as a further argument to divorce the link between rationality and deliberation, saying “if we were only to call people rational when their actions were caused by deliberation, we would have to call people rational considerably less often than we do.”⁶¹ The argument is as follows: if rationality and deliberation were necessarily linked, tennis matches would make no sense. Tennis players can make rational and irrational decisions in play, but never seem to ‘stop and think’ before they smash the ball over the net.

For us, the fact that she provides this as a central example is clear evidence that she considers deliberation to be something that requires a ‘stop and think’-moment. She never

⁶⁰ McIntyre, “Is Akratic Action Always Irrational?”, 398.

⁶¹ Arpaly, “On Acting Rationally Against One’s Best Judgment”, 506.

considers the option that some kind of deliberation may have happened regardless, and as such shows that she believes deliberation to be an explicit exercise. It cannot take place unless one takes the time for it, and explicitly directs their intellect in a deliberative effort.

The results of this need to be noted: as we brought up in chapter 3, both McIntyre and Arpaly end up at some point having to suppose a strange kind of semi-consciousness in their agents that allows them to act in accordance with the more rational path of action, despite not having explicitly deliberated on these paths. An intellectualist conception of deliberation is to blame for sandwiching inverse akratic agents between consciousness and unconsciousness of the reasons for their akratic action. If the agents were conscious of their reasons, why did they not enter their deliberation on their way to a best judgment? Because their deliberation was constrained by the requirements of intellectualism: it had to take place consciously and explicitly. Can we imagine it any other way?

Looking outside intellectualism: the evidence

When we consider the option of deliberation happening in other, less-than-explicit ways, we may find new consequences emerging. Those will be considered near the end of this paper. In this part, I simply aim to present some examples from the texts of McIntyre and Arpaly in which something seems to be going on that seriously *looks* like deliberation, but according to the authors, isn't.

Recall McIntyre's case of Mary, the impatient employer. Upon committing her act of inverse akrasia, she experiences an 'uneasy feeling', which McIntyre implies is a by-product of her "[responding] to the consideration that it just didn't seem right to add the appreciative comment."⁶² Something not '*seeming right*' would appear to imply that, on a certain level, Mary had considered the implications of her actions *beyond* her original best judgment. McIntyre then argues that what *really* prompted Mary to act akratically "was her *sense* that a lame appreciative comment would have [failed]."⁶³ This talk of a 'sense' again looks like it might mean that some consideration is taking place inside Mary, other than her explicit deliberation.

Next, I wish to consider McIntyre's following argument: while still maintaining her belief that properly carrying out her resolution to compliment her staff would be best, Mary nevertheless "*decided*" that she should "revise her practical aspirations downward".⁶⁴ To

⁶² McIntyre, "Is Akratic Action Always Irrational?", 393.

⁶³ Emphasis mine, *ibid.*, 394.

⁶⁴ Emphasis mine, *ibid.*, 395.

speak here of a ‘decision’ or a ‘revision’ is very strange from a viewpoint of intellectualist deliberation. Clearly there was in her example neither time nor opportunity to commit to an explicit moment of decisionmaking – and if there was, would this even be a case of akrasia? If an agent had the time to alter their decision based on good reasons, they can be presumed to have had the time to alter their best judgment. Could this ‘decision’ then have been influenced by deliberation, despite there not having been the time to ponder the situation?

For Arpaly, her ‘dawning’ cases are most interesting in this regard. In short, they are cases in which an agent suddenly has something ‘dawn’ on them that on its own is a completely rational belief, but was never deliberated on. In this light, she brings forth the example of Emily, who has always wanted to pursue a PhD in Chemistry, but has things not quite going her way. Interestingly, Arpaly frames her case as both one of classic (inverse) akrasia but also as a ‘dawning’ case. In this case, Emily finds herself experiencing a growing “sense of dissatisfaction”, and, when akratically quitting the program, an “inexplicable sense of relief”.⁶⁵

In the dawning case however, Arpaly describes Emily as ‘coming to her senses’ about her commitment to the PhD program, stating that “her conviction [...] disappeared, as if it had evaporated.”⁶⁶ Arpaly describes the process by which this happens as follows:

Gradually, she encounters more and more data that are inconsistent with her conviction, though it still fails to affect her deliberation, and at a certain point, not easy to define, she realizes that she doesn’t want to be in the program anymore. Alternately, she realizes suddenly that for a while she has been very unsure about what used to be obvious—the chemistry program being the right place for her. At this stage, Emily can articulate very good reasons for leaving. The facts she recites as reasons had to dawn on her gradually (rather than, say, be found out by laborious research), but since they are indeed good reasons, no one doubts the rationality of her decision.⁶⁷

Arpaly concludes from this that something undefined, outside of deliberation, produced reasons for Emily to quit the program. I find that to be a strange conclusion. Is it not far more sensible to explain the above as a deliberative process of reason generation? Then, instead of claiming the existence of a nameless process of fact-registration that worked on

⁶⁵ Arpaly, “On Acting Rationally Against One’s Best Judgment”, 504.

⁶⁶ *Ibid.*, 508.

⁶⁷ *Ibid.*, 509.

Emily's explicit convictions until they could be swayed, we can conclude that Emily's deliberative capacities worked just fine in eventually prodding her towards the rational course of action. They just didn't work as an explicit, intellectual exercise.

6. Williams, and the ‘rational relation’

At exactly this point, we need to once again consider what Williams has to say on the matter. The reason for that is quite simple: looking outside of explicit deliberation is all well and good, but we still need to make sure the reasons involved can be considered ‘internal’. After all, we are trying to prove the rationality of inverse akratic agents, and for that we need internal reasons generation – a process in which *deliberation is key*. The question is then: would Williams allow for any kind of deliberation besides intellectualist deliberation? Perhaps he would.

Part of the point about Williams’ vagueness regarding deliberation has been mentioned before when we presented McIntyre making her ‘claim to internal reasons’. However, it is worth presenting in full here. He makes no secret of this vagueness: on why it is unclear where, in practical reasoning, rational deliberation will bring an agent, he says

it is unclear, and I regard it as a basically desirable feature of a theory of practical reasoning that it should preserve and account for that unclarity. There is an essential indeterminacy in what can be counted a rational deliberative process. Practical reasoning is a heuristic process, and an imaginative one, and there are no fixed boundaries on the continuum from rational thought to inspiration and conversion.⁶⁸

Particularly the last sentence strikes a chord for us. On the question of what counts as a ‘proper’ deliberative process, Williams gladly allows for a great diversity, stating that his internal conception of reasons “shows that there is a wider range of states, and a less determinate one, than one might have supposed, which can be counted as [an agent] having a reason [for action].”⁶⁹

Still, while what we said here on McIntyre, Arpaly and Williams sounds like it could go together, we need to keep in mind the implications. If we step away from the idea of deliberation always being an explicit, intellectual effort, we appear to automatically be stepping into the territory of un- or subconscious deliberation. Looking back to chapter 3 and McIntyre/Arpaly’s ‘claim to internal reasons’, Williams said something on that matter, too.

⁶⁸ Williams, “Internal and External Reasons”, 110.

⁶⁹ Ibid.

The necessity of a 'rational relation'

Recollecting Williams' statement that an agent "may not know some true internal reason statement about himself" may put us in a good mood.⁷⁰ However, we also noted in our earlier discussion on proposition (iii(b)) about internal reasons that the notion of unconscious elements of one's subjective motivational set giving rise to internal reasons "may well not satisfy [the conditions for internal reasons provision]".⁷¹ What are we to make of this?

A few things spring to mind: first, we shall clarify the status of Williams' above statements in relation to McIntyre and Arpaly. We shall next see that special importance is lent in Williams to the notion of the unconscious elements of one's S being 'rationally related' to the action that the agent eventually carried out. Recall here the second origin of clause (iii(b)),

A may be ignorant of some element in S. But we should notice that an unknown element in S, D, will provide a reason for A to X only if X-ing is *rationally related* to D; that is to say, roughly, a project to X could be the answer to a deliberative question formed in part by D. If D is unknown to A because it is in the unconscious, it may well not satisfy this condition, although of course it may provide the reason why he X's, that is, may explain or help to explain his X-ing. In such cases, the X-ing may be related to D only symbolically.⁷²

Without coming to a complete conclusion, we shall finally in this chapter make a nudge as to the direction in which we should be thinking when considering this 'rational relation', and briefly run ahead on where this direction could be taking us.

We have made previous mention of Williams' statements on the (un)consciousness of internal reasons with regard to McIntyre and Arpaly in chapter 3. What we haven't done yet, however, is make the interesting observation that neither author ever made mention of these statements in their papers. How can that be? Both McIntyre and Arpaly grapple with exactly the same question that Williams is attempting to deal with in his own paper at that stage: can un- or subconscious elements in S give rise to un- or subconscious reasons for action?

⁷⁰ Ibid., 103.

⁷¹ Ibid.

⁷² Emphasis mine, *ibid.*

I can see two possible elements playing a part in their failure to mention this clause of Williams:

1. It does not particularly bode well for their theories should they be unable to claim some kind of (internal) reason generation outside of explicit consciousness, and
2. they may have just had a very hard time understanding what exactly Williams means, here. He seems to deny unconscious elements in S producing reasons for action, but does not do so very strongly (they ‘may well not’). He also adds the confusing remark that this element in S might still provide the reason for an agent’s action, but only in an explanatory sense of the word. The agent’s action is then related only ‘symbolically’ to the unconscious element in S.⁷³⁷⁴

If we are to even consider the option of stepping away from the notion of deliberation as an intellectualist effort, we must provide some kind of direction to the idea of a ‘rational relation’ between the unconscious element of S and the action it motivates. If we cannot do this, we have nothing to support the rationality of agents in inverse akrasia cases, and we are back to the same ‘aware-but-not-quite-enough’ paradox that arose earlier with regards to McIntyre and Arpaly. Can we think of a ‘rational relation’ strong enough for Williams to sway his ‘may well not’?

The possible nature of a ‘rational relation’

Considering the examples given already concerning the possibility of non-explicit deliberation that were given in the previous chapter (when discussing the ‘evidence’ for non-explicit deliberation), I would like to suggest that this ‘rational relation’ finds itself in something of a ‘sense’ felt by the agents in inverse akrasia cases that they should act in opposition to their best judgment.

We should first, however, make clear that the word ‘sense’ is mentioned in different contexts in our examples, both in the ones that have been mentioned and those that follow. On the one hand we have the above mentioned ‘sense’ as being felt (by Mary, for example), but on the other hand inverse akratic agents are also from time to time described as

⁷³ Ibid., 103.

⁷⁴ Taken at face value, one is tempted to assume that Williams himself knew that both a hard ‘yes’ and a hard ‘no’ lead to positions for which he did not want to, at that point, take on the burden of defence (as both allowing and disallowing unconscious internal reason generation have strong consequences, as we see throughout this paper).

‘coming to their senses’, which seems to imply something more along the lines of the effect of a ‘common sense’ providing our rational relation.⁷⁵

With this final way of speaking of the ‘sense’ in mind we can once again mention Arpaly’s example of Sam, the student who akratically decided to not become a hermit after all. When it comes down to his final moment of akrasia, he is described to not be a ‘weak-willed man’, but instead one for whom his “common sense prevails in real life.”⁷⁶ We can very well imagine it having been the kind of nagging feeling (the ‘sense’ that he should act differently) that eventually prompted Sam to akratically drop his ill-conceived plan, and thus he ‘came to his senses’.

If we put forward this ‘sense’ as the basis for Williams’ ‘rational relation’, what more does our account need? It still seems to miss some idea of what level, or what kind of consciousness we suppose in our agents when this ‘sense’ occurs. We can’t even say whether ‘consciousness’ is the right term for what we are looking for. ‘Awareness’ seems to be closer, but also somewhat off the mark. Both consciousness and awareness are terms that seem to imply the kind of front-of-the-mind presence of reasons that are in these cases evidently *not* at the front of the agents’ minds. Then again, unconsciousness or unawareness do not seem right either – after all, there is a *sense* being felt.

As a minimum requirement, we should keep in mind that this ‘sense’ does not seem to pop up by accident. It is not something that, out of the blue, occurs to an agent. A sense that we should act differently, or are on the wrong path, seems not to pop up when we are trying to wake up over a cup of coffee, or making small talk with friends, but rather when we are on some level involved in an exercise of ‘doing the right thing’.

What I am talking about is a kind of minimal effort: a state of wakefulness, and any reason whatsoever to want to ‘make the right call’, no matter how fast or fleeting the moment of making the call is, and no matter how much our intellectually deliberated reasons seem to have already ‘closed the case’ (as in Mary, who had a strong, intellectually deliberated resolution to preface her correction with a compliment – but didn’t). Does this imply awareness, or unawareness? Anticlimactically, it seems to be a little bit of both.

⁷⁵ And if we wanted to go further, we could even describe our deliberative capacities as a human ‘sense’ in and of itself. However, this implies a certain moral intuitionism that we cannot, in this paper, expand upon.

⁷⁶ Arpaly, “On Acting Rationally Against One’s Best Judgment”, 503.

7. Inverse akrasia revisited: solutions and problems

Let us recap for a moment. Having found significant problems with the intellectualist way of treating deliberation, we set out to see if there was any reason to look at deliberation through a slightly wider lens. In doing so, we encountered many agents experiencing an ill-defined ‘sense’ prefacing their moment of inverse akrasia, and also found them to be at least directing some attention towards an exercise of wanting to do the right thing. What does this mean for our inverse akrasia cases?

It means that we can now speak of deliberation in a way that is less rigid than before, allowing for, as Williams put it, a ‘wider range of states’ to count as a deliberative process. This means that we avoid the paradox of being aware of reasons, but not aware enough to have them feature in our best judgment. Mary the impatient employer, Sam the misguided student, Emily the doubting PhD candidate, and even our original Huck, the fateful friend, could all on some level have been deliberating towards their final path of action, despite not being fully aware of their own reasoning at the time. Seeing that Williams may not have wanted to commit himself to either a ‘yes’ or a ‘no’ on the topic of unconscious internal reasons generation, we can now say that we, in this paper, have at least explored the option of a (soft) ‘yes’. There are consequences to this exploration, however.

Remember: we have originally been discussing the potential *rationality* of (inverse) akrasia. That means that by leaving room for the possibility of some deliberative process producing reasons without us knowing about it, we are committing ourselves to saying that these reasons, that are being produced outside of our explicit awareness (and thus, explicit judgment) are *good* reasons.⁷⁷ After all, inverse akrasia cases involve taking the more rational path of action against one’s best judgment. However, is our ‘best judgment’ in these cases still our *best judgment*? Or, put differently: if we consider deliberation to be a part of our practical reasoning, then agents in inverse akrasia cases akratically commit to actions that are still in a way governed by their practical reasoning. Considering the fact that the presence of a *defect in practical reasoning* is, perhaps, *the* central characteristic of akrasia, *can we still call these agents truly akratic?*

We can decide not to rock the boat by saying that they are still akratic: akrasia, then, is about acting against your *explicit, intellectualist* judgment. This is not very satisfying, as

⁷⁷ Used here as explained earlier, with ‘good reasons’ being the kind of reasons that could potentially make akratic action more rational than following one’s best judgment.

it brings us back to where we started: with akratic agents seemingly acting rationally, but with no way to call them rational or praise them for their prudence in their actions. After all, they made a judgment to the best of their (explicit) abilities, and failed to follow through on it. They showed weakness of will.

Perhaps, then, we want to say that these inverse akratic agents are, in fact, not akratic at all: their practical reasoning led them to the rational course of action, which they took. Although this seems like the obvious solution at first glance, it leads to further problems down the line. How, for example, are we supposed to think of *regular* akrasia in a world in which you cannot always be aware of your own deliberative processes and where they might take you? In this case, can we still call anyone irrational at all? My akratic action to go home right now and cook dinner may be part of a wider deliberative effort, aimed at optimal intake of nutrients, so that I may be able to have a fruitful day of work tomorrow. Or, I might just be lazy and looking forward to a nice home-cooked meal.

There does not seem to be any way to tell whether I lack the *enkrasia* or strength of will to stick to my resolution, or whether in the back of my mind I am guided by a rational, deliberative process. The only things offered so far in this paper are the aforementioned ‘sense’, and a minimum effort requirement vis-à-vis the agent. Of the requirement, we will often be able to easily say that it is fulfilled; it is, after all, not very demanding. Of the sense, we really cannot say much at all: after all, it is up to the agent to say that he or she felt it. Even then, should the agent have felt it very strongly, we might once again be justified in wondering why he or she didn’t alter their original judgment in the face of such strong feelings of doubt. Worse, however, would be an agent attempting to justify their akrasia by appealing to such a sense. It seems we may have inadvertently created a very tempting avenue for rationalisation - as if we needed another one!

Closing remarks

Everything coming together, we can make some closing remarks. Having started with the question ‘is inverse akrasia rational?’, we looked to be hitting a premature yet solid ‘no’. After all, neither McIntyre nor Arpaly, despite following his conclusions, seemed to be able to satisfy the conditions set by Williams for internal reasons generation. This would have been an unsatisfactory answer, however, as we still have these well-described cases of agents acting akratically, but doing what is without a doubt ‘the right thing’. Looking for a way to ‘save’ these agents’ rationality, we turned to the concept of deliberation, and whether or not we may have grounds to look at it from a slightly less rigid point of view.

That, I feel, should be this paper's main takeaway. Not the bickering over different types of rationality, or 'goodness', and not the great ethical questions, but that first and foremost we should decide *what we (can) expect from our agents*. Focusing the debate on akrasia and practical reasoning on the stories of single agents brings to the fore a notion that, as we have seen, has been somewhat glossed over by authors on inverse akrasia: deliberation. It is only after analysing and appreciating this human capacity for reason, with all of its facets, that we can truly start speaking of the 'good' or the 'bad', and the 'rational' or the 'irrational'.

Bibliography

- Aristotle. *Nicomachean Ethics*. Cited in Annemarie Kalis. *Failures of Agency*.
- Arpaly, Nomy. "On Acting Rationally Against One's Best Judgment". In *Ethics*, 110 (Chicago: The University of Chicago Press, 2000). 488-513.
- , and Timothy Schroeder. "Praise, Blame and the Whole Self". *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, Vol. 93, No. 2 (1999). 161-188.
- Bennett, Jonathan. "The Conscience of Huckleberry Finn". In *Philosophy*, 49 (1974). 123–134.
- Davidson, Donald. "How Is Weakness of the Will Possible?". In *Essays on Actions and Events* (Oxford: Clarendon Press, 2001). 21-42.
- Döring, Sabine. "Emotion, Autonomy and Weakness of Will". In *Autonomy and the Self*, Philosophical Studies series 118. Edited by M. Kühler and N. Jelinek (Dordrecht: Springer, 2013). DOI 10.1007/978-94-007-4789-0_8. Accessed May 2017. 173-190.
- Kalis, Annemarie. *Failures of Agency*. Vol. 57 in series 'Quaestiones Infnitae' (Utrecht: Zeno Institute of Philosophy Publications, 2009).
- McIntyre, Alison. "Is Akratic Action Always Irrational?". In *Identity, Character, and Morality: Essays in Moral Psychology*. Edited by O. Flanagan & A. Oksenberg Rorty (Cambridge, Massachusetts: MIT Press, 1990). 379-400.
- Pettit, Philip, and Michael Smith. "External Reasons". In *McDowell and his Critics* (Blackwell Publishing, 2006). Edited by C. Macdonald and G. Macdonald. DOI: 10.1002/9780470776254.ch7. Accessed March 2017. 142-170.
- Plato. *Protagoras*. Cited in Sarah Stroud. "Weakness of Will". *The Stanford Encyclopedia of Philosophy* (Spring 2014 Edition). Edited by Edward N. Zalta. <https://plato.stanford.edu/archives/spr2014/entries/weakness-will/>. Accessed April 2017.
- Williams, Bernard. "Internal and External Reasons". In *Moral Luck: Philosophical Papers 1973-1980* (Cambridge: Cambridge University Press, 1981). 101-113.