

“Rational belief, reflection, and undercutting defeat”

Frank HOFMANN

University of Luxembourg

Author’s preprint

Abstract:

Philosophers disagree about the role of reflection for rationality, understood as the capacity to (properly) respond to genuine, normative reasons. Here, ‘reflection’ means the capacity for self-conscious normative meta-cognition. This paper develops and rejects a novel argument – the *argument from undercutting defeaters* – in favor of the ‘one-level view’ which holds that having the concept of a belief (and of a reason) is necessary for responding to reasons. It will be argued that the ‘two-level view’, which allows for rational subjects that can only non-reflectively respond to reasons, is supported by considerations to do with the role of responding to reasons for rational action. Rationality is not as unified as the one-level view wants to have it. We start with the non-reflective way of rationally forming beliefs and then grow into the reflective way.

Keywords: rationality, reasons, reflection, undercutting defeaters.

1 Introduction

Arguably, rational agency requires *belief*. Without beliefs one cannot rationally decide whether to act in a certain way or not, since one lacks a perspective on how things are. But does rational agency also require having the *concept of belief*? This is much less clear. Of course, having the concept of belief is a crucial prerequisite of the ability to think about one’s own beliefs in a way that allows for evaluative assessments of these beliefs and, thus, to ‘*reflect*’ on them (or to engage in what is called ‘critical reasoning’ in Burge 1996). But it remains a non-trivial and controversial issue whether rationality requires such an ability to reflect on one’s beliefs and, thus, the possession of the concept of belief. The cases of small children and some non-human animals seem to suggest that it is not required.

The main goal of this paper is to discuss some arguments concerning this issue. In particular, I will formulate and reject a novel (though quite natural) line of thought in favor of the view that rationality requires the concept of belief.

2 The issues and views at stake

In the following, by '*reflection*' I will mean self-conscious normative meta-cognition, roughly speaking. One reflects on one's attitudes, in this sense of 'reflection', by having first-person meta-beliefs about some normative feature of one's own attitudes. For example, if Samira believes that her first-order belief that *p* is not justified, she is in a state which is a paradigm case of reflection. (Please note that reflection in this sense is not to be identified with *introspection*. Reflection can also be based on *testimony*, for example; it need not be based on introspection.)

By '*rationality*' the following core phenomenon is meant. Subjects follow normative reasons, they base their attitudes on such reasons, they hold attitudes for these reasons. At heart, this is the phenomenon that will be understood as '*rationality*' here; it is *rationality as responding to genuine normative reasons*. (Other notions of rationality, such as so-called '*structural rationality*', i.e., rationality as coherence, will not be the target here. I will also exclusively focus on genuine normative reasons and leave out merely apparent reasons.).¹

Now, we can state an important question: *does all rationality require the presence of reflection?* Or is it reserved for a special kind of rationality, i.e., '*reflective rationality*'? The paper wants to contribute to the development of a '*two-level view*' of rationality: some exercises of rationality are free of any reflection while others involve reflection; reflection is not essential to rationality as such. A *one-level view*, in contrast, would hold that *all* exercises of rationality involve reflection and that reflection is

¹ The '*responding*' here is always meant to be a *proper* responding. As familiar from the literature, there might be cases of improper basing, i.e., responding to reasons that proceeds along irrational ways (such as irrational rules of inference). Cf. Turri (2010). Here I will take the '*responding*' to always be a *proper* responding, to keep the formulations simple. In addition, one can include *actions* as responses to normative reasons. (Cf. Dancy 2018, for example.) To simplify formulations, I will not mention this option in the following.

essential to rationality. The one-level view could also be called '*reflexionism*'. And the two-level view constitutes a kind of '*anti-reflexionism*'.

To make the issues simpler and more manageable, this paper will focus on the attitude of *belief* and reflection on one's own *beliefs*. Other attitudes will be left out. But surely, if a certain account of reflection about belief is established, there will be a good chance to generalize to other attitudes as well. To make things manageable, this paper will make a start by exclusively looking at belief.

In addition, the paper will focus on rationality as an *ex post* status and leave out *ex ante* status. (*Ex post* status is a status of an attitude that the agents holds; *ex ante* status, in contrast, is a status of an attitude that the agent need not hold.) Acquiring defeating evidence can change the *ex ante* status of a belief, of course. But the reflexionist's view is probably rather that rationality requires being able to properly *makes use of* defeating evidence, and so the natural focus in this debate is on the *ex post* status of a belief that results from treating some consideration as an undercutting reason.

Related to the just mentioned main question is the question of *whether rational believers have to possess the concept of belief*. Evidently, if they lack the concept of belief, they cannot enter into any reflective states, since meta-cognition requires beliefs about one's own beliefs. One cannot represent one's beliefs in meta-cognition without having the concept of belief, since any meta-cognitive thought, as understood here, is about belief (in the *de dicto* sense, of course). So, if we can show that some rational believers lack the concept of belief but are still in possession of some rational capacity (ability), the one-level view is refuted. On the other hand, if it can be shown that any rational believer must have the concept of belief, the one-level view has at least a chance of being true. What I would like to argue here is that a certain argument for this central implication of the reflexionist's view – that any rational believer must have the concept of belief – fails. This weakens the reflexionist position insofar as some other support would have to be found.

Why should anyone hold that reflection is essential to rationality? – Initially, it seems rather plausible to think that one can respond to reasons both in a reflective and a non-reflective way. Sometimes one responds to a reason, *x*, for believing that things are so and so while believing that *x* is a reason for one to believe so; and sometimes one responds to reasons without such a reflective belief. Why not? After all, there is a quite strong intuition that *small children* and *some non-human animals* can respond to

reasons, but most probably lack reflection (at the relevant time, at least).² *Prima facie*, this provides a strong case against reflexionism. Reflexionists will have to re-interpret what is going on in small children and these non-human animals, and that can be seen as constituting some significant cost. *Prima facie*, it is almost an obvious fact that small children can have *some* rational capacity or ability for belief. (Here and elsewhere, I will use the two terms ‘capacity’ and ‘ability’ synonymously.)

Proponents of the one-level view – reflexionists – think that believing for a reason requires believing for a reason *as a reason* for so believing. There is no such thing as believing for a reason without believing for a reason *as a reason* (or *as such*). John McDowell is a prominent proponent of this idea.³ In his words,

“The notion of rationality I mean to invoke here is the notion exploited in a traditional line of thought to make a special place in the animal kingdom for rational animals. It is a notion of responsiveness to reasons *as such*. (...) For that idea [of responding to a reason as such, FH] to be appropriate in this connection, we would need to be considering a subject who can step back from an inclination to flee, elicited from her by an apparent danger, and raise the question whether she *should* be so inclined – whether the apparent danger is, here and now, a sufficient reason for fleeing.” (McDowell 2006, 128; McDowell’s own emphases)

In other words, exploiting reasons for belief requires an *ability to think about their being reasons for belief*; and such an ability is an ability to have reflective meta-beliefs to the effect that some facts are reasons for one’s belief. In effect, it might even require meta-*knowledge* that they are reasons for belief (and this could perhaps further difficulties for the view). Without believing (or even knowing) that the fact that x is a reason for the belief that p, one just cannot count as believing p in response to the reason ‘as such’. (McDowell is inclined to weaken this requirement by turning it into

² It is well known that small children develop a capacity for belief attribution (as part of a ‘theory of mind’) only around age 3. Standardly, performance in the false-belief task, developed by Wimmer and Perner (1983), is taken to support this claim.

³ Apart from McDowell and Matthew Boyle, there is surely a number of other philosophers who subscribe to the one-level view. I am fairly confident to include Andrea Kern (2017) and Sebastian Rödl (2007), for example. (The conditions for having the concept of a belief or a reason for belief could of course be watered down quite a bit. This further issue of what constitutes concept possession need not be settled here.)

the corresponding *modal* condition: one must be able to believe (or know) that the fact is a reason. But even if this weaker condition is endorsed, having the capacity for reflection is still necessary.) One might be ‘tracking’ reasons, tracking the facts which are the reasons, or something like that. But this is not the real thing; it is not the same as believing for a reason ‘as such’.⁴ Or so thinks John McDowell.

Similarly, Matthew Boyle describes rationality as essentially reflective:

“We could therefore say ... that the concept *rational animal* is a ‘concept of reflection’ in Kant’s sense: it is a concept whose source lies in our reflective consideration of our own cognitive activity, rather than in our empirical observation of particular objects with which our cognition is concerned, cf. Kant (1998), A260/B316.” (Boyle 2012, fn. 38)

The essential connection between rationality and reflection is clearly stated. In addition, Boyle thinks that the essential involvement of reflection helps to see how we can know about our own rational activities. Though this further claim is very interesting, and probably part of the motivation for the view, here I will have to bracket any discussion of it, simply for reasons of space. But as one can probably glimpse, very interesting further claims open up as extensions or further elements of reflexionism as it will be examined in this paper.

If one looks at the contemporary scene of accounts of believing for a reason, however, one does find prominent and well elaborated accounts that do *not* require reflection, at least *prima facie*. An example of such an account is the one that Kurt Sylvan and Errol Lord have proposed recently.⁵ According to this account, roughly, one has to be *sensitive* to x’s being a (normative) reason for believing p in order to believe p for the reason that x. Such a sensitivity is a modal feature that can be understood as constituted by a *dispositional feature* (or *competence*), as follows:

⁴ Cf., e.g., Karen Jones’s distinction between ‘reason-responders’ and ‘reason-trackers’ in Jones (2003), 189-190.

⁵ Cf. Lord, Sylvan (2019). Another (similar) account that does not impose any reflective dimension on responding to reasons is Susanne Mantel’s account (Mantel 2018, esp. ch. 4). This is an account of *acting* for a reason, but *mutatis mutandis* it could probably be transferred to the doxastic domain.

“BSNR-Competence: What it is for S to believe q for a sufficient normative reason p that S possesses to believe q is for S’s believing that q to be sustained by her competence to revise her belief if p ceases to be a sufficient normative reason S possesses to believe q.” (Lord, Sylvan 2020, 163)

There is no reference to any belief about p’s being a reason for believing q in this analysis, and so this account of responding to reasons does not require any reflection. Competences are not constituted by *propositional* knowledge, arguably. And it is not easy to see why the relevant competence would require the agent to (be able to) believe or know that the fact that p is a (sufficient) normative reason for S to believe q in order to have the relevant competence and sensitivity – which is rather a matter of ‘know-how’ than ‘know-that’.

So, it seems that we are facing a *standoff*. A further argument is needed in order to make a case for the reflexionist one-level view. What argument could there be? To just claim that an account such as Lord and Sylvan’s merely analyzes ‘*reasons tracking*’ but not ‘*responding to reasons as reasons*’ would not be enough. More must be said.⁶

In this paper, I will propose and discuss such a further argument in support of the reflexionist one-level view. It has to do with the phenomenon of *undercutting defeat* that has recently witnessed an intensive wave of discussion. Drawing on this debate, I will work out a novel argument that purports to support the reflexionist one-level view. A critical evaluation, however, will show that the argument does not succeed.

⁶ Let me mention two further considerations. One of McDowell’s reasons for imposing the capacity of reflection on responding to reasons (‘as such’) is surely his *access internalism*, inherited from Wilfrid Sellars. Being an externalist, I am not willing to accept this as a good argument. A second consideration is related to his view that responding to reasons constitutes *reasoning*, and reasoning is something an agent is doing when *responsibly self-determining* her own beliefs. (Cf. McDowell 2006, 138–139.) – However, it is far from clear why non-reflective reasoning by means of suitable competences that do not require reflection should not count as a kind of doxastic self-determination. This may lead to the subsequent worry that such a form of reasoning would not allow for *responsibility*. To deal with this worry would require further investigations that go beyond the scope of this paper. Let me just say the following. If responsibility is constituted by the ability to do otherwise, one can have this ‘leeway’ ability by having suitable rational capacities that embody sensitivity or safety with respect to some reasons. But then, we can again appeal to the idea that these rational capacities need not be *reflective* rational capacities. (See Hofmann (2022) for some further development and defense of this idea.) This would not necessarily lead into incompatibilism, as recent developments of so-called ‘New Dispositionalism’ have argued, I take it. Of course, all of this deserves further investigation.

Therefore, we can at least conclude that the one-level view cannot get any mileage out of the phenomenon of undercutting defeat.⁷ Prima facie, this weakens the reflexionist view quite a bit. For the phenomenon of undercutting defeat is most congenial or amiable to this view, since it already is of a higher-order nature. An undercutting defeater is evidence for some higher-order proposition about a normative feature of one's belief and belief formation, and so it is in some straightforward sense a higher-order phenomenon. (More precise characterizations will be given in the next section.) If not even this phenomenon yields a good case for reflexionism, reflexionists will have to look for other phenomena that are not, or at least not ostensibly, of a higher-order nature and rather invite the thought that they do not require reflection. Simultaneously, this adds credibility to the two-level view. Responding to reasons is not essentially a reflexive phenomenon. Rationality is not unified by reflection. We can endorse the very plausible view that small children (and some non-human animals) can have *some* rationality without having the special, *reflective* form of rationality that is involved in thinking that one's own first-order belief is not properly based on reasons.

3 Terminology, assumptions, and a taxonomy of defeaters

Let us move on by further fixing our concepts and stating some assumptions.

As already indicated, in the following, a *reason* will always be a *genuine normative* reason (not a mere motivating reason).⁸ The term '*rationality*', as already mentioned, will stand for responding to normative reasons. And throughout this paper, the relevant responding will always be understood as *proper* responding (cf. footnote 1 above).

The entire discussion will be held in the framework of a *non-psychologistic factualism* about reasons for belief. So, reasons for belief (epistemic reasons) are (in

⁷ In principle, there might be some other argument revolving around undercutting defeaters, but I am unable to see what it could be like.

⁸ It has become standard to distinguish between motivating and normative (or justifying) reasons. For a recent presentation of this distinction as applied to the doxastic domain see, e.g., Littlejohn (2019).

general, at least) non-psychological facts or states of affairs.⁹ This is mostly for reasons of having a unified and convenient terminology. The main points, I believe, could be translated into other frameworks. In any case, the non-psychologistic factualist view has become popular and is currently one of the main contenders in the field.¹⁰ A further assumption in the following will be that reasons for belief are provided by pieces of evidence. This excludes pragmatic reasons for belief, and it constitutes a commitment to some form of evidentialism. Again, however, this commitment need not be essential to the main points of the line of thought below. (And I suspect that proponents of reflexionism will not bark at this point.)

Now, we can start looking at the phenomenon of *defeaters*. As has become standard, we can distinguish between *rebutting defeaters* and *undercutting defeaters*.¹¹ A *rebutting* defeater (RD) for the belief that p is a piece of evidence that speaks against the proposition p. In contrast to rebutting defeaters, *undercutting* defeaters (UDs) are more variegated. They know of three different species: reasons-undercutting, basing-undercutting, and possession-undercutting defeaters. In addition, there is a further notion of a *generic* undercutting defeater. These are to be characterized as follows.

Firstly, a reasons-undercutting defeater is a bit of evidence to the effect that the reason on which one's belief is based *is not a reason at all* (or at least *not a sufficient* reason) for that belief. The undercutting goes against the status of being a (sufficient) reason for what the subject treats it as a reason.¹²

Secondly, a basing-undercutting defeater is a bit of evidence to the effect that there is *no proper basing* between what is treated as a reason and the belief for which it

⁹ Let me point out that in this paper, the question whether possession of a reason requires conceptual content can be left open and is not under discussion. – For more on the relation between evidence and normative reasons for belief see Hofmann (2021).

¹⁰ The view is of course similar to the view in practical philosophy that reasons for *action* are facts (in general, nonpsychological facts), made popular by T. Scanlon, D. Parfit, J. Dancy, and others. More recently, the same kind of view for reasons for *belief* has been proposed and investigated by a number of epistemologists. See, e.g., the collection of essays in Mitova (2018).

¹¹ The distinction is standardly credited to John Pollock. See, e.g., Pollock (2001), (1987).

¹² Here and in the remainder, the case where one has not yet formed a belief but is in the business of making up one's mind and a reason or candidate reason has come into view is meant to be included. Undercutting reasons, thus, can already be at work during the process of forming or not forming a belief (i.e., of settling the question whether p) and not only after the fact of having formed a belief. To simplify formulations, I will only mention the 'post facto' case in the following.

is treated as a reason. Basing-undercutting defeater target the feature of being properly based.¹³

Thirdly, a possession-undercutting defeater is a piece of evidence to the effect that *one does not possess* what one treats (or tries to treat) as a reason for one's belief. If one does not possess the fact that is a reason – even if it is a reason for one's belief – one cannot properly base one's belief on it, since possessing the relevant reason is a necessary condition of responding to a reason. So possession-undercutting defeaters target the possession of reasons.¹⁴

What unifies undercutters? The most general undercutting defeater is a bit of evidence to the effect that one's belief is not properly based on reasons that one possesses. This leaves it open which aspect is concerned, being a (sufficient) reasons, being properly based, or being in possession of the reason. We can call these unspecific undercutters '*generic undercutting defeaters*'.

How about the independence of the various kinds of undercutting defeaters? The first three kinds – possession-, reasons-, and basing-undercutting defeaters – seem to be independent of each other. A consideration can be about whether some consideration is a reason without concerning the way in which this consideration is treated as a reason; and similarly vice versa. A consideration can also be about whether one really possesses a consideration as a reason without addressing the status as a reason or how one treats it as a reason; and similarly vice versa. (Admittedly, the 'treating as a reason' then has to be possible without possession of the reason – which some philosophers might find questionable.) Arguably, however, all of these three kinds of specific undercutting reasons seem to be *generic* undercutters. (So it is particularly important to not confuse basing-undercutters and generic undercutters. A reasons-undercutter is automatically a generic undercutter, but not automatically a basing-undercutter.)"

Finally, there are the so-called '*higher-order defeaters*' ('HODs', for short) that have figured prominently in the recent debate about higher-order defeat.¹⁵ In some sense, all four of the defeaters mentioned above – reasons-, basing-, possession-, and generic undercutting defeaters – could be called 'HODs'. But a quick survey of the

¹³ A similar notion – of 'side defeat' – has been proposed by Ram Neta. See Neta (2015).

¹⁴ The phenomenon of possession-undercutting has recently been described by De Prado (2020).

¹⁵ See, for example, the collection of essays in Skipper, Steglich-Petersen (2019).

cases that are taken as paradigms in the recent discussion shows that it is mostly or predominantly *basing*-undercutting defeaters that are at stake here. For example, the case of the sleepy detective who acquires evidence that his reasoning is not properly working because of his fatigue is quite clearly to be classified as a basing-undercutter. The status of the facts (that are treated as reasons) as reasons is not targeted. And presumably it is also not the property of being in possession of these reasons that is targeted (though admittedly, this is not entirely clear). Similarly for cases in which some subject acquires evidence that she has taken some reasoning-distorting substance. In any case, from a systematic point of view, the cases in the HOD literature belong to one of the three species or to the generic form. This may vary to some extent from case to case, but the predominant kind of defeat is rather clearly the *basing*-undercutting one. Thus, our taxonomy of undercutters is complete and does not have to be expanded any further.

4 The argument from undercutting defeaters

Relying on the terminological and substantive assumptions elaborated in the last section, we can now state the *argument from undercutting defeaters*. The argument is supposed to show that the capacity for responding to reasons requires the possession of the concept of belief and, simultaneously, also the possession of the concept of a reason. Our focus will be on the concept of belief, but the entire argument is actually working for both concepts, if it works at all.

If the argument succeeds, it provides some support for reflexionism (the one-level view). In addition, the argument is interesting in itself, independently of its bearing on the truth of reflexionism.

The argument can be approached by means of the idea of the *unity of rationality*. Reflexionists tend to think that rationality is a unified phenomenon. To have rationality is to have a whole package of capacities, and it is not possible to have only a proper part of it and still remain rational. Such a partial possession of these capacities would amount to a collapse that would no longer be recognizable as having rationality at all. To have the capacity to respond to reasons, then, requires a whole bunch of sub-capacities that are inseparably bound up together in the capacity known as rationality. The non-reflexionist, in contrast, thinks of some of the elements of the bundle as being separable; one can possess a (central) part of it and yet remain rational.

Now, here is the official formulation of the argument from undercutting defeaters.

- (1) The capacity to respond to reasons requires the capacity to *acquire* beliefs on the basis of *reasons for* belief (including rebutting reasons).
- (2) The capacity to respond to reasons requires the capacity to *give up* beliefs on the basis of *undercutting reasons against* belief.
- (3) Being able to give up beliefs on the basis of undercutting reasons against belief requires possession of the *concept* of belief (and of a reason).
-
- (4) The capacity to respond to reasons requires possession of the concept of belief (and of a reason).

Some comments on the premises of the argument seem in order. The first premise is basically an unshakable and uncontroversial platitude about rational belief formation. Acquiring beliefs by responding to reasons for belief is the very first thing that comes to mind when one starts thinking about rational belief formation. (Rebutting reasons are included, since they are also reasons *for* believing a proposition and not essentially higher-order.) Let me note that the first premise is not needed for deriving the conclusion, of course. However, since the difference between the two capacities mentioned in (1) and (2) is of prime importance, it is useful to have this claim on the table as well.

Probably, the second premise is less obvious. But if one thinks of rational believers as subjects who are able to steer their beliefs in the space of reasons, as it were, then there are of course also the reasons *against* belief, and not only the reasons *for* belief. And quite obviously, a reason against the belief that p is not necessarily a reason for believing not p. Undercutting reasons exactly are reasons of that sort. So it is rather plausible to think that a rational believer should also be able to respond to undercutting reasons by giving up some belief of hers. (The undercutting reasons are

provided by undercutting defeaters of the various sorts discussed above. We will soon have a look at which sort might be of primary interest.)¹⁶

The third premise sounds also initially plausible. Since an undercutting defeater targets some aspect of the subject's response to a first-order reason (i.e., being a reason, being possessed, being properly processed, or all of that indiscriminately), it is already of a higher-order nature. How should a subject be able to respond to such a consideration if she lacks the concepts to understand the higher-order content of the proposition that an undercutting defeater is evidence for? It is not easy to see how this might be possible. So (3) is very plausible, it seems. – Thus, given that the entire argument is valid, we seem to be confronted with a sound argument.

Premise (3) will not be further questioned in the rest of this paper. (Sturgeon (2014) contains some considerations in favor of this premise.) Let me note, however, that it is actually not obvious at all. It might very well be that having the concept of a belief (and a reason) is merely an *acquisition* requirement but not a strict *entailment* of the capacity to respond to undercutting reasons. It might be *metaphysically possible* that there is a subject which possesses the disposition that constitutes safety (sensitivity) towards undercutting defeaters and, therefore, could give up beliefs for undercutting reasons against belief simply by manifesting these dispositions (competences), without having the concept of a belief. If it is possible to non-reflectively respond to reasons by forming a belief, then it should also be possible to non-reflectively respond to reasons by giving up a belief. Why not? On the level of metaphysically necessary conditions, the relevant dispositions can be possessed without having the concept of belief, and this is so for all kinds of undercutting reasons. To put it blatantly, the undercutting reason might simply be a red light turning on. The subject just needs to have the disposition to no longer believe *p* on the basis of the consideration that she treated as a reason for believing *p* in order to follow this undercutting reason - be it possession-, or reasons-, or basing-undercutting. However, it is very hard to see how an ordinary human subject could *acquire* the relevant dispositions without having the concept of a belief and of a reason. In the ordinary, normal course of things, one would have to figure out, by some

¹⁶ The undercutting defeaters might be *misleading*, as familiar from the current debate about misleading higher-order evidence. But there are of course also the *non*-misleading ones, and these are the ones that uncontroversially provide reasons against the relevant first-order belief. Here we can leave it open what the ultimately correct view about how to react to misleading HODs is.

kind of negative feedback or testimony presumably, that the presence of a certain fact (which is the UD) makes one's continuing to believe that *p* on the basis of the consideration that one treated as a reason no longer appropriate. That fact could be possession-, or reasons-, or basing-undercutting. In any case, the subject would have to acquire some understanding of how the presence of this fact indicates some important negative aspect of belief formation. And it is hard to see how such an understanding could be had without having the concept of belief. I am not in a position to prove this point, but it does seem very plausible to me. If testimony is the main ordinary road to learn about undercutting defeaters, how is a child who does not possess the concept of belief supposed to succeed at learning about a reasons-undercutting defeater? (Let us start with *reasons*-undercutting defeater since they seem to be the least demanding.) Suppose that initially the child has learned to treat the thermometer reading as a reason for believing that the temperature is as the thermometer represents it. (The representation might be very coarse-grained, perhaps only indicating 'okay' and 'too hot to touch', say.) Then the parent tells her to ignore the thermometer when the red light is on (where the red light is the reasons-undercutting fact, for example). But what does this mean? The right way to understand the parent's message is to no longer rely on the thermometer *for forming beliefs about the temperature* when the light is red. The child cannot understand that, however, since that requires the concept of belief. Therefore, the child might just as well react by treating the red light as a rebutting defeater (and responding by believing that the temperature is not as displayed) instead of a reasons-undercutting defeater. To learn to treat some condition as a reason-undercutting defeater thru testimony is rather demanding and seems to require an understanding of belief. (For possession- and basing-undercutting defeaters the situation seems to be basically the same, and at least not less demanding.) As an alternative to testimony, we can consider the possibility that the child becomes sensitive to the red light (as something whose presence makes the thermometer no longer trustworthy) merely thru trial and error etc. and without any testimony. Again, however, it is difficult to see how the child could distinguish between rebutting and undercutting defeaters. The child might just as well end up being sensitive to the red light as a rebutting defeater.¹⁷ In sum, the acquisition of a capacity to respond to an undercutting defeater

¹⁷ I admit that this is not obvious. But I think it is very plausible. Please note that the point is not about the (metaphysical) possibility of developing dispositions whose manifestations would amount to treating something as an undercutter. The point is

specifically seems to almost require the concept of belief. So, in the end, the diagnosis is that premise (3) is probably false and that acceptance of premise (3) is likely to depend on a confusion of metaphysically necessary conditions and normal acquisition conditions.¹⁸

Finally, we can come back to our taxonomy of undercutting reasons and address the question *which sort* of undercutting reason is concerned in the argument from undercutting reasons. – The answer is a bit tricky. Initially, it may seem that only *basing-undercutting* defeaters are concerned. For to respond to a basing-undercutting reason one has to form the belief that one's belief is not properly based on reasons, and that surely requires the concept of belief. Thus, we can agree that this kind of undercutting reason is clearly concerned. However, the other kinds of undercutting reasons are also relevant, arguably. For whenever we think of something as a reason, we have to also think of it as a reason *for a certain kind of attitude* – an intention or a belief or It is natural to speak of an 'implicit parameter' here. But arguably, to have the concept of a normative reason requires some understanding of *what it is a reason for*. And if one does not have the concept of belief, it seems one cannot have the concept of a *reason for belief* either. So, contrary to initial impression, the other kinds of undercutting defeater are also concerned in the argument. We often do speak elliptically when we speak of 'a reason', leaving out mentioning what it is a reason for. But this should not blind us to the actual requirements that hold on the level of conceptual thought.¹⁹ Thus, we can conclude that the argument is probably not only about basing-undercutting defeaters but, instead, about the whole range of undercutting defeaters. – Not a problem for the proposed criticism of reflexionism as it is not restricted to a special subset of undercutting reasons.

about learning. Without an understanding of the difference between an undercutter and a rebutter it is very hard to see how someone could learn that a particular condition plays the role of an undercutter rather than a rebutter. – Thanks to an anonymous referee for valuable discussion on this point

¹⁸ Special thanks to Daniel Whiting and an anonymous referee for helpful discussion and comments on this point.

¹⁹ Perhaps, possessing-undercutting reasons seem to be an exception. However, even there it is not entirely clear whether one can respond to such a reason without having the concept of belief, since it is about possession of a fact in a certain role – the *role of a reason for belief* (and not for intention or emotion). However, if possession of a fact in the role of a reason is invariant as to what it is a reason for, possession-undercutting defeaters might indeed be an exception.

5 An objection against the argument from undercutting defeaters

Though the argument seems sound, a critical examination yields the verdict that *the second premise should be given up*. To argue this is the first main goal of this section.

Rejecting the second premise calls for further explanation of how we are to understand responding to reasons. How can someone be merely in the business of responding to reasons *for* belief but be incapable of responding to undercutting reasons *against* belief? How are we to conceive of such a rational believer? Could we really call her a rational believer? Or wouldn't it rather be, at best, a pseudo-rational believer? The second main goal of this section is to provide a story, or at least a sketch of its main pillars, that addresses these explanatory demands.

There is a rather simple answer to these questions. It will bring up some objections that we will take a look at in what follows. But in the first place, we can state this simple answer as follows: forming and giving up beliefs is essential to being a rational agent. But it can be done without any *higher-level assessment of reasons*, and thus does not require the special ability to give up beliefs in response to *undercutting* reasons against these beliefs.

We can bring out this priority more precisely by the following consideration. Our ability to give up a belief for undercutting reasons against it requires having higher-order evidence about our first-order belief formation. But if all our beliefs were formed without the exercise of an ability to form beliefs for reasons *for* belief, we could never acquire a well-founded, rational belief that could *serve as the basis for rational action*. We could eliminate some non-rational beliefs (supposing that the necessary higher-order evidence is available), but not acquire any rational, well-founded belief. Arguably, however, rational action requires rational belief as a basis. (Some philosophers believe it requires even knowledge, but perhaps knowledge is simply the special case of rational belief which is formed for conclusive reasons for belief.)²⁰ The principle that rational action requires rational belief is very plausible. So we can infer that a subject without the ability to form beliefs by responding to reasons for belief would not be capable of rational action. And that is surely a strong reason to think that the ability to form beliefs

²⁰ See, for example, Hawthorne, Stanley (2008) who argue that knowledge is necessary for treating a consideration as a reason for action.

for reasons *for* belief lies at the base of rational subjectivity. At the same time, there is no corresponding need for having the – very special – ability to give up beliefs by responding to *undercutting* reasons against belief. The function of rational belief, to serve as an input to rational action, assigns a *priority* to the ability to respond to reasons for belief, but not to the ability to respond to undercutting reasons against belief.

In other words, once we look at the roles that abilities to respond to reasons play for subjects, we can realize a priority of the ‘positive’, ‘positing’ ability to respond to reasons for belief over the ‘negative’ ability to respond to undercutting reasons against belief. (This is not to deny that, of course, having both would be best.) To have some beliefs based on reasons is more important to rational agency than to eradicate beliefs by means of a grasp of undercutting defeaters.²¹

The just mentioned fundamental priority provides us with a reason to think that the requirement of the first premise is non-negotiable whereas there is no corresponding reason (of more or less the same kind) to think that the requirement stated by the second premise is equally non-negotiable. Is there any other reason to think that the second premise is true? In order to address this issue, we have to bring in some further considerations having to do with how various sorts of responding to reasons operate in our lives.

One further consideration turns around the role of *giving up beliefs*. Suppose that a subject S has the ability to form some rational beliefs (by responding to reasons for belief) but no ability to give up beliefs by responding to reasons against belief. Now, she accumulates belief after belief but never gives up any belief for a reason *against* belief. She either preserves the belief or gives it up arbitrarily, non-rationally (i.e., not by responding to reasons against belief). Wouldn’t this be a very odd situation?²²

Of course, there is a mistake in this line of thought. A look at some cases can be very fruitful here. Suppose that a subject, Liz, acquires the belief that someone is

²¹ One might think that this is not generally true but depends on the general characteristics of the relevant environment or situation. But this is not so. The argument appeals to the strict necessity of having rational beliefs for acting rationally, it does not matter how often one has beliefs that are not properly based on reasons or how often higher-order evidence is available.

²² Of course, if x is a reason *for* believing that p, it is at the same time a reason *against* believing that not p (or other propositions that are made unlikely by p). But the scenario we are considering here can be specified further by adding that the subject never comes across any rebutting reason. (We will come back to rebutting reasons soon, at the end of this section.)

standing at the door by responding to the bell's sound as a reason for that belief. Now, she does not go and look at the door but instead continues her phone call, say. Then, after a while, her belief that someone is standing at the door becomes outdated and should be given up. But how could she rationally do this if she did not have the ability to respond to undercutting reasons *against* belief? – The answer is not that hard to find. In cases like Liz's, the subject acquires what we might call an *episodic belief*, and she does so on the basis of an *episodic reason*. The fact that the bell is ringing is only an episodic reason, i.e., a reason for having an episodic belief. This episodic belief is short-lived. One could express this point by saying that its content has a 'now-form'. Liz comes to believe, at time t1, that *now* someone is standing at the door. The reason that she is responding to does not license any belief with a more permanent temporal dimension that transcends the *now*. (This *now* need not be instantaneous but can last for a while, of course. And there is some amount of vagueness, to be sure, here as elsewhere.) It is a reason for an episode of belief merely, but not for a standing belief. And therefore, giving up the belief once the *now* has passed by is part and parcel of (properly) responding to this episodic reason. *No further, distinct reason for giving up the belief is needed*. To assume otherwise would amount to an over-intellectualizing misunderstanding of the nature of episodic reasons and beliefs.²³

It is not hard to conjecture that episodic reasons and episodic beliefs are legion in our lives. The small child is mostly engaged in forming such episodic beliefs. 'Now Mommy is here', 'Now there is a red ball on the table', 'Now it's cold', and so on. Particular short-lived facts provide plenty of the – episodic – reasons for these beliefs. This remains true in adult cognition, though some non-episodic reasons and non-episodic, permanent beliefs show up as well. All of this, I take it, is rather known by reflection on common sense and needs no specialized experimental tools to be uncovered. But it could easily be overlooked if we focus too hard on *undercutting*

²³ Please note that the notion of 'giving up a belief' employed here is a light one. One does not have to actively make the belief go away (as, for example, if one responds to an undercutting defeater). One can give up a belief merely by no longer holding the belief. To me it seems plausible that the simplest episodic beliefs, as rational responses to episodic reasons, have this in-built self-consuming nature. An episodic now-belief may, but need not lead to a corresponding memory or then-belief after the now has passed away. – As an anonymous referee pointed out to me, this is likely controversial. But since the main idea does not seem entirely hopeless to me, it still makes sense to put it on the table, I believe. I am grateful to this anonymous referee for challenging me to clarify this idea a bit more.

reasons that serve as a rational basis for giving up beliefs. In many cases, no such undercutting reason is needed in order to do the rational thing, i.e., to give up the relevant belief after the relevant time (fixed by the episodic reason for acquiring it). A change calls for a doxastic event, an episodic belief. And giving up the belief is part of this (proper) response.²⁴

It would involve a kind of overintellectualization to think of these cases as involving the exercise of a concept of a *particular* time, e.g., the concept of I am on 1st of October 2022. Often, such a concept of a particular time will not be available since the subject does not know what time it is. And arguably, small children lack such concepts while already possessing the concept of now.

Another possible line of thought in favor of the equal importance to rationality of the ability to respond to undercutting reasons appeals to the two-foldness of our epistemic goal: to attain truth and to avoid falsity. Undercutting reasons would allow us to discover that some consideration we rely on is in fact no reason that we could properly rely on, and that would help us to avoid false beliefs. So, responding to undercutting reasons is needed to serve the second basic goal, avoiding error. – However, it is quite clear that following reasons can already serve both goals equally well. If one relies on a consideration which in fact is a good reason, one is likely to attain true beliefs and to avoid false beliefs. One could perhaps do even better if one could follow undercutting reasons as well. But the nature of following reasons for belief already provides some substantive mileage toward avoiding error, and not only toward attaining truth.²⁵

We can adduce another point to prop up the preceding consideration. *Rebutting* defeaters can also call for *giving up* a belief. For, a rebutting defeater – against the belief that p – is not only a reason for acquiring a new belief, the belief that not p is the case, but also a belief against the old belief that p. Suppose that Shahid responds to the bell's sounding by forming the belief that someone is at the door (p); he follows his reasons for believing so. Later on, he acquires a decisive reason to believe that it is *not* the case that someone is at the door (not p), simply by going to the door, opening it, and

²⁴ What we are belaboring here is in the vicinity of what Gareth Evans has described as 'cognitive dynamics'. (Cf. Evans 1982.) Please note, however, that the dynamics of the simplest episodic beliefs that I am drawing attention to might not conform to Evans' conception in all respects.

²⁵ Thanks to an anonymous referee for suggesting this line of thought to me.

taking a close look. By vision he becomes aware of a strong rebutting defeater. And responding to this rebutter consists not only in acquiring the belief (that not p), but also in giving up the old belief (that p). Thus, giving up a belief can be a response to a *rebutting* defeater, and no additional *undercutting* reason needs to come into play. This adds to the previous point that giving up an episodic belief is part and parcel of responding to an episodic reason. One can also give up a belief in response to a rebutting defeater. (In case the rebutter is equally strong as the initial reason for believing that p, the subject can respond by giving up the belief that p and suspending on p, of course.)²⁶

6 Conclusion

Let us take stock. We have seen the simple answer to the question of how to make sense of subjects that have an ability to form and give up beliefs (for reasons for or against belief) but not to give up beliefs for *undercutting* reasons. The fundamental priority grounded in the belief-action connection provides the essential insight. The phenomenon of giving up beliefs does not yield an argument for the necessity of having the ability to give up beliefs for undercutting reasons. For in many cases, giving up a belief is just a matter of the cognitive dynamics of episodic reasons and episodic beliefs or is part of the proper response to first-order reasons for and against belief. So, the second premise stands unargued yet. And *prima facie*, I take it, no further reason can be spotted in the neighborhood of undercutting defeaters.²⁷ To simply assume that rationality must be ‘unified’ and, thus, requires the very special rational ability mentioned in (2) would be question begging. We can state that there is no principled

²⁶ Special thanks to Thomas Raleigh for helping me to see this point.

²⁷ One could think of the distinction between *prima facie* and *ultima facie* justification (well-foundedness). Roughly, a belief is *ultima facie* justified only if it is *prima facie* justified and this *prima facie* justification is not defeated. But if a subject lacks the ability to conceptualize undercutting defeaters (since she lacks the concept of a belief and of a reason), how could she have *ultima facie* justified beliefs? – The answer to this question is not hard to find. The *distinction* itself does not collapse if there is a subject who lacks the ability to respond to undercutting defeaters. When she forms beliefs by responding to reasons for belief, we can say that her belief will be *prima facie* justified and also automatically *ultima facie* justified. We can say that the two standings *coincide* for such a subject. The distinction itself, however, is still in good shape and need not be given up, of course.

reason for accepting the second premise, though there is such a reason to accept the first premise.

This concludes the defense of the recommended view: accept (1) and deny (2). The role of forming rational beliefs by responding to reasons for belief is much more basic than the role of giving up beliefs by responding to *undercutting* defeaters. So it is perfectly possible and makes perfect sense (from the point of view of cognitive architecture) to have the former and lack the latter. Plausibly, this is the case with small children and some other non-human animals. We can keep our practice of ascribing rational beliefs to them, i.e., beliefs that are formed in response to reasons for belief, since we are not forced to re-interpret these cases as non-rational, as the reflexionist has to. And we have made room for the idea that only *later*, children can acquire reflection and the – quite special – ability to respond to *undercutting* reasons. It is a further form or level of rationality that we grow into when we acquire reflection and, thus, become able to exploit higher-order evidence (most of it provided rather by testimony than by introspection, I believe) in the pursuit of weeding out non-rational beliefs.²⁸ This provides a picture of how rationality can be acquired that is in line with a plausible evolutionary and developmental perspective. All forms or levels of rationality can be understood as capacities to respond to reasons. But they are not all bound together in just one cluster; rather, they can be acquired in a stepwise fashion. Thus, some exercises of rationality are non-reflective, some are reflective.²⁹

Acknowledgement:

For valuable discussions and comments I am grateful to Yannick Kohl, Susanne Mantel, Thomas Raleigh, and Daniel Whiting.

²⁸ This does not mean that reflection is infallible, of course. It might very well happen that one mistakenly gives up a fully rational belief in response to misleading higher-order evidence. But the acknowledgement of fallibility is no reason to become skeptical about reflection. Fallibility is the rule rather than the exception.

²⁹ How about the *reflective* way of responding to a reason *for* belief? – A promising idea is that in these cases, one responds to the normative fact that the fact which is the reason is a reason for the relevant belief. Then one needs an ability to safely/sensitively form the belief with respect to this normative fact, just as when responding to the fact which is the reason. To explore this idea would require further investigation that cannot be undertaken in this paper.

Funding support statement:

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.

References:

Boyle, Matthew 2012. "Essentially rational animals." In: *Rethinking Epistemology Vol. 2*, edited by Günter Abel & James Conant, Berlin: Walter de Gruyter 395-428.

Burge, Tyler 1996. "Our entitlement to self-knowledge." *Proceedings of the Aristotelian Society* 96, 91-116.

Dancy, Jonathan 2018. *Practical Shape*, Oxford: Oxford University Press.

De Prado, Javier González 2020. "Dispossessing defeat." *Philosophy and Phenomenological Research* 101:2, 323-340.

Evans, Gareth 1982. *The Varieties of Reference*, Oxford: Oxford University Press.

Hawthorne, John, Stanley, Jason 2008. "Knowledge and action." *Journal of Philosophy* 105:10, 571-590.

Hofmann, Frank 2022. "Explaining free will by rational abilities." *Ethical Theory and Moral Practice* 25:2, 283-297.

Hofmann, Frank 2021. "Is evidence normative?" *Philosophia* 49, 667-84.

<https://doi.org/10.1007/s11406-020-00241-2>

Kant, Immanuel 1998. *Critique of Pure Reason*. Cambridge Mass.-London: Cambridge University Press.

Littlejohn, Clayton 2018. "Reasons and theoretical rationality." In: *Handbook of Reasons and Normativity*. Edited by Daniel Star, Oxford: Oxford University Press, 529-552.

Lord, Errol; Sylvan, Kurt 2019. "Prime time (for the basing relation)." In: *Well-Founded Belief; New Essays on the Epistemic Basing Relation*, edited by Joseph Adam Carter & Patrick Bondy, London: Routledge, 141-173.

Mantel, Susanne 2018. *Determined by Reasons. A Competence Account of Acting for a Normative Reason*. London: Routledge.

McDowell, John 2006. "Conceptual capacities in perception." In: John McDowell 2009. *Having the World In View*. Cambridge-Mass.: Harvard University Press, 127-144.
(Reprinted from: *Kreativität*. Edited by Günter Abel, Hamburg: Felix Meiner, 1065-1079.)

Mitova, Veli (ed.) 2018. *The Factive Turn in Epistemology*. Cambridge-Mass.-London: Cambridge University Press.

Neta, Ram 2015. "Basing as conjuring", *Minds Online Conference*, available on the internet: <https://mindsonline.philosophyofbrains.com/2015/session3/basing-is-conjuring/>

Pollock, John 2001. "Defeasible reasoning with variable degrees of justification." *Artificial Intelligence* 133, 233-382.

Pollock, John 1987. "Defeasible reasoning." *Cognitive Science* 11, 481-518.

Rödl, Sebastian 2007. *Self-Consciousness*, Harvard University Press.

Skipper, Mattias, Steglich-Petersen, Asborn 2019. *Higher-Order Evidence: New Essays*, OUP.

Sturgeon, Scott 2014. "Pollock on defeasible reasons." *Philosophical Studies* 169, 105-118.

Turri, John 2010. "On the relation between propositional and doxastic justification." *Philosophy and Phenomenological Research* 80:2, 312-326.

Weatherson, Brian 2008. "Deontology and Descartes' Demon." *Journal of Philosophy* 105:9, 540-569.

Wimmer, Heinz; Perner, Josef 1983. "Beliefs about beliefs." *Cognition* 13, 103-128.