Beyond 'Rationalism': how reasons explain action

Jason Bridges
University of Chicago

9/24/23

1. Introduction

A recurring theme in Joseph Raz's writings is that it is an error to assume our reasons for action will single out one option, from all those available to us, as *the* thing to do. On the contrary, "in normal circumstances", the agent, facing a choice, finds "more than one response supported by reasons, with none of them supported to a higher degree than any of the others" (2011, p.5).

Raz deems this point consequential enough to earn the label, "the *basic belief*" (1999, p.111). Its chief significance lies in frustrating what is otherwise an attractive view of how our thought about reasons for action leads us to act. On the "rationalist conception of human agency", "the paradigmatic human action is action taken because, of all the options open to the agent, it was, in the agent's view, supported by the strongest reason" (1999, p.47). In fact, the paradigmatic human action is one taken when there is no option "supported by the strongest reason". So the rationalist conception cannot be correct.

But if reasons do not typically identify a unique option as that which is to be chosen, how do they explain the particular choices agents actually make? Raz's answer to this question is to bite the bullet: an agent's conception of her reasons, in the paradigmatic case, is *not* sufficient to explain why she does what she does. We are led to "the classical conception of human agency", on which the agent's conception of her reasons helps explain her action only in tandem with factors external to that conception. If "reason" is "the capacity to recognize and respond to

reasons" (2011, p.92) and the "will" is "the ability to choose and perform intentional actions" (1999, p.48), then the upshot is a picture of "a will that is informed and constrained by reason but plays an autonomous role in action" (1999, p.48).

I will argue that Raz is right to embrace the *basic belief* and reject the *rationalist conception* on its basis, but wrong to conclude from this that the *classical conception* is correct. The choice between these alternatives is false one. It is an artifact of Raz's retaining from the *rationalist conception* an assumption we ought rather to abandon, namely: that explaining an intentional action must involve explaining why the agent chose that action rather than the alternatives. Rejecting this assumption will enable us to effect a more radical break from rationalism than Raz is able to muster. But to get into a position to see through the assumption, we will need to examine the logic of practical reasons from a vantage point which is rarely taken up in contemporary practical philosophy.

2. Raz against the rationalist

The rationalist reasons as follows:

> To rationalists, the fact that intentional action is undertaken in the light of the agent's understanding of his situation suggests that the agent must always be capable of finding an answer to the question, 'What am I to do?' There are always factors—we call them reasons—that guide the agent's choices and decisions. If there were incommensurabilities, then actions would be unintelligible to the agents who perform them. They would not be able to explain why they performed the action they did rather than one of the other options open to them. All they would be able to say is: 'We saw that there is no reason to prefer *A* to *B*, or the other way around, and we did *A*.' The obvious gap in this explanation will baffle not only the observer who is trying to explain or predict people's behavior. It will defeat the agents themselves, who would regard their choices as a mystery, as something that happens to them rather than something they do. According to the rationalist, incommensurability is inconsistent with the fact that intentional actions are under the control of the agents, that they are determined by their choices. (Raz, 1999, p.49)

In this passage the rationalist is said to be worried about the prospect of "incommensurable" options, options which admit of no positive comparative assessment. But the logic of their argument, as Raz makes clear elsewhere, means that they will worry no less about the prospect of commensurable actions assessed as equally good. The problem the rationalist sees would arise in any case with "none of [the options] supported to a higher degree than any of the others" (2011, p.5).

The skeleton of the rationalist's argument is this. An intentional action is one undertaken in thought of what you have reason to do. Doing *A* intentionally, you can explain why you do *A* in terms of the reasons as you see them. But suppose you did not see stronger reasons for doing *A* than for doing *B*. Then you could not appeal to your reasons to explain why you do *A* rather than *B*. But then you could not after all explain why you do *A*—that particular thing—in terms of your reasons either. It follows that intentional action would not be possible in such a circumstance.

Let us say that a conception of one's reasons *singularizes* upon an action *A* when it identifies stronger reason for doing *A* than for any available alternative. Then the conclusion of the rationalist's argument may be put this way:

> *Rationalist's conclusion*: If it is to be possible perform an action intentionally, one's conception of one's reasons must singularize upon that action.

We may also use this terminology to reframe the *basic belief*:

> *Raz's "basic belief":* In paradigmatic cases of intentional action, the agent's conception of her reasons does not singularize upon her action.

The *basic belief* flatly contradicts the *rationalist's conclusion*. So Raz, holding to the *basic belief*, must find an error in the rationalist's argument.

One conspicuous premise of the argument is that if the reasons the agent sees cannot

explain why she does *A* rather than some other available alternative, they cannot explain why she does *A*, period. Raz does not reject this premise. But he gives it a twist, and this twist opens conceptual space he takes the rationalist to have overlooked. It is in this space that the *classical conception* of human agency is situated.

Raz shares the rationalist's premise to this extent: if agents' conceptions of reasons do not singularize upon particular actions, then "reason can neither determine nor completely explain their choices and actions" (1999, p.48). Here the word "completely" is not a dispensable flourish. It expresses a subtlety that is key to Raz's reasoning. The rationalist assumes that if the agent's conception of reasons cannot explain why she does one thing rather than another, it could not be enlisted in the explanation of her action at all. But this, for Raz, is to overlook the prospect of an explanation that is partly in terms of the agent's conception of her reasons and partly in terms of something else—where it is this further part of the explanans that carries the burden of explaining why the agent takes the action she does rather than the alternatives.

To explain *why an agent does A rather than B* in terms of her reasons, we would need to find considerations which show doing *A* to be preferable to doing *B*, i.e., to be supported by stronger reasons. The *basic belief* tells us such considerations are not typically to be found for all of the alternatives available to the agent. But why should we insist that the contrastive question be answered in terms of reasons? There seems room for a different thought. An intentional agent responds to the reasons she has, yes, but insofar as those reasons fail to single out a unique alternative as best, there must be something else—say, some further psychological force or tendency—that gets her all the way to a determinate choice.

In sum: the rationalist is *right* to hold that an explanation of an intentional action must appeal to the agent's reasons, and *right* to insist that such an explanation must explain why the agent does what she does rather than the alternatives, but *wrong* to assume the explanation

must accomplish the latter via the former. Taking this line, we join the rationalist—and indeed, the whole tradition since Anscombe and Davidson—in conceiving intentional action as action explicable in terms of the reasons the agent sees for it. We grant that an agent who acts intentionally acts in light of a consideration she knows or presumes to support so acting, and that one may thus cite that consideration in explaining her action. We grant, in a nutshell, that doing *A* intentionally = doing *A* for a reason. What we deny is only that an intentional action must be such as to be explained in terms of reasons *alone*. The *basic belief* shows us to the contrary that, in "the standard situation for choice and action":

> …reason cannot determine what we do. Whatever we choose to do we will do for a reason, that is, for the reason which, as we see it, supports the option we pursue. But as we believe that other reasons, not defeated by it, support alternative options, that cannot be the full story of why we do what we do. (Raz, 1999, p.111).

We may summarize Raz's argument, labeling the premises for convenience, thusly:

> 1. *Intention-reason principle*. When an agent does *A* intentionally, the complete explanation of *why she does A* will be at least in part in terms of her conception of her reasons.
>
> 2. *Contrastive claim*. The complete explanation of *why an agent does A* must explain *why she does A rather than the alternatives*.
>
> 3. *Singularization premise*. The agent's conception of her reasons can explain *why she does A rather than the alternatives* only if it singularizes upon her doing *A*.
>
> 4. *Basic belief*: In paradigmatic cases of intentional action, the agent's conception of her reasons does not singularize upon any available action.
>
> ∴ 5. *The classical conception of human agency*. In paradigmatic cases of intentional action, the complete explanation of *why the agent does A* will be partly in terms of her conception of her reasons and partly in terms of something else.

As I indicated, Raz likes to expound the *classical conception* in terms of a distinction between faculties. On the one hand, there is "reason", the capacity "to recognize and respond to reasons" (2011, p.92). And on the hand, there is the "will", which is "the capacity for intentional action"

(1999, p.111). We might hope to identify reason, at least in its practical operations, with the will. The *classical conception* teaches that this is an error. Reason contributes to exercises of the will. Its contribution is traced in that part of our complete explanation of an action which appeals to the agent's conception of her reasons. But because there will typically need to be a further part to this explanation, the will is not merely an expression of reason. It has the freedom to respond to factors outside reason's province:

> [T]he will plays a role in human agency separate from that of reason, a role that neither kowtows to reason by endorsing its conclusions nor irrationally rebels against it by refuse to endorse them. (Raz, 1999, p.65)

How are to understand the workings of the will insofar as they transcend reason's influence? Raz gestures toward one idea. This is of a psychological phenomenon of "thick desire" (1999, p.110), of a "want" understood "in the thick sense" (1999, p.111). A thick desire or want is a species of "attachment not required by reason", but which may nonetheless conform with what we have reason to do or to seek. But Raz does not regard his treatment here as sufficient. He says that "much work needs to be done to analyze the different ways in which our will leads us to do one thing rather than another" (1999, p.65). "It is a subject deserving much more exhaustive examination" than he provides; indeed, he worries his focus on topics having to do with reason and reasons "is liable to distort the greater picture by overrating the role of our rational capacities" (2011, p.5).

3. Explanatory completeness: what is at issue?

The *contrastive claim* we found to be at work in Raz's argument is that the complete explanation of an action will say why the action was taken rather than the alternatives to it. Such significance and plausibility as this claim can possess for us will depend upon our knowing

what is here supposed to be meant in speaking of the "complete explanation" of an action. Do we know this?

There may seem no special problem about what is meant. The "complete" explanation of why something happens or comes to be so will be the explanation which leaves nothing out. It will tell us everything there is to be known about why the thing in question happens or is so. Or, if this seems too ambitious, it will tell us everything essential.

But these glosses do not help us understand the contrastive claim. On the contrary, they show that without further clarification, the claim cannot be understood.

Two related points may be observed. On the one hand, on actual occasions for explanation, what we think essential for understanding why something happens or is so will depend upon such circumstances as what we are interested in, what our purposes are in seeking understanding, and on what we already know or can assume as common ground among some intended or notional audience. And on the other hand, there is simply no such thing as an explanation of why something happens or is so which enumerates everything anyone might conceivably think to know, or find important, in answer to that question—the ideal of "absolute" completeness, as we might try to put it, is a mirage. Taken together, these points imply that the contrastive claim has nothing to tell us until it is determinately given what kind or form or style of explanation of action the claim is supposed to concern.

Consider a fragment of conversation:

Y: Why did you drive downtown this afternoon?
M: My reason was to get a haircut.
Y: But you could have taken the train instead. Why didn't you?
M: The new train schedule, reduced as it is, meant the options were too inconvenient.
Y: Why did the city switch to such a feeble schedule?
M: I'll tell you why. It's the mayor's shortsighted handling of the budget.
Y: Ugh. Why was he even elected?
M: Therein lies a tale. I was just reading a book on the history of machine

politics in this city, and…

Everything *M* says in this conversation, including whatever she goes on to report from her reading, may reasonably be viewed as relevant to the question why she drove downtown that afternoon. Absent any indication of the particular concern of our inquiry, there can be no basis either for saying that *M does not* reach a complete explanation of her action before her contrastive second answer, or that she *does* reach a complete explanation with her second answer, so that everything said afterward is superfluous. Often enough, *Y* would be perfectly content with the first answer alone. It would tell her just what she wanted to know in asking after *M*'s action. In other cases, it will rather be some answer downstream of the second answer that is seen to contain the crucial information. As *Y* might later find occasion to say, "My friend *M* drove into our traffic-clogged downtown the other day, and I'll tell you exactly why she did that. It's our mayor's shortsighted disinvestment in public transportation. Throw the bum out!"

Wittgenstein tells us: "an explanation serves to remove or to avert a misunderstanding—one, that is, that would occur but for the explanation; not every one that I can imagine" (*PI* §87). The point he is making does not require reference to a *mis*understanding per se; we may equally say that an explanation is designed to fill a particular gap in our understanding, not every conceivable gap.

Wittgenstein made this point because he thought that, when doing philosophy, we were prone to hanker after a "final" explanation, one which does not depend for its being grasped upon any understanding that it does not itself provide. He reminds us that this prospect is an illusion. All explanation succeeds, if it does, only in exploiting a background of prior understanding. By the same token, talk of an explanation as "complete" or "incomplete" has no content absent a determination of the particular gap in understanding the explanation is designed to fill. If, in particular, the *contrastive claim* is to have import, we need to know what

gap in understanding—or again, what kind or form of explanation—it purports to condition.

There is a related point to be made here. Saying *why I do A rather than the alternatives* can be a requirement on a complete explanation of *why I do A* only if there is actually something to be said on the former score. The fact that one can formulate a question does not show that it is well-taken or has an answer. It is true that *M* could answer the contrastive question in our example. But must the agent, or anyone, always be able to find something to say here? We should not take this for granted.

Only if it can be shown that answering the contrastive question is required for a complete explanation of an intentional action will we have reason to believe that this question must have an answer. But if we are to be shown this, we will first need to know what kind or form of explanation of intentional action is at issue.

4. Explanatory completeness: reasons for action

Let us set aside the argument for the *classical conception* for the moment. It is clear, at any rate, the kind of explanation the rationalist has in mind when they impose *their* contrastive requirement. The rationalist thinks an explanation of an action *in terms of the agent's reasons* is incomplete—indeed, they would say, it has not even got off the ground—if it does not show the agent's conception of reasons to singularize upon that action. Their thought is that the contrastive requirement is internal to the reason-giving form of explanation as such—to "rational explanation", in the parlance of the field. The rationalist invites us to recognize that when we endeavor to explain an action in terms of the agent's reasons for it, we have taken on a task whose successful execution depends upon showing the action to be preferable, in the agent's view, to all alternatives. That is why to accede to the *basic belief* would be to give up on the prospect of rational explanation.

Is this a plausible conception of the shape of the rational-explanatory project? It is certainly a pervasive one. Raz is right that his rationalist makes explicit an idea which lies in the background of much thinking in contemporary practical philosophy. And so quite apart from an interest in assessing Raz's arguments, it is worthy of note that the rationalist's idea has no plausibility. The design the *rationalist conception* would impose upon rational explanation is imagined whole cloth. It bears no relation to the pattern that is in fact there to be found.

Let us begin with a consideration of the primary act of rational explanation: the giving of *a reason for which* the agent does what she does. This act is primary in that (a) it is present in every instance of rational explanation, and (b) it is often all that is present. On the one hand, you cannot answer the question, "For what reason does *S* do *A*?", without giving a reason for which *S* does *A*. And on the other hand, it would be perfectly natural should *M* say only, "My reason was to get a haircut", and leave the matter at that.

What can be said about the reason for which an agent acts? One simple observation is that in giving such a reason, we identify something the agent took to support, to favor or justify, her action. So much everyone agrees, or should agree. And for our purposes we can ignore the debates, perhaps not especially illuminating, about whether the "reason" for which the agent acts is precisely the same as the "reason" which, in so acting, she sees to favor her action.

It might seem that the foregoing observation is the only thing that can be said with complete generality about the reasons for which we act. It is certainly true, as philosophers of action have documented, that we can find a great deal of variety in reasons for action if we look closely enough. But in dwelling on this diversity, we are prone to miss the possibility of a further general observation about reasons for action. The observation is abstract, but it will turn out not to lack for significant consequence.

The further observation is this: to give a reason for which an agent acts is to identify a *point*

the action has for her. It is to say, *here is what, in so acting, she seeks to achieve, to accomplish.* Or again (for this is just another way of saying the same thing), it is to give an end—a purpose, an aim—for the sake of which she acts.

Let us call what has just been stated in three equivalent ways the *ur-observation* about the reasons for which agents act. This might seem a misnomer. The claim might seem not a basic observation but a tendentious philosophical thesis. In the contemporary literature, the much-debated "teleological conception of practical reasons" holds that all reasons for action concern the "promotion" of "ends".[1] This might look identical to the ur-observation. It is not. The formulation of the "teleological conception" hinges upon the assignment of a technical meaning to "end", on which it is equivalent to "consequence" or "state of affairs".

Those who object to the "teleological conception" challenge it to make sense of claims about reasons of which the following may be taken representative: "while I realize there will be only negative consequences if I resign, it would be dishonest for me to remain in this position now that I know I got it unfairly, and that is all the reason I need." We do not need to consider what resources the "teleological conception" may have for analyzing this statement. What matters for us is just to observe that the statement is not even a putative counterexample to the ur-observation. It is rather a clear supporting instance. The agent who makes the remark tells us the point they see to resigning; they tell us what, in resigning, they seek to accomplish. By resigning, they will act with honesty, with integrity. Integrity is not a state of affairs. It is, as we might put it, a way of life. And its connection to particular actions is not to result from them, as a state of affairs might, but to be manifested in them. This is evidently a significant difference in metaphysical category. But it is not a difference which prevents us from having integrity as an end of what we do.

[1] See Scanlon (1998), Portmore (2018).

Here I use the word "end" with the meaning it has in our language. But the point is not about "ordinary language". It is not about the use, or meaning, of words. Our topic is not what "end" or other such words mean, but what, using these words with the meanings they have, we find we are able to say about the general character of acting for a reason. What we find we are able to say, with complete generality, is that to act for a reason is to act in pursuit of an end.

One may wonder how we are to know this equation to be true. Or conversely, one may wonder whether its statement comes to anything at all. If we refuse to assign "end" (or correlatively "pursuit" or "aim") a special philosophical meaning, can the equation tell us anything about reasons for action that is, as it were, philosophically actionable?

We can see both the truth and the significance of the ur-observation by noting one of its corollaries. If I am to do something for the sake of an end, I must think the action apt for pursuit of that end. The ur-observation thus implies that one who does something for a reason always acts with a certain judgment. The judgment will be of the form *by doing A, I will (or may) ….* What goes into the ellipsis will not merely repeat the description which substitutes for "*doing A*"; the judgment is not tautological. Rather, the specification will be such as to render the judgment substantive, with its truth turning on whether doing *A* really is such, in the circumstances, that this further thing may come of my doing it.

Such specifications, as we have already begun to notice, can take a variety of forms. Some examples which show the variety are:

> *By kicking the garden gnome I will knock it over.*
> *By signing up for this class I will improve my German.*
> *By throwing out all my alcohol, I will benefit my health.*
> *By taking this stand, however futile, I will be true to my code.*
> *By lending her the money, I'll help her through this difficult time.*
> *By coming along I'll have fun.*
> *By relocating the murder weapon, I will bring it about that suspicion falls on someone else.*

If we take this diversity at face value, we will say that an end of an action can be, at least: a

further action, an ongoing activity, a quality (or perhaps form or mode) of life, a rule of conduct, the wellbeing of another, an experience, a state of affairs. We do not need to consider the questions of taxonomy and systemization that here arise.[2] What matters for present purposes is just the necessary presence of a judgment of the general form.

It's true that we may cite judgments which are not of this general form when asked to give the reasons for which someone does something. But reflection on the possibilities here will reinforce, rather than undermine, the present point. For while it is true that anything we judge true might intelligibly factor into practical reasoning and thus be said to be a reason to which the agent responds in acting as they do, this figuring has a condition. The condition is that what the agent judges true be seen by her to bear in some way or another on the prospect of her achieving some end with her action. A single example will perhaps be a sufficient illustration of this mechanism. You see me coming into the house and ask my reason for coming inside. I say, "It's raining." In being told this, can you be said to have learned my reason for coming inside? We are inclined to say yes. But that inclination is inseparable from our tacit assumption that one who responds to the fact of rain by coming inside is one who *seeks to avoid getting wet*. It is because assumptions such as this belong to the conversational common ground that we can, in such a way, be told the reason for which the agent acts. Consider that my answer would have been misleading if, in fact, I was going inside rather *to get wet*. I am, let's say, eight-years-old, and my plan was to grab my rain boots so as to do some puddle stomping.

[2] The question of how to understand, and systematize, these various categories of end might seem a natural subject for the philosophy of action. In fact, this subject has been more or less entirely ignored. Part of the reasons for this, I think, is the influence of Davidson (2001). The judgment to which I am drawing attention as a consequence of the ur-observation corresponds to the belief component of Davidson's concept of a "primary reason". Davidson use of this concept was almost entirely to explore the question of whether rational explanations can be treated as causal-psychological. This framing turned subsequent attention away from a consideration of the forms that go into specifying the diverse contents of such beliefs/judgments. Add to this a widespread disinclination, symptomatic of the so-called "naturalistic" worldview, to take teleological concepts at face value, and the neglect is perhaps unsurprising.

A parent who mistakenly thinks I have the sensible end of staying dry is right in thinking that, in coming inside, I am responding to the fact that it is raining. But she is wrong about the reason for which I act.

This is as much as I will do here to motivate the ur-observation.[3] I have worked to bring it into view because it has an immediate implication for our present topic. The implication is this: to see a reason for an action is not thereby to see a reason against alternatives to that action. On the contrary, to see a reason for an action is to see the possibility, and often likelihood, of there being the *same reason* for alternatives to that action.

Why so? An agent sees a reason for an action in having an end which she thinks the action may serve. That a given action may serve a given end has no tendency to entail that it is the only available action that may serve that end. There is no principle which says that an end cannot be multiply realized. On the contrary, ends invite realization in any of the ways the environment and agent are jointly equipped to provide, and this invitation is open-ended. To see a reason for an action is not to see a reason *against* alternatives; it to see the prospect of a parallel reason *for* alternatives.

The point is general but has its sharpest expression in cases in which we conceive the end of an action as itself an action. When an action is one's end, pursuing the end means seeking a *sufficient means* to it: an action or course of action by which the end-action is executed in full. I aim to make a dinner for my family with the ingredients on hand. I realize I can do so by making a shrimp and leek stir-fry with rice. Of course it does not follow from this realization that I could not make a different dinner instead. That I have the ingredients and knowhow to make the stir-fry does not entail I lack the ingredients and knowhow to, say, make a broccoli-cheddar soup and serve it with a baguette. In general, that doing *A* is a sufficient means of

[3] For further discussion, see Bridges (forthcoming).

doing *E* does not entail that doing *A* is the only sufficient means of doing *E*. How could it? The first thought directly concerns only the relationship between doing *M* and doing *E*, and in identifying *E*, invites the prospect of other means. The second thought rejects this prospect, and in so doing advances beyond the first thought. But it is only the first thought that is needed to make it intelligible that I do *M* in order to do *E*. So acting for a reason is perfectly consistent with seeing an equivalent reason for an incompatible alternative.

One may object that a conscientious agent will not be content merely with finding a sufficient means to an end but will seek rather to take the *best* sufficient means to that end—thus after all imposing a comparative condition upon their choice. In fact, conscientious agents do not always, or even usually do this. I seek to show my neighbor a token of appreciation for his past kindness and consider that making him a cake would be a sufficient way of doing so. It is simply false that I have made any kind of mistake at all in asking only what a sufficient way would be expressing of my appreciation, rather than what would the very best way to do so.

There is more to be said on this matter. But for present purposes the topic is moot, for even the judgment *there is no better means of achieving E than doing A* does not entail that doing *A* is the only act which may be so judged. Best-ness is an inherently comparative status, but it is not an inherently singularizing one. There is nothing to prevent a case in which the right metric for ranking alternative means to a given end is one which allows for ties in the top position.

Note further that the point I am making about reasons for action does not depend on assuming that all possible ends generate reasons for actions that are means to them. Such an "instrumentalist" view of our reasons for action is unsustainable. It's often perfectly sensible to ask what reason an agent has to pursue an end they propose to take up, and the answer may well be that there is none. If, on the other hand, we can identify a reason in answer to this question, this reason will itself concern something further that may be accomplished by

achievement of the end. And so the same question will arise one level up. As Aristotle noted, such an order of questions cannot go on forever. At some point we must arrive at a specification of an end that is worth pursuing, not for any further reason, but simply "in itself".

These considerations trace out further consequence of the observation that reasons for action lie in the ends the actions serve. They do not undermine the present point. If an end is worth pursuing, either in itself or for the sake of some further end, then it is apt to generate reasons for actions that are means to it, and it is apt to find such means in actions which are incompatible alternatives. This point is all that is needed for the current argument.

We can get the point more clearly into focus by seeing that it entails a fundamental logical difference between theoretical and practical reasons. A reason for a judgment is just as such a reason against any incompatible judgment. Why? Because a reason for a judgment is a consideration speaking to its truth, and what renders judgments incompatible is the thinker's consciousness that the truth of the one precludes the truth of the other. Reasons for action do not exhibit a parallel structure. A reason for an action speaks to its serviceability to a worthwhile end. But what renders it incompatible with another action it is not the second actions' *disservice* to the end. On the contrary, two actions which equally serve an end may be incompatible with each other, say, because taking the one would use up the time that would be required for taking the other. In the sense in which reasons for judgment are *bipolar*—such that favoring one judgment entails disfavoring opposing judgments—reasons for action are not.

Note that this situation persists even when we conceive the space of alternative actions as {doing *A*, refraining from doing *A*}. Of course these choices are incompatible. And it may seem natural to say that a reason for doing *A* must be a reason against refraining from doing *A*. But in fact there is no call to say this. I have reason to make my neighbor a cake. But I have the same reason to go out to the shop to get him a nice bottle of wine. Given that my free time is

finite, I can do the latter only if I refrain from doing the former. This means that the same end which generates a reason for me to make my neighbor a cake generates a reason for me to refrain from doing so, as the latter is a necessary part of something I have, given the end, reason to do. If the end were to generate a precisely opposing reason against refraining from making the cake, then it would nullify the reason it generates in favor. The upshot would be, absurdly, that no end could generate any positive reason for means to it in a case in which more than one means is available.

All of this is now to say: an explanation of an action that consists in giving the agent's reason for the action does not, in general, singularize. Such an explanation explains the action as taken in pursuit of an end. The only implication this has for other alternatives is the prospect that some might also admit of justification on the same ground.

We could achieve singularization, in explaining an action by giving the reason for it, if we could add, "And by the way, this was the only means to the end available." We will not in general be able to say this. But even when we are, we articulate a contingency whose obtaining is manifestly irrelevant to the character of the explanation.

The truth of Raz's *basic belief* entails that the rationalist's tempting conception of rational explanation—namely, that such an explanation does its work by showing the agent's conception of reasons to singularize upon their action—cannot be correct. It is one thing to have reason to think a view must be wrong, and another to understand exactly how it diverges from what is right. In this section I have tried to say how. The primary act of rational explanation is to reveal the action to be taken in service of an end. Revealing this does not require showing reasons to singularize upon the action; on the contrary, the form of the explanation is such as to suggest the opposite.

5. Explanatory completeness: goodness on the whole

The primary act of rational explanation—giving the reason for which the agent acts—does not require the agent's conception of reasons to singularize upon that action. If anything, it suggests the agent's conception will not so singularize. So I have argued. But our question was whether a *complete* rational explanation must appeal to a conception of reasons which singularizes. If a basis can be found for distinguishing between the primary act of rational explanation and the furnishing of a complete such explanation, we can regard our question as still open.

And indeed, there seems to be a natural candidate for a standard of rational-explanatory completeness which would require more than merely giving the reason for which the agent acts.

It is one thing to see a reason for doing something and another to be in a position to regard that action as justified unconditionally, justified full stop. A reason for doing something might be countered by a stronger reason against doing it. To ground an unconditional verdict of justification, you need to do more than register this or that reason in its favor. You need to take stock of a range of considerations bearing on the question of what you are, on this occasion, to do. And you need to find a way to bring those considerations together into a unified assessment of the action's merits.

At any rate, it may be said that this is what you will do if you are being a careful and conscientious practical reasoner. The closer your thinking approximates to this ideal, the firmer will be your justification for doing what you do. And given that it is the office of rational explanation to show how the action issues from the agent's understanding of it as justified, the more satisfying will be the rational explanation of your action.

I will say that an agent who gives due consideration to relevant reasons and finds an action

to pass muster on that basis judges the action *good on the whole*. Let us adopt as the standard of completeness for rational explanation that the explanation reveal the grounds on which the agent judges her action good on the whole. Does *this* conception of rational-explanatory completeness entail, per the rationalist's conception, that such completeness requires singularization? That question comes to this: does finding an action good on the whole entail finding it preferable to all alternatives?[4]

Let us return to our example. I seek, reasonably enough, to come up with a suitable token of appreciation for my neighbor's kindness. I consider that I may do so by making him a cake. Suppose I proceed to make the cake. Were you to ask my reason for this action, the correct answer would be that I sought to provide my neighbor with a suitable token of appreciation. But so far, this is just the primary act of rational explanation. We are looking for an explanation of my action which shows me to have had a ground for deeming that action good on the whole. What sort of additional considerations will need to enter into my conception of reasons for it to provide such a ground?

This is not a speculative question. We are ourselves conscientious agents, or approximate to so being, and we know the kinds of considerations that can lead us to deem an action good on the whole insofar as that status means anything to us at all. The present example, indeed, is not fictional. In the actual event, what I added to my thought of the action's suitability to my end was simply the further thought that, so far as I could see, my making the cake would not present a problem from the standpoint of my other ends. It *might* have done so. I might have promised my wife I wouldn't use up the flour. Or I might have had a wrist injury that would be aggravated by the beating and mixing. But no such thing was so that I could see, and that was

[4] I adopt the phrase "good on the whole" from Engstrom, who defends a positive answer to the question I here raise (2009, 102).

enough for me to find the action not merely supported by a reason, but, as I might have put it, justified full stop.

My ground exhibits a bipartite structure, which may be idealized as follows. First, there is an end that I seek in my action to realize. Call this my *active end*. Perhaps some prior practical reasoning set this end as my current agenda. But it is now for me a given. It prompts the initial question my reasoning is to address: *with what course of action may I achieve this end?* Second, there is a collection of ends which are mine despite their not being what I am on this occasion actively pursuing. Call these my *background ends*. They fund a second question, to be asked of any action given in answer to the initial question: *will this action undermine my contemporaneous or future pursuit of my background ends?*

Let us say that an action which adequately answers the initial question and yields a negative answer to the second question achieves *elementary rationality*. My reasoning on this occasion thus shows allegiance to the following principle: *elementary rationality* may suffice for *goodness on the whole*. It was enough for me to find my action as good on the whole that it was a sufficient means to my active end while posing no discernible hindrance to my background ends.

Suppose I am reasoning rightly in all this. Then it is clear that an adequate ground for a judgment of goodness on the whole need not singularize. That my making my neighbor a cake could achieve my active end without obstructing pursuit of my background ends obviously does not entail that it was the only possible action which met these criteria. Indeed, I knew this was not so: I knew that getting my neighbor a bottle of wine was another trouble-free alternative. The judgment *action A possesses elementary rationality* is simply not a singularizing form of assessment. There may be cases in which only one alternative can be correctly so assessed, but even when this is so the uniqueness will have nothing to do with the ground and truth of the assessment.

Of course, I described a simple and low-stakes scenario. On other occasions agents will need to chart a course in the face of limited resources, insufficient information, competing ends, and difficult tasks. Often enough it will be impossible to determine any alternative to possess elementary rationality, either because there is no clear sufficient means to one's active end, or because one is compelled to actively pursue disparate ends at once, or because there is no alternative that will not undermine one or another of one's background ends. Then one will need to find another way to deem the action good on the whole.

I cannot consider the range of complexities that may here arise. But there are two points to keep in mind. First, if elementary rationality *can* suffice for goodness on the whole, in those cases in which it is a feasible standard, that suffices to show that our candidate standard of rational-explanatory completeness does not involve singularization. Second, the crucial underlying fact, persisting in even more complex scenarios, is that reasons for and against action will always have to do with the bearing of those actions on the agent's pursuit of her ends. Accommodating a complex constellation of ends in demanding circumstances raises much harder questions than how to use an hour on a Sunday to come up with a trouble-free token of appreciation for my neighbor. But it does not change the fundamental logic of the reasons with which we are concerned.

It remains to ask whether I was indeed right in treating elementary rationality as sufficient for goodness on the whole. It is tempting to say: look in your heart, and you will know that it is so. The chief obstacle to this resolution, I think, is the distorting influence of the broadly decision-theoretic tradition on our practical-philosophical imaginations. I will mention two aspects of this influence.

First, formal decision theory treats acting itself as unproblematic. An action is modeled, say, as a choice among lotteries. This model foregrounds prior calculations of odds and values,

while obscuring the time, effort and attention that go into action as such. The very idea that the intentional agent is always solving discrete questions about what to choose is an idealization of what is in fact a dynamic and continuous mode of engaged thought.

One implication of these considerations is that there will be a sharp limit on how many disparate ends we may simultaneously keep in view on a single occasion for action. Sometimes we can see how to get several things done in one fell swoop. But as different ends typically require different means, and as taking a means itself requires focus, attention, and thought, the prospects for such multitasking will be limited. The paradigm of a single active end, with the rest of one's ends relegated to the background, reflects not a lack of ambition, but due diligence and care.

Second, the focus in the tradition on articulating and debating technical principles for measuring the aptness of a choice, or the quality of a means, makes it very difficult to avoid the impression that the true business of the human being is to meet a standard of faultless choosing. Of course, this thought will seem strained and ridiculous, as Hume might say, when we leave our desk and enter into the commerce of actual life. But it remains difficult to keep in view, in the context of current work on the subject, that the point of acting is not to be able to say that what we do is "justified", "maximal", or "rational". It is to achieve our actual, concrete ends. And so long as we are pursuing those concrete ends tolerably well, we are free to regard any and all formal metrics of assessment as an irrelevance.

6. The fate of the classical conception

Raz reaches the *classical conception of human agency* with the help of the *contrastive claim*, the claim that the complete explanation of why an agent does something must say why she does that thing rather than the other alternatives available to her. But the claim itself stands in need

of explanation. What do we mean when we speak of an explanation as complete?

If by "the complete explanation of why" (c.f., "the full story of why" in the quote from Raz in section 2), we mean "the explanation which includes everything anyone might conceivably wish to learn about why", then, as this description describes nothing, so too is the claim empty. No conclusion could be drawn from it, not a conception of human agency nor anything else.

It is, however, intelligible to hold that a complete *rational* explanation of an action must answer the contrastive question. For we can have a determinate idea of what it might mean to speak of completeness with regard to this explanatory form. Indeed, we found two candidates. We can say that a rational explanation achieves completeness in giving the agent's reason. Or can say that completeness is secured only once we have shown the agent's grounds for viewing her action as good on the whole.

Taken in either of these ways, the contrastive claim has sense. It is susceptible of truth or falsity. And, as it turns out, it is false. Neither giving an agent's reason, nor giving the ground of a judgment of goodness on the whole, requires showing the agent's conception of her reasons to singularize upon her action. The explanation for this lies in to the teleological, non-bipolar logic of reasons for action.

So much ground we have already traveled. But now we must consider the ground where Raz wishes to stand. For Raz accepts—it is a consequence of his *basic belief*—that a full accounting of the agent's conception of her reasons will typically fail to contain anything sufficient to answer the contrastive question. And yet he insists that when we give the agent's reasons, we contribute to an explanation which, when it is complete, does answer the contrastive question. This is the amalgam of premises that yields the *classical conception of human agency.*

Can this amalgam be made coherent? We have observed that a claim about what is required

for an explanation to be complete must be pegged to some determinate form or kind of explanation. Whether we can follow Raz will thus depend upon whether we can make good the implicit presumption underlying his whole construction: that there is a kind or form of explanation of an action which we elicit in asking, "Why does *S* do *A*?", and which meets two further conditions: (a) at least part of an explanation of this form or kind must involve citing *S*'s reasons for doing *A*, but (b) we will have a complete explanation of the form or kind only when, for each alternative to doing *A* of which the agent was aware, we have an explanation of why the agent does *A* rather that alternative. We have established that the envisioned explanatory form or kind is not rational explanation per se. And so the question arises: what is it?

Raz never answers this question, and I take it that is because no satisfactory answer can be given. We are now in a position to see why not. Once we recognize that the rational-explanatory form has its own standards of completeness, and that these standards do not involve answering the contrastive question, there can be no basis for holding that, in giving a rational explanation, we are contributing to an explanation which is designed to answer that question. A rational explanation answers the question "Why?" when that question is asked in the sense in which what is sought is an account of the agent's reasons, and it answers that question in no other sense. Anscombe said that "what distinguishes actions which are intentional from those which are not…is that they are the actions to which a certain sense of the question 'Why?' is given application", and that, "of course", "the sense is…that in which the answer, if positive, gives a reason for acting" (1957, p.9). We have here examined the logic of this question and its answers, and we have found that they have no necessary involvement with the contrastive question. This is sufficient to rule out not only the *rationalist conception*, but the *classical conception* as well.

In justifying the *basic belief*, Raz cites the prospect of incommensurable values and actions

much more frequently than he does the more mundane possibility of alternative means to the same end. He does not clearly connect the *basic belief* to the considerations of teleological logic I have raised here. And he is prone to speak as if the *basic belief* entailed a gap of some kind between one's reasons and the actions they are reasons for. For example:

> [O]n many occasions, I am inclined to say in normal circumstances, there would be more than one response supported by reasons, with none of them supported to a higher degree than any of the others. Needless to say, in all situations our actual response is more definite than that. That is it goes beyond what reasons require of us. And that is the key to the understanding of the limits of reasons, and of the richness of the sources of our response to the world, which include very much more than our rational capacities. (2011, p.5)

What reasons requires of us, he goes on to say, may be "only the alternative consisting in the disjunction of all alternatives that are not ruled out" (2011, p.6).

And again:

> In typical situations, reason does not determine what is to be done. Rather it sets a range of eligible options before agents, who choose among them as they feel inclined, who do what they want to do or what they feel like doing. (1999, p.65)

An illusion seems to have taken grip here. If I make a cake for my neighbor as a token of appreciation, then my reason for what I do reaches all the way to the particular action I take. That is to say, it is part of why I so act that I judge *making a cake for my neighbor is a means of showing him appreciation.* If I were instead to buy him a bottle of wine, I would not thereby act on *this* thought, but another. My "actual response" is exactly as definite as the thoughts in terms of which I find a reason for it. But because this definiteness resides in thought which is teleological, it does not require the ruling out of other "eligible options".

Even though we can find no room for Raz's imagined form of explanation, it is not hard to surmise what he wants out it. When reasons fail to singularize, he says, they cannot "determine" or "dictate" (1999, pp.48, 111) what we do. These verbs express compulsion,

necessitation. We may say, I think, that what Raz seeks is an explanation via a specification of a condition of the agent, prior to acting, which includes as a component their conceiving of their reasons as they do, and which further *guarantees* the performance of their action.

Many philosophers have supposed that an understanding of intentional action requires an explanation in terms of guaranteeing prior conditions or causes. In the 20th-century, the thought tended to be that this part could be filled by rational explanation as such. Hempel, to take a conspicuous example, held that explanations which "explain an action in terms of the agent's reasons and his rationality" conform to the "deductive-nomological template". That is, they cite pre-existing circumstances and lawful regularities such that "given the particular circumstances and the laws in question, the occurrence of the phenomenon was to be expected; and it is in this sense that the explanation enables us to understand why the phenomenon occurred (965, p.337). Such a view, as Raz would agree, cannot be squared with an appreciation of the *basic belief*. For by what law of "reason and rationality" could we predict which of a set of equally rational options, equally well-supported by reasons, the agent will choose?

But there is an older, we might say more "classical", tradition. Leibniz, while certain that "there is always a prevailing reason which prompts the will to its choice" (1985, §45), did not think that this "reason" must be a consideration which occurs to the agent, or indeed something she or any other human being will ever be in a position to know. Something will invariably ensure we do not occupy the state of perfect "equipoise" exemplified by Buridan's imaginary ass, but what that is might easily be a vast concatenation of causes "complex and indecipherable to ourselves", knowable only to God and the angels (1985, §49).

In denying that the agent's conceiving her reasons as she does can suffice to constitute the sought-after prior condition, Raz sides with Leibniz against Hempel. In insisting that a legible account of the condition will nonetheless be available to us and will include an accounting of

our reasons, Raz sides with Hempel against Leibniz. Occupying the ground between them, he is in the neighborhood of Schopenhauer. Schopenhauer denied that a specification of the agent's "motivations" alone could be enough to show their action to have been inevitable, but he held that if we were to couple such specifications to an account of the agent's empirically determinable and lawlike "character", we could in principle predict their actions (2005, chapter 3). Raz envisions similar explanatory materials, to much the same effect.

One benefit of getting clear on the teleological logic of practical reason is thus to see the unintelligibility of the middle position Raz and Schopenhauer seek to occupy.

Works Cited

Anscombe, G. E. M. (2000). *Intention*. Harvard University Press.

Bridges, Jason (forthcoming). "The unity of a practical inference". In Jim Conant and Dawa Ometto (eds.). *Practical Reason: Historical and Systematic Perspectives*, De Gruyter.

Davidson, Donald (2001). "Actions, reasons, and causes". In his *Essays on Actions and Events: Philosophical Essays*. Clarendon Press.

Engstrom, Stephen P. (2009). *The Form of Practical Knowledge: A Study of the Categorical Imperative*. Harvard University Press.

Leibniz, Gottfried (1985). *Theodicy*. Open Court Publishing.

Schopenhauer, Arthur (2005). *Essay on the Freedom of the Will*. Dover Publications.

Raz, Joseph (1999). *Engaging Reason: On the Theory of Value and Action*. Oxford University Press.

Raz, Joseph (2011). *From Normativity to Responsibility*. Oxford University Press.

Raz, Joseph & Heuer, Ulrike (ed.) (2022). *The Roots of Normativity*. Oxford University Press.

Scanlon, T. M. (1998). *What We Owe to Each Other*. Harvard University Press.

Wittgenstein, Ludwig (1973). *Philosophical Investigations*. Pearson.