



Explaining Free Will by Rational Abilities

Frank Hofmann¹ 

Accepted: 3 February 2022 / Published online: 22 March 2022
© The Author(s) 2022

Abstract

In this paper I present an account of the rational abilities that make our decisions free. Following the lead of new dispositionalists, a leeway account of free decisions is developed, and the rational abilities that ground our abilities to decide otherwise are described in detail. A main result will be that the best account of the relevant rational abilities makes them *two-way* abilities: abilities to decide to do or not to do *x* in accordance with one's apparent reasons. Dispositionalism about rational abilities will be vindicated, since it helps to see why and how these two-way abilities entail an ability to decide otherwise and, thus, free will.

1 Introduction

A promising and important view about free will is the *rationalist leeway* view according to which free will is grounded in certain rational abilities that enable us to do otherwise. Here, and throughout this paper, 'free will' will be understood as the control component of moral accountability, the component which makes it the case that something is 'up to us' or 'in our control'. Here, 'moral accountability' is understood as the status of being an appropriate target of blame and praise. It is the kind of moral responsibility that comes with the engaged reactive attitudes that Peter Strawson has brought into the focus of attention in his classic "Freedom and Resentment" (Strawson 1962). The cash value of free will, then, is usually

✉ Frank Hofmann
frank.hofmann@uni.lu

¹ Institute of philosophy, FSLHASE, University of Luxembourg, 11, Porte des Sciences, L-4366, Esch-Belval, Luxembourg

taken to consist in the ability to do otherwise (the ADO, for short).¹ So the view is a *leeway* view as well as a *rationalist* view.²

For a number of years, several philosophers have been developing accounts that belong to this kind, more or less at least. One of the best worked-out version of this approach is Kadri Vihvelin's dispositionalist account, according to which the abilities that make us free are rational abilities and these are, like all abilities, identical to (or grounded in) dispositions (cf. Vihvelin 2013). Similarly, Michael Smith has tried to account for responsibility in terms of rational abilities, called 'rational capacities' by him (Smith 2003). Susan Wolf and Dana Nelkin have put forward views that focus on the ability to do the right thing for the right reason (cf. Wolf 1990, Nelkin (2011)).³

In this paper, the focus will be on *free decision*. Plausibly, free intentional action requires free decision (free choice), i.e., the free formation of the intention to act in a certain way.⁴ If the intention was not formed freely, it is hard to see how the action guided by it could count as free. In the case of free decision, according to the rationalist leeway view, we have the ability to decide otherwise by having a certain rational ability related to decision. Vihvelin holds that it is the ability to decide on the basis of reasons. In certain situations of choice (called 'Moorean Choice situations' by Vihvelin) the subject.

“deliberates, decides, and acts successfully on the basis of her decision and has, at some time before she decides, the *wide* as well as the *narrow* ability to decide and to act otherwise. She has the narrow ability to decide otherwise by having the narrow ability to make decisions on the basis of reasons; she has the narrow ability to act otherwise by having the narrow ability to act on the basis of her decision.” (Vihvelin 2013, p. 169; emphasis in the original).

The appeal to an 'ability to make decisions on the basis of reasons' is what makes her view *rationalist* in the relevant sense. But what is it to decide on the basis of reasons?, we may ask. Vihvelin does not say much about it, and she does not provide a general account of reasons and the ability to decide on the basis of reasons. The main question here is not so much the (deep) metaphysics of these abilities, since they may very well be subsumed under the general dispositionalist account of abilities. (In my view, Vihvelin's arguments and defense of the dispositionalist account are successful.)⁵ The question is rather *what the successful*

¹ As Clarke notes, even if free will did not require the ADO, it would still be a very interesting question whether we have the ability to do otherwise. Cf. Clarke (2015), n. 39.

² We can weaken this a bit by saying that an ability to do otherwise is only needed in the bad cases in which one does the wrong thing. This would be an asymmetrical leeway view, but still a leeway view. For an asymmetric view see Nelkin (2011) and Wolf (1990).

³ Surely there are many more proposals that are significantly similar, and there are many traditional views that could be seen as precursors, such as Aristotle's or Kant's view.

⁴ At least since Widerker (1995), the focus is on the free formation of intention (choice, decision). As Widerker nicely puts it, “mental acts, such as deciding, choosing, undertaking, forming an intention ... constitute the basic *loci* of moral responsibility” (Widerker 1995, p. 247).—There remains then a question of what to say about actions that are not intentional, supposing there are any (like, perhaps, jumping up simply out of joy and without any further intended effect). Probably, one could make headway by looking at what's possible: a corresponding intention could and would have been formed if the subject had deliberated, and then this intention could and would have been freely formed.

⁵ However, there is a part of Vihvelin's account that I do not accept, namely, her endorsement of a refined Conditional Analysis (RDA) à la Lewis. More on this below, in footnote 26.

manifestation consists in, the so-called ‘deciding on the basis of reasons’. What does it mean to ‘*decide on the basis of reasons*’? If we look at Michael Smith’s account, rational abilities are appealed to (called ‘capacities’ by Smith, which I take to be synonymous with ‘abilities’). But similarly, we can ask what these ‘rational abilities’ have as their successful manifestations. Are they the ‘rational decisions’? But how are rational decisions related to decisions based on reasons? Susan Wolf and Dana Nelkin, finally, speak of ‘doing the right thing for the right reasons’. In the case of decisions, this would mean that in order to decide freely, one has to have the ability to decide the right thing for the right reasons. Again, it is not entirely clear what ‘deciding rightly for the right reasons’ means. What are the reasons that we are talking about here? Are they objective normative, justifying reasons or rather some variety of subjective reasons? And what do we do, what do we attain if we successfully ‘make the right decision for the right reasons’? These questions surrounding reasons and rationality are what the first part of this paper will address. To some extent, the rationalist leeway view of free will is *incomplete* without good answers to the just mentioned questions. Some work remains to be done even if so far, the existing rationalist leeway accounts are very promising and can help us to preserve many important intuitions, like the compatibilist intuition that free will can be had in a deterministic world, just to mention a very prominent feature. Here I would like to do some of this remaining work and make the rationalist leeway view more complete.

The second part of this paper will be concerned with a dispositionalist account of these rational abilities, following recent ‘new dispositionalists’ like Michael Smith, Michael Fara, and Kadri Vihvelin. Here the focus will be the question of whether the relevant rational abilities have to be *one-way* or *two-way* dispositions (powers). The result will be that we indeed need *two-way powers* in order to ensure that the relevant rational ability guarantees the ability to do otherwise. The preferred dispositionalist version of the rationalist leeway account will thus be a two-way power account, since free will has essentially been identified as the ability to do otherwise. To some this may come as a surprise, and it is surely not an altogether obvious point. So doing the careful work of spelling out the dispositionalist picture is worth the effort. And without it, the first part would be much less significant since it leaves out an argument to the effect that the envisaged rational abilities indeed make us free.

2 What are the Relevant Rational Abilities to Do?

The main question of the first part of this paper is which ‘reasons’ we should appeal to and how rational abilities are to be understood and how, in particular, they are connected to these ‘reasons’. Since in the recent three decades quite a lot of views about rationality and reasons have been proposed and discussed, a complete investigation considering all of these views would go much beyond the scope of this paper. Instead, therefore, I will present one independently promising view of reasons and rationality and argue that it fulfills the job of supplementing the rationalist leeway view of free will in a satisfactory way.

How do we pick?—The main desideratum that any candidate view has to satisfy, is the intuition that even a subject that is very much disconnected from external reality—e.g., a brain in a vat—could behave rationally. This desideratum is the ‘*biv intuition*’, as we can call it. Consider Biff, an ordinary subject whose brain has been temporarily envatted unbeknownst to him. Biff’s perceptual experiences are hallucinatory, triggered by stimuli

coming from a controlling computer. If Biff has a hallucinatory perceptual experience as of a green apple in front of him, he still rationally responds by forming the belief that there is a green apple in front of him, not realizing his bad conditions for perceptual belief formation. Similarly, since he desires to eat an apple, he decides to grab and eat what he takes to be an apple; he forms the intention to grab and eat it. He engages in ordinary instrumental rational reasoning, it seems, and he clearly might be able to do so.

As Biff's case shows, nothing prevents a brain in a vat from having and exercising some form of rationality. (Of course, a biv need not behave rationally—but it's not impossible.) This is a strong intuition, the biv intuition. Many epistemologists have emphasized this point, in the case of *belief*, and used it to distinguish between justification and rationality, where justification arguably requires some factive mental states whereas rationality does not.⁶ And they have pointed out that among the brains in vats, we can distinguish between those who behave rationally, like Biff, and those who do not behave rationally—even if both do not achieve some further status like justification. The same should be true of decisions, which we can take to consist in the formation of intentions to act in a certain way.⁷ In other words, rationality does *not* presuppose *factive* mental states that make the subject possess genuine reasons in the sense of facts that genuinely favor certain attitudes and actions. So if we want to say that the ability to do otherwise is grounded in the rational ability to decide on the basis of reasons, it cannot be genuine reasons whose possession requires factive mental states like knowledge. Here the genuine reasons are, of course, the famous facts that favor various ways of acting, familiar from the works of Parfit, Dancy, Scanlon, and Skorupski, most importantly.⁸

Apparent reasons to the rescue! Apparent reasons are either genuine reasons or considerations that would be reasons if veridical. Incidentally, Derek Parfit has appealed to apparent reasons in his account of rationality.⁹ Basically, then, rational abilities can be understood as abilities to decide in accordance with apparent reasons. This is my proposal for how to complete the rationalist leeway view of free will. In the following I will spell out this view further and describe why and how it is preferable to some alternatives. We must have a precise and clear understanding of the manifestation of rational abilities in order to arrive at

⁶ See, for example, Weatherson (2008), Williamson (2017), and Williamson (forthcoming). Recently, in the debate on higher-order evidence, the distinction between rationality and justification is considered to lie at the root of the so-called 'acratia trilemma'. Cf. Kappel (2019) and other essays in Skipper, Steglich Petersen (2019). Interestingly, Littlejohn has argued quite convincingly that given the supervenience of rationality on internal properties ('standard mentalism', as he calls it), reasons and rationality cannot always point into the same direction—for both the practical and the theoretical domain. Cf. Littlejohn (2019).

⁷ Following philosophers like Pamela Hieronymi, we can say that forming intentions is settling the question of what to do, and this is exactly the same as taking a decision. Cf., for example, Hieronymi 2013, 2009. The notion of a decision that I am relying on here might not entirely match the ordinary life notion of a decision since it does not require any (conscious) deliberation or consideration of alternative options. One can form intentions that are in line with one's (apparent) reasons, I take it, without considering alternatives. The ability to 'decide' otherwise that makes us free is not necessarily a deliberative one. All we need is the ability to form alternative intentions. Therefore, we can equate decision and intention forming for the present purposes.—I am grateful to an anonymous referee for drawing my attention to this point.

⁸ The generic assumption then can be called 'factualism' about reasons, following Alvarez (2016).

⁹ Cf. Parfit (2011), p. 35.

a full, complete view of free will. To take *accordance with apparent reasons* as having this role is the missing idea that can do the job.¹⁰

To have a running example, we can use Parfit's hotel case of a person, call her 'Delia', who is in the first floor of a hotel and falsely believes that there is a dangerous fire closeby (cf. Parfit 1997). Intuitively, her jumping out of the window is rational, even though Delia lacks a genuine reason for doing so, since there is in fact no dangerous fire.

Let us first take a look at the Wolf-Nelkin view. Basically, it says that we decide freely if we have the ability to *decide the right thing for the right reasons*.¹¹ A very natural interpretation of this would be that the 'right reasons' are genuine reasons, i.e., reasons that really favor the decision (or the course of action that one decides for). This cannot be right, however, since it would exclude the brain in a vat from freedom, and to do so is highly counterintuitive. The brain in a vat may not be able to *act otherwise bodily* (even though it might mentally). But as we have already noticed, freedom of *decision* is the crucial issue. And because the biv is not able to possess genuine reasons (mostly, at least), it is not able to decide for the right reasons. Deciding for the right reasons requires possession of genuine reasons (in a very natural understanding of this notion). One is able to decide for the right reason only if one is able to possess the right reasons. But plausibly, in all relevantly close worlds, the biv is still envatted (even though undergoing different stimuli generated by the computer that is feeding it with inputs) and could not possess genuine reasons. The biv is screened off from genuine reasons because it is no position to enter into factive mental states (about particular contingent features of its environment, at least).¹²

Vihvelin's formulation is very similar. She speaks of making decisions "on the basis of reasons". Again, a natural interpretation is that "reasons" are genuine reasons. But also again, this cannot be so since it would exclude the biv.

Things are slightly different with Michael Smith's view. He speaks of rational abilities simply (which he calls 'rational capacities') and in general, he thinks of rationality in terms of coherence (cf. Smith 1994, 2003). How these are related to reasons is not specified by Smith. So in a sense, his view is also not complete, or at least leaves an interesting question open, namely, in what relation free will stands to reasons.

For all three accounts, then, the general argument given above—the 'argument from the biv intuition'—shows some need for revision or reinterpretation, or at least some further

¹⁰ Ability *ascriptions* are often context-sensitive and implicit (thus not containing context-invariant explicit representations of some aspects of the manifestation). Here, we are concerned with *abilities* themselves, not with ability *ascriptions*. Thus, it is necessary to provide a complete explicit description of the manifestations. This is what I am trying to do in this paper.

¹¹ For sure, I have modified the account which primarily is an account of action, not of decision. But this adaptation seems legitimate and what we need to focus on free decisions.

¹² In this context, a quick look at the case of the dictator Jojo, introduced by Susan Wolf in Wolf (1987), might be helpful. Very briefly, Jojo is the son of a dictator who has been raised in a way that has made him into a cruel dictator too. He is unable to respect the (moral) reasons for not acting as he does, and so he is not acting for genuine reasons, in general. Now, the main point about Jojo is that he is very much different from the BIVs we are considering. By stipulation, we take the envatted brains to be brains of *ordinary, sane* people. If these BIVs were put in Jojo's situation, they would decide very differently (in a sane way); they would decide rationally. Jojo does not decide rationally. The difference is explained by a rational ability that these BIVs have and that Jojo lacks: to decide in accordance with one's apparent reasons. (In Jojo's case the apparent reasons are genuine reasons, and they have entered his mind, but he does not 'see' their genuine normative implications and is not deciding and acting in accordance with them. In the BIV's case they are merely apparent reasons, and the BIV decides in accordance with these apparent reasons.) Jojo is unfree, as Wolf says, but he is so since he is lacking the rational ability to decide in accordance with his apparent reasons.

supplementation. My proposal is that the relevant reasons are *apparent reasons*, and the rational abilities are abilities to decide (form intentions) in accordance with apparent reasons. Whose apparent reasons? The subject's, of course. So we can say:

(RAAR) The relevant rational abilities that make us free with respect to decisions are abilities to make decisions in accordance with our apparent reasons.

What is '*accordance*'? This is the last thing we need to fix. Fortunately, this is not too hard to accomplish. It is quite clear that apparent reasons have things they 'favor', just as genuine reasons favor certain things. An apparent reason favors an attitude just in case it would favor that attitude if it were veridical (i.e., a genuine reason). So in Parfit's case of Delia, what Delia believes—that there is a dangerous fire in the hotel—favors jumping out of the window (in the sense in which apparent reasons can favor responses); it provides an apparent reason for Delia's jumping out of the window.¹³ One's decision then will be in accordance with one's apparent reasons if and only if it is favored by these apparent reasons.¹⁴

What is not answered by RAAR is the *metaphysical priority*. Do reasons explain rationality, or is it the other way round? RAAR is neutral on this grounding question. So both reasons-firsters and rationality-firsters can accept it. This is as it should be, it seems. The direction of priority concerns a deep metaphysical issue that we do not have to settle in order to say which rational ability one needs to have in order to decide freely.

Other things are also not yet specified. Are we talking about abilities that somehow include opportunity conditions? And what does '*making decisions* in accordance with one's apparent reasons' mean exactly? Does it mean that one can decide for or against?—We will address these issues in the next section.

By now, to sum up, we have seen that views that build on *genuine* reasons (such as, arguably, Wolf's or Nelkin's view, and possibly other rationalist views) cannot be on the right track since this would require too much for freedom. The more adequate view requires 'merely' the ability to decide in accordance with one's *apparent* reasons.¹⁵

¹³ It is a hypothetical favoring, if you like, but that is good enough for the present purposes. We are not engaged in figuring out whether apparent reasons have any genuine normativity here. It may be conceded that they do not. Still, they 'favor' or 'select' certain attitudes, and that is all we need.—In case appeal to apparent reasons as considerations that would be genuine reasons if true turned out to be problematic, one could substitute some alternative conception. (Cf. Wodak (2019), for example.) This paper's main proposal about how to understand the relevant rational abilities is not committed to any analysis of the deep nature of apparent reasons. (I am grateful to an anonymous referee to make me more aware of this issue.)

¹⁴ We have not yet fixed the strength of the apparent reasons. It seems plausible to take overall or all-things-considered apparent reasons to be the relevant type of reasons. Various views could be taken here, and we need to decide the issue conclusively for the present purposes.

¹⁵ At this point one might wonder if deciding in accordance with one's apparent reasons is sufficient for having the relevant kind of *control*, namely, the control that is a component of moral accountability, or whether some ability to recognize genuine reasons is required. The answer has to be that no such further ability is required *locally*. That is, in each instance of decision, the relevant control is furnished by having the ability to decide in accordance with one's apparent reasons and one need not be able to recognize the genuine reasons (as the biv's case illustrates). *Globally*, however, there might be some requirement since, plausibly, being a subject that has apparent reasons might require some ability to recognize genuine reasons. For the present purposes, I believe that we can refrain from trying to spell out this global connection.

3 Rational Abilities and Dispositions

It is strictly speaking not necessary to bring in dispositionalist about abilities at this point. But many philosophers are inclined to bring it in at this point, and it can be very constructive and helpful to see what happens if we do so. And in addition, it will bring out one important and non-trivial, surprising result. So I will try to spell out a dispositionalist view of abilities and apply it to the RAAR abilities in this section. The dispositionalist view that I will present is a kind of core view that can be found in the works of a number of philosophers, more or less. It is close to Vihvelin's view, for example. The main point of this section will be to see what happens and follows if we add this dispositionalist view to the rationalist leeway view presented above. If some philosophers prefer some minor modifications, this could be done quite easily, I believe, and it would neither change the main picture nor the interesting result about the ability to do otherwise that we will eventually arrive at.

Before we can look at dispositionalism, however, we need to *taxonomize abilities* in a more fine-grained way. So here comes a suitable and plausible taxonomy of abilities and, in particular, of rational abilities of the RAAR sort. This is again a kind of core account that several writers have developed or adopted.¹⁶

Let us then begin with abilities *in general*. Fortunately, the recent debate about abilities (in the context of issues of rationality and responsibility, at least) has converged nicely on a kind of characterization (or 'definition') of abilities that we can rely on for the present purposes. This is the *capacity-opportunity conception* of abilities, as we can call it. According to this consensus, abilities to M can be taxonomized in a onion-like fashion. First, there are internal (complex) properties (internal structures) that constitute one's having a capacity to M (frequently also called 'skill' or 'know-how'). This is, or grounds or constitutes, the *general ability* to M. (A specific neurophysiological structure, like a brain structure, might be the kind of thing that is a typical internal complex property for human beings that constitutes the subject's skill or know-how.) Second, one can in addition be in internal opportunity conditions for exercising this capacity to A. Then one has the so-called *narrow ability* to A. The internal opportunity conditions are comprised of something like the physical and psychological conditions that allow the subject to produce the internal bodily and mental activities that M-ing on the basis of the capacity to M involves (i.e., lack of anesthesia, drugs, damages to the nervous system, and extreme psychological conditions like those present during psychosis, or the like, that undermine the interaction of one's skill with the rest of one's body and one's non-factive mental states and processes). Third, if in addition external opportunity conditions obtain, one has the *wide ability* to A. The external opportunity conditions include having all the means and external conditions necessary for an unimpeded interaction with relevant objects of one's environment during the exercise of one's skill, such as, for example, having a tennis racket, a ball, and a court for exercising one's skill to play tennis. So, one gets from the general ability to the narrow and to the wide one by adding internal and external opportunity conditions.¹⁷ One's narrow abilities do not change by changes in the environment, whereas one's wide abilities may. One could then say that if (in context c) a subject has the narrow ability to M, 'she has what it takes' to M; and a subject has the wide ability to M if she 'has what it takes to M' and, in addition, she

¹⁶ See, for example, Clarke (2015), Franklin (2015), Vihvelin (2013), and Whittle (2010). It is important to note that this paper is not concerned with any semantic project concerning disposition *language*.

¹⁷ 'Interference-free' abilities as understood by Nelkin are wide abilities. Cf. Nelkin (2011), pp. 66-67.

‘has the means and opportunity to M’.¹⁸ (These characterizations are to be understood as helpful paraphrases merely and not as definitions.)¹⁹

Please note that for all these kinds of ability—the general, the narrow, and the wide abilities—what it is an ability to do is always the same: an ability *to M*, say. One can also generate different sorts of abilities by changing the ‘content’ of the ability (i.e., what replaces the ‘M’). For example, instead of the ability to ride a bicycle one can consider the ability to ride a bicycle when blindfolded, or the ability to ride a bicycle when wearing sunglasses. These are more fine-grained, ‘local’ abilities, since the ‘contents’ specifying the manifestations are more fine-grained.²⁰

Now, it is quite clear and important to keep in mind that this kind of differentiation is independent of the other kind of distinction between different types of ability (general, narrow, and wide). The commonly used terminology may cause confusion and is not entirely free of misleading connotations. But it is clear that the two ways of distinguishing kinds of ability are very different and have nothing to do with each other. In the very recent debate, the more fine-grained, local abilities have come to take center stage on the question of which ability to do otherwise subjects in Frankfurt-style cases have or lack. Ann Whittle, for example, has argued that the crucial and relevant ability in the controversy between compatibilists and incompatibilists is the ability [*to do otherwise when in a Frankfurt-style condition*], a *local* (wide) ability, not the global (wide) ability to do otherwise.²¹ (Here the brackets are supposed to make it easier to read the ability characterization as one that concerns, in its ‘content’, a fine-grained thing to be done, and not as one that specifies the context in which the subject is actually placed. Whittle uses “–“ hyphenation instead, and she is one of the few philosophers who are fully aware of this distinction and explicitly make it.)

Let us now turn to the special case of *rational abilities*. The basic task here is to fix what counts as the ‘M’, the ‘success’ or manifestation of the ability and disposition. Since what is at stake is no more and no less than rationality, the issue will be controversial, to be sure. But still, a suitable and plausible specification of the manifestation can be found, as I will argue now.

Relying on what has been said above, the manifestation is an accordance with one’s apparent reasons, as (RAAR) says. But we have to note that this is merely a schematic or generic specification. The manifestation is still significantly underdetermined. For it leaves open what the *first relatum* is supposed to be. The second relatum of the accordance relation is the subject’s apparent reasons. But what is the first relatum? One might think it is easy to specify the first relatum: it is simply a *decision to do x* (to raise one’s arm, to buy milk in the grocery store around the corner, etc.). This is one option, of course. But another option is to say that the first relatum is ‘two-way’, so to speak: it is a decision *to do or not to do x*. This would mean that the abilities become *two-way abilities*, similar to the so-called two-

¹⁸ Cf. Vihvelin (2013), p. 11.

¹⁹ In addition to abilities (accounted for by dispositions), one could also postulate *teleofunctional* entities, i.e., entities that have the teleofunction to bring about certain effects. Deviating from the terminology used in this paper, one could call these teleofunctional entities ‘capacities’, in order to distinguish them from mere abilities that do not have any teleofunction or purpose. For the present purposes, however, we do not need to enter the discussion of capacities teleofunctionally understood.—For an interesting discussion of capacities, abilities, and dispositions see Millikan (2000). Space does not allow me to adequately deal with Millikan’s views here.

²⁰ Calling these abilities ‘local’ is taken from Whittle (2010).

²¹ Cf. Whittle (2010).

way powers (to act or not to act in a certain way) that have figured prominently in the recent literature and that can be found in Aristotle's work.²² Please note, however, that here we are talking not about action but about *decisions* primarily. So things can be a bit different, though the same in structural respects.

Maybe it is helpful at this point to say a bit more about *decisions*. I take a decision to be the formation of an intention. Following Hieronymi, we can say that intentions settle practical questions. The question that one is often faced with is whether to do or not to do *x*. Famously, Aristotle held that the abilities we have in virtue of having *nous* are abilities to act or not to act in a certain way, i.e., abilities with 'disjunctive manifestations'. (To call them 'disjunctive' is merely meant to provide a nice and handy label, nothing more.) The second option sees the relevant rational abilities to decide as being of a similar two-way sort: they are abilities to decide whether to do or not to do *x*. There are then two quite distinct possible manifestations, deciding to do *x* and deciding not to do *x*.²³ So the second option is the two-way conception of rational abilities like those referred to in RAAR.

Since no further plausible candidate comes to mind, we have basically two options. Which one is the right one? We can significantly advance our account of free will by answering this question about the right decision parameter. The way I will approach the question will be the following one. I will now bring in dispositionalism about abilities and spell it out for the rational abilities that we have identified so far. Then I will ask if the resulting account *will yield the leeway claim*, i.e., the claim that rational abilities entail abilities to do otherwise. We will then see that only the disjunctive conception of the decision parameter can fit the bill. So the alleged, and desirable, implication of abilities to do otherwise by rational abilities can really be vindicated—by saying that the relevant abilities are abilities to settle two-way questions of the form 'whether to do or not to do *x*'. This is the plan for the remainder of this paper.

4 Dispositionalism About Abilities

Let us begin with dispositionalism about abilities in general. The application to the case of rational abilities will then not be too hard. And it is of course desirable to provide a dispositionalist account of rational abilities that fits into a general dispositionalist view of abilities, in order to avoid *ad hoc* assumptions.

²² See, for example, Alvarez (2013), Frost (2013), Steward (2012), and Pink (2009). Alvarez also presents the standard Aristotelian argument why the two-way power is not just a conjunction of two one-way powers (cf. Alvarez 2013, p. 109).

²³ Why not take the following two ones: deciding to do *x* and not deciding to do *x*? Well, the argument is flat-footed. Not deciding is not a decision, and so it cannot be a manifestation of an ability to decide ... (whatever we fill in after the verb "decide"). This is different for actions. Sometimes the ability to act or not to act in a certain way—with acting or not-acting as the two manifestations—may be the relevant ability. If not acting (omission) does not count as an action in the relevant sense (in some context), we should probably not take this ability as the relevant one. For decisional abilities, however, it seems clear that the manifestation is deciding to do *x* or deciding to not do *x*. Well, we can also remain undecided; decisional suspension is a further possibility. But it does not count as a further type of making a decision. 'Not deciding' should be understood as neither deciding for nor deciding against. As long as one leaves things undecided, one has not decided. What matters for freedom is being able to take some real decision, for doing something or against it, when apparent reasons favor doing it or favor not doing it, respectively. Sometimes suspension is permitted, but it does not constitute a *settling* of the question whether to do or not to do *x* (just like doxastic suspension does not constitute a settling the question whether *p*).

Many philosophers favor a general dispositionalist account of abilities that starts with general abilities which are identified with an intrinsic disposition or set of intrinsic dispositions.²⁴ It is intrinsic since the ability is grounded in an internal (complex) property of the subject, the ‘skill’ or ‘know-how’, as it is sometimes called. (That we should consider intrinsic rather than extrinsic dispositions is clear from our biv intuition.) To have the general ability to M, then, is to have a set of intrinsic dispositions to M. If we add the internal opportunity conditions, we get the narrow ability. So we can say:

(NA) To have the narrow ability to M is to have a set of intrinsic dispositions to M and to be such that the internal opportunity conditions (for this ability) obtain.

In the following, I will always take the internal opportunity conditions to be obtaining and drop mentioning them, to make the formulations simpler.

If we add the external opportunity conditions, we get the wide ability:

(WA) To have the wide ability to M, then, is to have a set of intrinsic dispositions to M and to be such that the internal and external opportunity conditions (for this ability) obtain.²⁵

From now on, unless noted otherwise, we will talk about narrow or wide abilities, not general abilities. But for both narrow and wide rational abilities, we can simultaneously discuss the question whether they entail abilities *to do otherwise*.

As a next step, let us make use of a plausible view of dispositions that has many good arguments in its favor. It analyzes dispositions not in terms of counterfactuals (which is problematic because of masks and finks, as has been realized since long) but in terms of possible worlds.²⁶

(D) For S to have the intrinsic disposition to M is for S to be such that there is a wide range of worlds in which S Ms.

Just as a window pane’s fragility consists in there being a wide range of possibilities in which it breaks, so in general, the possession of an intrinsic disposition requires an analogously suitable modal profile. This account can be called a ‘*modal account of dispositions*’. And it can deal with many questions very successfully. For example, masks and finks are no problem since if an intrinsic disposition is masked, there will still be a wide range of

²⁴ Cf., for example, Vihvelin (2013), p. 11.

²⁵ It is to be noted that it is not absolutely clear what the ‘obtaining of the external opportunity conditions’ mean. The stronger reading is that they actually obtain, the weaker reading would be that they could (very) easily obtain. For example, if a tennis court and racket is just around the corner, one might say that the external opportunity conditions for the ability to play tennis obtain. I will assume the stronger reading here, but this does not lead to any problematic restrictions. I will also assume the stronger reading for the internal opportunity conditions.

²⁶ Please note that on this point, I depart from Vihvelin’s account of dispositions which is a combination of the modal account and Lewis’s refined Conditional Account (LCA). Therefore, Vihvelin’s account remains a sort of conditional account and, consequently, is hostage to the Lehrer/Chisholm objection of cases in which the subject is unable to try to do x, due to some psychological blockage—an objection that is very clearly presented in Clarke (2009, 328–29). I propose to get rid of conditionality entirely, and I do not find Vihvelin’s reasons for sticking to some sort of conditional analysis convincing. (Cf. Vihvelin (2013), p. 185.)—Thanks to an anonymous referee for pointing me to this objection and urging me to take an explicit stand on this.

worlds in which the mask is absent, and this will suffice for the object's having the intrinsic disposition.²⁷ And if the mask 'becomes intrinsic', becomes a part of the object, it will lose the intrinsic disposition, since will be no longer any wide range of worlds in which the object Ms.²⁸

At this point we are finally in a position to continue the line of thought presented at the end of the last section. What happens if we use this general account of dispositions and apply it to the relevant *rational* abilities of the RAAR kind? To begin with, we can spell out the account of *one-way* rational abilities as follows:

(1WRA) S has the one-way rational ability to decide to do x in accordance with her apparent reasons iff there is a wide range of standard worlds in which S decides to do x in accordance with her (sufficient) apparent reasons.²⁹

Deciding to do x will be in accordance with one's reason just in case one has (sufficient) apparent reasons for doing x.³⁰ (I will assume that reasons for decisions are content-related reasons, i.e., reasons for the action that one decides for, and not state-related reasons. In our context, a very plausible and harmless assumption, I believe.)³¹ So we get.

(1WRA*) S has the one-way rational ability to decide to do x in accordance with her apparent reasons iff there is a wide range of standard worlds in which S decides to do x and has (sufficient) apparent reasons for doing x.³²

Analogously, we can analyze the two-ways rational abilities:

²⁷ Cf., e.g., Manley, Wasserman (2008).

²⁸ Further specifications are needed in order to avoid counterintuitive results. A kind of standard world with 'standard' or 'normal' conditions is needed. (This is, in effect, what Manley and Wasserman do by talking about 'test cases C' in their Manley, Wasserman 2008.) We need to look at what happens at standard worlds. So, in effect, having a disposition to M amounts to there being a wide range of standard worlds in which the object Ms. (Sometimes the standard is specified to some extent by mentioning stimulus conditions, but not always, I take it.) Only standard worlds are relevant since otherwise there might be M-worlds somewhere far out there in modal space that, intuitively, do not matter. And an object can be fragile, e.g., even if it actually is in conditions that are very much atypical or abnormal.—In what follows I will mention this further qualification in brackets.

²⁹ It is very natural to take rational decisions to be decisions that are in accordance with *sufficient* (apparent) reasons. In what follows, I will sometimes leave out this qualification and sometimes keep it in parentheses as a reminder.

³⁰ Why choose deciding *in accordance with* apparent reasons instead of deciding *for* apparent reasons?—There are some views of what it is to decide (or act) for a reason which would only allow for *genuine* reasons to play that role. (I have in mind views like Maria Alvarez's which hold that all reasons are facts. Cf. Alvarez 2010.) Deciding for a reason would make that decision (pro tanto) *justified*, and arguably that requires a *genuine* reason. But one can also tailor an account that allows for some kind of deciding for apparent reasons (by appeal to some motivational connection between an apparent reason and the resulting intention to act in a certain way). (Here I have in mind views like that presented in Mantel (2018), suitably adapted to decision instead of action.) Then, however, one could simply add this further condition to the proposed account if one wanted to. Roughly, the relevant possible worlds would then be worlds in which the subject decides to do x and has (sufficient) apparent reasons for doing x and the additional motivational connection between the apparent reason and the decision obtains. The rest would stay the same.

³¹ Piller (2001) and Parfit (2001) were the first to make this distinction.

³² One could of course add a causal relation between the possession of the apparent reasons and the decision. For simplicity's sake I will refrain from such a modification.

(2WRA) S has the two-ways rational ability to decide to do or not to do x in accordance with her apparent reasons iff there is a wide range of standard worlds in which S decides to do or not to do x in accordance with her (sufficient) apparent reasons.

What does this amount to? A natural and plausible answer is this:

(2WRA*) S has the two-way rational ability to decide to do x or not do x in accordance with her apparent reasons iff in the vast majority of standard worlds it is the case that S decides to do x iff S has (sufficient) apparent reasons for doing x and S decides not to do x iff S has (sufficient) apparent reasons not to do x.³³

Deciding to do x or not to do x can also be described as deciding *whether to do x or not to do x*. This brings a natural understanding of decision to the foreground, namely, that decisions can go one way or the other way. Deciding is *settling* the question whether to do or not to do x. The manifestation of an ability to decide *whether* to do x or not is, thus, either a decision to do x or a decision to not do x. Either counts as having decided whether to do x or not, and nothing else does. We can therefore rephrase our answer (2WRA*) as follows:

(2WRA)** S has the two-way rational ability to decide whether to do x or not in accordance with her apparent reasons iff in the vast majority of standard worlds it is the case that S decides to do x iff S has (sufficient) apparent reasons for doing x and S decides not to do x iff S has (sufficient) apparent reasons not to do x.

The crucial point is to ‘split off’ the required set of worlds into two sets, the ‘plus worlds’ where the question (whether to do or not to do x) is answered affirmatively and the ‘minus worlds’ in which it is answered negatively. Why should we do this? Well, it is a natural and intuitive move to make here, it seems to me. I lack any proof or decisive argument, apart from my intuition. The following argument in its favor, however, can be given. To leave out the plus worlds would mean that in a case in which one decides not to do x one could do this freely even without there being a wide range of worlds in which one has reason to do it and does it. This, however, could be satisfied by one’s being somewhat forced or compelled to decide not to do x. Analogously, to leave out the minus worlds would mean that in a case in which one decides to do x one could do this freely without there being a wide range of worlds in which one lacks reasons to do x and decides not to do it. Either idea is in conflict with our basic intuitions about free decisions, I submit. The point, of course, is that free decision making requires an ability to do otherwise. And if the ability that is at stake is the ‘disjunctive’, two-way rational ability to decide to do or not to do x, then we can preserve this ability-to-do-otherwise intuition by, and only by, opting for the just-given analysis, i.e., (2WRA*) or (2WRA**).

³³ This definition interprets the ‘wide range of worlds’ in a more demanding way, namely, as concerning the vast majority of standard worlds. This is necessary in order to avoid the possibility that, though there is a wide range of standard worlds in which decisions match apparent reasons, there is an even bigger number of standard worlds in which there is no such match. Second, the possibility of (almost) only having (e.g.) standard worlds with sufficient reasons for deciding to do x (and none with sufficient reasons for deciding not to do x) poses no problem since by just using one quantifier over possible worlds and biconditionals in its scope it remains open how many of the standard worlds are worlds with sufficient reasons for deciding to do x and how many are worlds with sufficient reasons for deciding not to do x. This is intuitively adequate.

Thus, I submit, two-way rational abilities as characterized in (2WRA*) or (2WRA**) are the right choice. This characterization will have the desirable consequence that the relevant rational abilities entail an ability to do otherwise and, therefore, can play the role of grounding our free will. The dispositionalist analysis of abilities can thus help us answer the question that remained to be settled at the end of Sect. 3, viz., whether the one-way or the two-way conception of abilities to decide is the right one. We should opt for two-way powers to decide whether to do or not do something, and then our RAAR abilities will entail the ability to do otherwise. In this way, a rationalist leeway account of free decision can be upheld and stands out as the first choice to take. At least, anyone who wants to endorse some other conception of rational abilities now has the burden of proof to show that this other conception is at least as good as the proposed rationalist leeway account. Not an easy task, I submit.³⁴

It remains to note that two things have not been settled. First, the Nelkin-Wolf asymmetry issue has not been addressed. But it is quite clear that one can implement it in the proposed account. All one needs to do is to add the claim that only in cases where one decides to do the wrong thing is it necessary to have the ability to decide to do or not to do it in accordance with one's apparent reasons, but not in cases in which one does the right thing. This issue could be tackled by looking at further arguments, and the investigation could remain within the proposed account of rational abilities. Second, we have not addressed the issue of compatibilism versus incompatibilism. But this problem could also be approached with the help of the proposed account of rational abilities. My hunch is, to venture at least this much, that compatibilism is not at all out of question.³⁵ But again, this would have to be argued on the basis of further considerations that go beyond the scope of this paper.

5 Conclusions

To think that rational abilities bestow on us the crucial ability to decide otherwise is an attractive view of free will that has been the main topic of this paper. I have first investigated what the rational abilities that makes our decisions free are abilities to do. (RAAR) provides the answer: they are rational abilities to decide in accordance with our apparent reasons. Then a second interesting result concerning recent 'new dispositionalists' accounts of abilities has been argued for.

If we think that abilities can be accounted for along these lines, then an attractive option is available: the ability that makes us free is a *two-way ability* to decide to do x or not to do x. For if the relevant rational (narrow) ability to decide to do x or not to do x in accordance with one's apparent reasons is understood in this way, i.e., as the corresponding two-way (intrinsic) disposition, it entails the ability to decide otherwise. A one-way ability, understood dispositionally, will not guarantee such an ADO.

³⁴ One could try to tailor an account that appeals to *one-way* rational abilities and a *further* theoretical ingredient that makes the view imply an ability to do otherwise. I take it that such an account would be inferior in that it exhibits less theoretical unity or elegance. In addition, it is not clear how to do this, since the additional ingredient would have to be coherent with the one-way rational abilities. Unless an intimate connection between the additional ingredient and rational abilities could establish such harmony, such an add-on account would describe the phenomenon as quite dis-unified.

³⁵ The topic of whether two-way dispositions are compatible with determinism is lucidly discussed in Frost (2013), for example. I believe that Frost is right in his interpretation of Aristotle as a compatibilist.

Acknowledgments I am grateful to many helpful discussions to Christian Loew, Yannick Kohl, Susanne Mantel and Christian Piller.

Funding This work was supported by the Fond National de la Recherche Luxembourg (FNR) through the RREXAM project (OPEN grant O18/12880605). I am very grateful for this support by the FNR.

Declarations

The author has no competing interests to declare that are relevant to the content of this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alvarez M (2013) Agency and two-way powers. *Proceedings of the Aristotelian Society* 113, part 1, 101–121
- Alvarez M (2016) Reasons for action, acting for reasons, and rationality. *Synthese* 195:8
- Alvarez M (2010) *Kinds Of Reasons*. OUP
- Clarke R (2015) Abilities to act. *Philos Compass* 10/12:893–904
- Clarke R (2009) Dispositions, abilities to act, and free will: the new dispositionalism. *Mind* 118:323–351
- Franklin C (2015) Everybody thinks that an ability to do otherwise is necessary for free will and moral responsibility. *Philos Stud* 172:2091–2107
- Frost K (2013) Action as the exercise of a two-way power. *Inquiry* 56(6):611–624
- Hieronymi P (2013) “The use of reasons in thought (and the use of earmarks in arguments)”. *Ethics* 124, 114–27
- Hieronymi P (2009) “Two kinds of agency”. In: O’Brien L (ed) *Mental Action*. OUP, pp 138–162
- Kappel C (2019) “Escaping the acratric trilemma”. In: Skipper, Steglich-Petersen (eds), *Higher-order Evidence*. *New Essays*, OUP, 2019, 124–43
- Littlejohn C (2019) Being more realistic about reasons. *Philos Phenomenol Res* 99:3:605–627
- Manley D, Wasserman R (2008) On linking dispositions and conditionals. *Mind* 117:465
- Mantel S (2018) *Determined By Reasons*. Routledge
- Millikan RG (2000) *On Clear And Confused Ideas*. CUP
- Nelkin D (2011) *Making Sense of Freedom and Responsibility*. OUP
- Parfit D (2011) *On What Matters*. OUP
- Parfit D (2001) “Reasons and Rationality”. In: Egonsson D, Petersson B, Josefsson J, Rønnow-Rasmussen T (eds) *Exploring Practical Philosophy: From Action to Values*, Aldershot, Ashgate, P 17–41
- Parfit D (1997) *Reasons and Motivation*. *Proceedings of the Aristotelian Society Supplementary Volume* 71: 99–130
- Piller C (2001) Normative practical reasoning. *Proceedings of the Aristotelian Society Suppl Vol* 75:1: 175–216
- Pink T (2009) “Power and moral responsibility”. *Philosophical Explorations* 12: 2, June 2009, 127–149
- Skipper M, Steglich-Petersen A (eds) *Higher-order Evidence*. *New Essays*. OUP, 2019
- Smith M (2003) “Rational capacities”. In: Stroud S, Tappolet C (eds) *Weakness of Will and Practical Irrationality*. OUP, pp 17–38
- Smith M (1994) *The Moral Problem*. Blackwell
- Steward H (2012) *A Metaphysic For Freedom*. OUP
- Strawson P (1962) “Freedom and resentment”, *Proceedings of the British Academy* 48, 187–211
- Vihvelin K (2013) *Causes, Laws, And Free Will: Why Determinism Doesn’t Matter*. OUP
- Weatherson B (2008) Deontology and Descartes’ Demon. *J Philos* 105:9:540–569
- Whittle A (2010) Dispositional abilities. *Philosophers’ Impr* 10:12

- Widerker D (1995) Libertarianism and Frankfurt's attack on the principle of alternative possibilities. *Philos Rev* 104:2:247–261
- Williamson T (forthcoming), "Justifications, excuses, and sceptical scenarios". In: F. Dorsch, J. Dutant (eds) *The New Evil Demon*, OUP
- Williamson T (2017) "Ambiguous rationality". *Episteme* 14:3, 263-74
- Wodak D (2019) An objectivist guide to subjective reasons. *Res Philosophica* 95:2:229–244
- Wolf S (1987) "Sanity and the metaphysics of responsibility", reprinted in: R. Shafer- Landau (ed). *Ethical Theory: An Anthology*, Blackwell, Oxford, 2007, 363-72
- Wolf S (1990) *Freedom Within Reason*. OUP

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.