KNOWING ONE'S OWN MOTIVATING REASONS

Seyyed Mohsen ESLAMI

ABSTRACT: Reasons are not the same. Normative reasons need to be distinguished from non-normative reasons. Then, due to some considerations, we have to draw a distinction between explanatory reasons and motivating reasons. In this paper, I focus on a rather implicit assumption in drawing the explanatory-motivating distinction. Motivating reasons are mostly characterized as those reasons that the agent takes to be normative. This may imply that the agent always knows the reasons their motivating reasons. This I call the infallibility or transparency assumption. This suggests that there is some sort of report condition on motivating reasons. In contrast, one may ask whether it is possible for the agent to be mistaken about what their motivating reasons are. I argue that we can distinguish motivating reasons and satisfy the motivations for the explanatory-motivating distinction without committing to this infallibility assumption. I briefly sketch a character-based approach to motivating reasons as an alternative. Next, I argue that, in addition to being important on its own, this account has implications for other debates. I illustrate this by mentioning cases such as recalcitrant actions as well as critically discussing one kind of counter-example presented against the guise of the good thesis.

KEYWORDS: theory of reasons, motivating reasons, normative reasons, self-knowledge, the guise of the good

1. Introduction

Reasons do a lot in our thinking and theorizing about the normative. However, soon it becomes clear that there is more than one kind or sense of a 'reason' we use and need. Speaking with a friend, we say "I know you had your reasons, but there is no reason for what you did!" Similarly, while there are good reasons for donating to charity, one may do so for the wrong reasons. These simple cases show that normative reasons may come apart from the agent's reasons, i.e., non-normative reasons (Parfit 1997, 99). If someone is in the building and the building is on fire, there is a normative reason for her to jump out of it, though she may lack a non-normative reason to leave. Alternatively, thinking that the building is on fire, she can jump outside, while that belief is false and there is no such normative reason for leaving the building (Parfit 2001, 17, 31).

The normative-non-normative distinction is important and raises many questions. For one, a classic argument has been that normative reasons need to pass

some motivational condition (Williams, 1995), which grounds an argument against external reasons. Relevantly, others have argued that normative reasons must be able to play other roles beyond merely favoring (Dancy 2000, 103). Yet the distinction holds and is now established. Plus, this has changed the scene of the contemporary ethical theory: despite the centrality of motivation in the metaethical debates of the twentieth century (Scanlon 2014, 1), now many authors take normative reasons at the heart of the theory of reasons (Scanlon 1998, 20). However, non-normative reasons matter and we cannot do only by normative reasons.

On the one hand, there are distinctions about, say, objective and subjective normative reasons (Parfit 2001, 25). On the other hand, we need to elaborate on the main classification, as the non-normative class of reasons may cover too much. We have to distinguish different kinds of non-normative reasons.

The non-normative category of reasons used to be introduced as 'explanatory or motivating'—as if 'explanatory' and 'motivating' are interchangeable. Now, however, the two are considered as different. It is now widely endorsed that there is a distinction between explanatory and motivating reasons (Alvarez 2010; Mantel 2018). Perhaps talking about what goes on in the neurons and muscles can explain the action, but we may be looking for something else, other kinds of explanations, or something else altogether. In such cases, as the common view goes, we appeal to motivating reasons.

Two points are due before reviewing the motivations for the explanatory-motivating distinction. First, the distinction is usually introduced in this way: motivating reasons are not *merely* explanatory reasons. This suggests (and is explicitly endorsed by many) that motivating reasons are a subset of explanatory reasons. One can argue that this is a substantial claim worth exploring, but I do not tackle it here. Second, it doesn't affect what follows whether what we get from explanatory reasons is indeed any explanation of *action* or not. That is, there are further questions to be asked about explanatory reasons as well.

This being said, let's review some considerations that motivate drawing the explanatory-motivating distinction. First, motivating reasons tell us something about the agent's deliberation process leading to the action. It has something about what operates at the agential level. Explanatory reasons may completely miss this point as they may include mere causal explanations. Therefore, explanatory reasons are not enough for there to be an action– there are always such explanations. But to judge whether there is an action we look for motivating reasons. Put differently, there may be 'reasons' for what I do that are not important to the question of whether I have done an intentional action. Anscombe's why-question is directed at motivating reasons ([1957] 2000, §16).

Second, motivating reasons are about the agent's perspective on the situation. When engaged with motivating reasons, we think about the agent's rationality. We expect agents to respond to reasons, but we don't expect them to respond to all normative reasons regardless of their perspective, as it would be impossible. Similarly, this is not about responding to explanatory (i.e., causal) reasons, as they do their job and do not need the agent's response. So, we can say that we care about motivating reasons, not normative or explanatory ones, in such questions. (Parfit 1997, 99; Sylvan 2011, 588).

Third, the explanatory-motivating matters normatively. In evaluating actions, we do not only look into the causal history of action (Muller 2004). Relevantly, consider the distinction between acting based on moral duty in opposition to merely acting in accordance with duty (Kant [1785] 2002, AK4: 398)—it matters to the worth of the action for which (motivating) reasons one has acted. Regret provides another example. Thinking about the fittingness of regret, it seems that the distinction is needed. If I have had a behavior for which I did not have any motivating reasons, the kind of regret that may be fitting (if any) would be different from when I have acted similarly but based on my motivating reasons. The thought goes back to Aristotle and his discussion of non-voluntary and involuntary (NE 1110b).

These and similar considerations push us to distinguish motivating reasons from explanatory reasons. Here is how:

> A motivating reason is a reason for which someone does something, a reason that, in the agent's eyes, counts in favour of her acting in a certain way. When an agent acts motivated by a reason, she acts 'in light of that reason' and the reason will be a premise in the practical reasoning, if any, that leads to the action. (Alvarez 2016, sec. 1)

One element in this picture I want to focus on is the knowledge of the agent about her own motivating reasons. The point may also be explained in terms of transparency. As noted, we err about normative reasons. It is also clear that we are mistaken about many explanatory reasons for our actions—it is no surprise that we do not know many causal facts. However, this question about motivating reasons is less discussed and its answer is less clear.

In Section 2, I describe the implicit assumption, i.e., infallibility or transparency about motivating reasons. I argue that this assumption is not necessary for the explanatory-motivating distinction. Plus, we have to avoid it if we can. In Section 3, I suggest an alternative approach to motivating reasons independent of the assumption and in line with the aforementioned motivations we have for the

distinction. In Section 4, I critically evaluate one kind of counter-example to the guise of the good thesis. In doing so, I explain how the counter-example relies on the standard understanding of motivating reasons and the suggested approach here blocks the challenge. This is partly to illustrate that the proposal about motivating reasons is relevant. Finally, before the conclusion section, some objections are briefly addressed in Section 5.

2. Characterizing Motivating Reasons

A positive answer seems natural to the question, "Are people infallible about their own motivating reasons?" I call this the infallibility assumption about motivating reasons. Alternatively, we can put this in terms of transparency. Importantly, this is not to say that writers on the subject defend the infallibility assumption. However the idea is mentioned in passing and is not explored on its own, and there are signs that the views favored are in line with the infallibility assumption. All this is mostly due to the way the motivating reasons are distinguished from explanatory reasons. Again, here is an example:

> A reason is called a 'motivating reason' because it is something that motivates an agent, that is, it is what he took to make his φ-ing right and hence to speak in favour of his φ-ing, and which played a role in his deciding to φ. (Alvarez 2010, 35)

Considering these remarks, Maria Alvarez's view doesn't allow generosity to be a motivating reason for the agent since "it is not what he took to make his action of giving money right" (Alvarez 2010, 36). The same thought is expressed focused on what the agent reports:

> This is not to say that the reason that motivates me to go about incognito is also that I believe that I am being followed. If asked what my reason was, I'd say that it was that I was being followed by MI5. (2010 49, n. 16)

So, this is a rough picture of the common view: not all reasons are normative; we act when there is no normative reason, or we miss them and act on other reasons. But not all non-normative reasons are the same. There is a distinction between explanatory and motivating. Motivating reasons are a subclass of explanatory reasons but have a specific relation with the agent. That relation is that of believing or being taken as or the like. If that relationship is what makes motivating reasons what they are, one cannot be mistaken about one's own motivating reasons. As if there is a report condition on motivating reasons—the agent needs to be able to report something as the ground for action for it to be a motivating reason.

It may help at this point to briefly consider *where* and *when* the agent may err. The main aspect of this picture I am opposing is something along the lines of a report condition on motivating reasons. The infallibility here refers to knowing about motivating reasons, not knowing the considerations themselves. What matters, thus, is transparency.

Next, about when the agent may go wrong. The issue is clearer when introspecting about why I have done something in the past. The agent obviously can misidentify the reason why she has done something. But I think the same is true when the agent is deliberating and acting. While deliberating, I may think that I am doing something out of kindness, but actually I am acting in fear; I may think that I am doing something because it makes my friend happy, but in fact I am responding to the fact that doing so will bother someone else. True, through some reflection I may be able to correct myself. I can think of the relevantly different situations and see if I would act the same way. Be that as it may, the possibility of correcting my mistakes doesn't suggest that I have never been mistaken. Plus, it does not imply that I can always correct myself.

So far for the illustration of the target assumption. Now, let's see if there are motivations for challenging the assumption and seeking an alternative approach to motivating reasons.

First, it would be preferable if the explanatory-motivating distinction is not committed to the infallibility assumption. Currently, it may seem that if one rejects the infallibility assumption, then the explanatory-motivating distinction vanishes: These are both operating inside the agent and the difference is that the agent can report some of them as motivating reasons. Therefore, even if nothing problematic arises from this picture, it still makes sense to develop alternatives to draw the distinction but not commit to the assumption. Well, infallibility is usually problematic, thus not a much innocent assumption as well.

Second, ordinarily we think people are sometimes mistaken about their own motivating reasons. I may sincerely claim that I did something for some reason, and others may disagree and suggest that the motivation has been something else. Note that this is not necessarily to change the talk from motivating reasons to explanatory reasons. Consider this: What others say would not provide an excuse by pointing to some explanatory reason. For example, I may blame a friend for a racist comment and say a few words about it. I think that I did that because the comment was offensive. Others may see that my motivating reason has been the speaker, not the comment. What we all are dealing here with are motivating reasons. Compare: Others could suggest that I am reacting because I am sensitive due the pain I feel from a recent surgery. This is to point to some explanation behind what I did and

may provide some excuse. This is not the only way others may know better about my reasons than myself. Sometimes they can know better my motivating reasons as well.

Third, and relevant to the previous point, there are cases of cognitive biases or racism (Srinivasan 2020). These are problematic and even blameworthy. At the same time, it is not that the racists can always observe and report their racist motivating reasons at work. Yet, as long as such actions (i.e., actions based on such states or attitudes) are blameworthy, there must be motivating reasons behind them. Compare: If racism is due to some poison, it would not be blameworthy in the same way.

Fourth, there may be further motivations for developing alternative approaches to motivating reasons depending on one's other theoretical commitments. For example, the alternative to be discussed here is a character-based approach to motivating reasons. If one is convinced to give a substantial role to the character in one's ethical theory, then (other things being equal) this alternative approach would be an attractive option.

With all this in mind, let's consider a less-committed approach to motivating reasons that allow fallibility.

3. Not Knowing One's Own Motivating Reasons

Here is a rough view: motivating reasons are those considerations that the agent is not alienated from, have some relation with the agent's character, and the action is done in light of them. This is merely meant to be suggestive enough about the direction of the alternative approach to motivating reasons.

First, we need actions to be that of the agent. There must be some specific relation between the agent and the action for it to be an intentional action. Not feeling alienated is a minimal way of saying this. Consider a radioman (Quinn 1993), horrified and confused about what is happening to him. Such an attitude toward action can raise worries about the relationship between the agent and the behavior. Yet, this is not to claim that one cannot simultaneously do something and find it alienating.

It is noteworthy that alienation here is to be understood rather objectively. In fact, according to the proposal, one *can* be psychologically to some degree alienated from one's actions. This has interesting implications as well. The usual discussions of action take it as well under control. However, the suggested approach to motivating reasons arguably allows *recalcitrant action*, where one may not consciously and at the moment want to act in such a way, though one cannot control oneself and the result is an action anyway. Compare this with the case of blame.

Usually, authors endorse a judgment requirement for belief. However, it may be that something such as seeming fits better with the theory, as it among others allows for recalcitrant blame and explains it (Portmore 2022, 55).

Second, we must explain *why* the agent doesn't find the action alienating. If the radioman has always been a radioman and remembers all his history, or if he has decided at a specific point to become one, he would not find things alienated. Similarly, if one desires something and the desire is somehow derived from one's previous desires, it would seem acceptable to her. It appears that the agent's character can help with this. The issue here is not what the agent thinks about the situation. The agent may be mistaken. The dependence in question is supposed to be at the agential level, yet not necessarily infallibly accessible or transparent to the agent.

For example, consider Bernard Williams on integrity, which is the ability to originate action (Smart and Williams 1973, 117). Consider a sad situation in which the agent, due to all sorts of oppressions, accepts a job, while she has for years resisted it and has always expressed his opposition to accepting that kind of job. Now, she says that she is acting as she wants and thinks the job is worth taking. However, those who know the person through years of conversations and interactions would judge that she is suppressing what she would authentically do and that the motivating reasons for the action are not the considerations she reports or thinks, but she is acting worrying about consequences. Her integrity is 'attacked', even if she doesn't see the matter this way. Integrity is lost as the action does not fit the agent's character—or 'projects and attitudes' in Williams' terms.

Finally, a consideration needs to be effective if it is a motivating reason. We have all sorts of considerations and don't act upon them all. The way a consideration becomes effective is important. It would not be a merely causal role. Rather, it may work in the agent's reasoning and similar processes.[1]

[1] The question of how a consideration must be effective, though important, is not specifically important for the suggested approach to motivating reasons. The same issue is raised with regard to the standard conception. More importantly, here is one way the issue may be complicated: First, consider the case for hypothetical blame (Shoemaker and Vargas 2021, 597)—where the agent may be blamed not for what she has done, but for what she would do. (Something similar may be arguably true about standing to blame.) Now, we can wonder about the idea of hypothetical motivating reasons. Then, the hypothetical blame fits well with the reasons behind the action, actual and hypothetical. Plus, hypothetical motivating reasons may help to make sense of Aristotle's distinction between non-voluntary and involuntary, mentioned in Section 1. The character-based approach discussed in this paper can easily make sense of the idea. Still, whether we want to allow such hypothetical reasons is another matter.

The crucial point is the difference in the kind of relation between the agent and the consideration. In the standard conception, the relation is an epistemic one. Here, the relation is somehow character-based. Of course, what makes something part of my character may be explained in part by the epistemic relations involved. Still, we can see how this conception of motivating reasons differs in that it is not committed to infallibility and allows the agent to make mistakes about her own motivating reasons. In all this, I am merely suggesting and motivating an approach to distinguishing motivating reasons from explanatory reasons.

Many elements mentioned in passing do need further development, to be sure. More is needed about 'fit,' 'character,' or 'alienation,' merely to mention a few. Yet, I hope this is suggestive enough to point to the kind of approach in mind.[2] Moreover, this approach has expected implications. For example, if motivating reasons are tied to the agent's character, to judge cases of action (whether there is action and what action there is) we cannot merely rely on what happens at the moment, but need to consider the past and even the future (cf. Tenenbaum 2020 on ratioanlity and agency).

Now, we have to examine this alternative approach and see if it is an option at all. That is, the explanatory-motivating distinction is drawn for some considerations. This alternative would be an option only if it can satisfy those considerations leading to the distinction in the first place.

The first consideration to draw the explanatory-motivating distinction is about deliberation. The thought is that all those things that may explain the action are not part of the agent's deliberation. In this regard, if we allow that deliberation is not limited to conscious thoughts, motivating reasons are not limited to considerations the agent consciously considers as reasons.

We can also add that the conception of deliberation that requires a conscious awareness of all the steps or premises would be a target of intellectualization objection. We do a lot that we are not conscious of at the time as in the case of habits. Although it may be the case that usually, or in many cases, we know our own motivating reasons, this is not necessarily the case.

A misconception is to be addressed here. If I act due to envy, it may be due to motivating reasons that I do not acknowledge are normative reasons and thus reject that they are motivating me to do anything. This is not to say that one can be motivated to do something for considerations completely inaccessible to her. Consider this in the case of anger: "Sometimes we lack the insight to know what our anger is really about; knowing that p is distinct from knowing that one is angry that

[2] Notably, such a picture is already familiar in some relevant literatures, such as Arpaly (2003) on virtues. It is not similarly developed in the context of reasons.

p" (Srinivasam 2017, 8, n. 32). The same holds for the proposal about motivating reasons and reasoning. Some thoughts may be part of my reasoning, though I may think that they are not. This is not to say that those elements are not in no way present to me at some level.

The second consideration for the explanatory-motivating distinction is that explanatory reasons cannot capture the agent's point of view, i.e., how she sees the world and finds it appropriate to act in it. A poisoned person can 'do' things because she is poisoned, though this doesn't reflect her perspective. Can this character-based approach satisfy this motivation? Yes. There is nothing wrong with this possibility of being mistaken about one's own motivating reasons, while the action based on those motivations says something about the agent's perspective, how she sees things. There is a sense in which this conception is more reflective of the agent's perspective since it is not focused on one aspect and the present time, but the character is more complicated and more encompassing. For one, it is sensitive to the agent's past and her ways of acting in other situations.

Finally, the third consideration is the normative implications of identifying motivating reasons. In this case, again, not only does this approach do the job, but it sounds to be advantageous. The thought is familiar and developed in the virtue ethical traditions.

Nothing here is to deny that there are also challenges to be met. For one, in this character-based approach to motivating reasons the line between the explanatory reasons and the motivating reasons seems to be less clear-cut. However, this would not be a problem in itself. Perhaps a too simple view on this issue would not fit how we normally consider the complexity of the issue.

One may object that this conception captures only explanatory reasons, not motivating ones. I think this issue may be turned into a verbal dispute at some point. I am fine with using an 'explanatory' label for the kind of reasons I have been talking about. But then we have to add that there is more than one kind of explanatory reasons, some normatively more important than others. Then, we would think that some explanatory reasons also provide excuses, ground blame, and the like.

In sum, having in mind the considerations for which the explanatory-motivating distinction is drawn, we have alternative approaches to motivating reasons available instead of the one committed to some form of infallibility assumption. Furthermore, not only alternatives are available, but there are motivations for preferring the alternative suggested here, as mentioned earlier in the chapter and will be explained in the next.

4. Motivating Reasons and the Guise of the Good Thesis

One may still worry if the character-based approach to motivating reasons has any philosophically relevant implications. To address this worry, as an example, I use the approach in response to a kind of counter-example against the guise of the good thesis.

In a recent work, Arto Laitinen (2018) presents some counter-examples to the guise of the good thesis. I argue that the counter-examples presented only work by assuming the standard conception of motivating reasons committed to the infallibility assumption. Put differently, if we have the character-based approach at hand, the counter-examples are easily blocked. This is an example of the relevance and perhaps importance of the infallibility assumption.

Here is one of the counter-examples put forward, which is structurally representative of the others:

> *Ollie in a burning building*: Ollie is in a burning building. He is in danger. He could flee. His family is also in the building and in danger. Ollie could try to save his family. Ollie thinks that "in this situation, the danger is a reason for me to flee, but it is overridden". Ollie flees however, and the reason that motivates him to flee is that he is in danger. (Laitinen 2018, 28)

We are supposed to read this as a case of intentional action, intelligible to us, although the action is not done for the desirability aspect, as if the action is done for a 'motivating consideration' rather than a 'motivating reason.' Here is the distinction Laitinen has in mind between these two:

> The stipulation is to distinguish a motivating reason and a (mere) motivating consideration. We can continue to call a 'motivating *reason*' the consideration that the agent took to be a good normative reason, and which the agent acted on. In happy cases, motivating reasons are identi'al with normative reasons, in unhappy cases the agent is mistaken about them. A "(mere) motivating *consideration*' is one that the agents did not even take to be a good normative reason, but acted on it anyway; typically knowing fully well one shouldn't. Importantly, this is not a mere causal-explanatory 'reason why,' and does not apply to snowmen melting. It requires the same psychological capacity of acting in light of a consideration, as acting for a motivating reason does, the only difference is that the agent does not think the consideration normatively favours the action. (Laitinen 2018, 25)

In a footnote, we are told about what goes on in the agent's mind: "The agent's attitude is rather that 'I did not have any reason to do it; that I did get some benefit was a silenced consideration. A virtuous agent would not have acted on such a consideration.'" (Laitinen 2018, 29, n. 24)

So, Laitinen aims to argue for this view: at times, intentional action is not done for reasons that can be normative but for motivating considerations. It seems that

this is the general view behind the argument: If the guise of the good thesis requires that intentional actions be done for motivating reasons (i.e., considerations that the agent takes to be normative), there are counter-examples. That is to say, contra the guise of the good thesis, there are cases in which there are no motivating reasons at work since, as the argument goes, there are merely motivating considerations, and the agent doesn't take them as normative reasons. For our purpose, here is the critical assumption: if the agent doesn't take a consideration to be normative, it cannot be a motivating reason.

Let's focus on this assumption. First, it seems that we need to learn more about the distinction between motivating reasons and motivating considerations, as Laitinen draws it. An important issue is that they seem to be able to do the same job—having a special bearing on the action or intentional action, more specifically. So, why are they two completely different categories, one of which can challenge a substantial thesis? Note that they are both different from merely explanatory reasons and have something to do with motivation. But motivating consideration is not yet a motivating reason, according to this view. Why? We can imagine that the thought behind this is that if the agent *sees*, considers, or *treats* a consideration as normative, she would be able to report so or think so—which may be described in terms of transparency. That is, it seems that the distinction holds only if the agent is in an infallible situation with regard to her own motivating reasons. However, if we put aside this infallibility assumption, there would be no distinction, and the counter-examples dependent on the distinction disappear.

There are two ways to deal with such counter-examples against the guise of the good thesis. First, one can argue that a merely motivating consideration, if it has nothing to do with the agent (i.e., the kind thoughts that invite a distinction between the explanatory and motivating reasons), then actions done on them are *not* intentional actions. Therefore, they cannot be counter-examples for the target thesis. Second, one can argue that there is no crucial difference between 'motivating reasons' and 'motivating considerations' but the way the agent thinks about them. This has nothing to do with the intentional action and how the agent treats some considerations. Therefore, it would not be a problem for the guise of the good thesis.

Let's see how this applies to the quoted case of Olli and the burning building. Olli leaves the building to save himself and leave his family alone. Ignoring the reason-consideration distinction, we can say that the action is done for a motivating reason. Olli thinks with himself that "the family must be a priority," but he does not act accordingly here. To see the point, imagine if we knew Olli for a longer time and have seen him acting similarly in similar situations. These may suggest that he, at

heart, does *not* see things that way. He reports such a priority for his family, but this does not fit him. He prioritizes saving himself over saving them.

However, perhaps this is not the case. Olli as we know him may be always (or usually) very caring about his family. He risks for their sake and does what he can to protect them. Meanwhile, he just left the building without paying attention to his family. In such a situation, perhaps this is not a good case for intentional acting. We can imagine that the situation has some aspects new to Olli, thus he is shocked and unable to perceive the situation properly and merely reacts rather reflexively. This doesn't fit what we know of him. In other words, although Olli is in some (non-agential) sense 'motivated' to save his life, this is not the motivation we hoped for, which reflects more on the agent.

Be that as it may, it is not our main concern here whether the guise of the good thesis holds and how one must understand the counter-examples against it. The point is to see how the conception of motivating reasons plays a central role in such debates and that a main element of that conception concerns the infallibility or transparency assumption.

5. Objections

Let me add further clarifications by addressing some objections.

First, to allow fallibility is not to allow being mistaken in all cases about one's own motivating reasons. The proposed approach allows the possibility of mistakes. Meanwhile, it may remain acceptable that in many cases, or in normal cases (whatever they are), the agent has at least a rough picture of her own motivating reasons. That there are unconscious parts to our minds or that transparency of the mind is mistaken doesn't suggest that we have no conscious level or that we don't know anything about ourselves in a privileged way.

Second, the label and presentation may suggest that this conception is committed to there being a 'character.' This may be problematic considering the recent attacks on the very idea of a character as something stable or deep-structured or else (Alfano 2013; Miller 2017). These debates about the character do not threaten the suggested approach to motivating reasons. First, it is not that the character challenge necessarily suggests that we should be eliminativists about character. Second, what is needed here is rather minimal. Roughly, the thought behind the character challenge is that there is no character with features such as stability. Still, the challenge is compatible with a sense in which agents act in certain ways and that those patterns (even if minimal and shaky) reflect something about them. The character challenge would be about the depth of such patterns and how we rely on

them for evaluation. We can incorporate what we learn from that debate into this alternative approach.

Furthermore, the questions from the character challenge are not limited to any specific conception of motivating reasons. For example, consider the practice of evaluating agents for their actions based on their motivating reasons. And assume that some form of infallibility. So, we may want to praise an agent based on the normative reasons for the action and that the agent takes those reasons to be normative and acts on them. Still, a question may be raised: "That the agent sees this consideration as a normative reason is very much dependent on the specific details of the situation. With minor changes, she could refrain from saving. Therefore, there is no ground for praising the agent." That is to say, the character challenge leads to questions and worries about luck and its role in our practices; there is not much especial to the character-based approach to motivating reasons.

Third, another worry would be about the relationship between motivating reasons and normative reasons. As already noted, a familiar thought in the literature is that there must be some relationship between motivating and normative reasons, as we wish that, at times, one can do good, acting both in accordance and for normative reasons. One way to hold the relationship is identity—motivating reasons must be such that they can be at the same time normative reasons, and normative reasons must be such that they can be motivating reasons. One reason must be able to play both roles (Dancy 2000, 103). Such a picture would find the suggested approach to motivating reasons problematic—the kind of things that can be motivating reasons (such as some cases of anger and envy) cannot simultaneously be the normative reasons—not in a straightforward, direct way, at least.

This is not necessarily a problem for the suggested approach as long as the identity view is challenged independently (Mantel 2018, Ch. 6). What is hard to deny is the need for some relationship between normative and motivating reasons. However, identity is not a good option here. And it seems that we can use the resources in our theory of virtue or moral worth to provide the required relation. Plus, it is not a necessary element of the character-based approach to motivating reasons to allow (say) envy to be a motivating reason, although it can.

6. Conclusion

Reasons are at the center of contemporary normative inquiries, and normative reasons attract the most attention. However, we need to know more about the non-normative reasons as well. In this paper, I raise and discuss a question about motivating reasons: can agents infallibly know what their motivating reasons are?

Here is how we get to the question: we need to distinguish normative reasons from non-normative reasons. Furthermore, there are considerations to distinguish motivating from explanatory reasons. One way to do so is to appeal to some epistemic relation between the agent and reasons - motivating reasons, we are told, are those explanatory reasons that the agent considers to be normative reasons in a way that can report it—thus, the report condition.

An alternative approach can still make sense of the explanatory-motivating distinction. This alternative can satisfy the motivations for drawing the distinction, yet it is not committed to the infallibility or transparency assumption. Plus, there are motivations for preferring this alternative—it has fewer theoretical commitments, fits well with our ordinary thinking, and leaves room for more possibilities such as recalcitrant action and blameworthiness of (say) implicit biases. Furthermore, the alternative approach has its own attractions and is relevant to some other debates. For example, it can block some counter-examples to the guise of the good thesis, where something similar to the infallibility assumption is assumed in the counter-examples.[3]

References

Alfano, M. 2013. *Character as Moral Fiction*. Cambridge University Press.

Alvarez, M. 2010. *Kinds of Reasons: An Essay in the Philosophy of Action*. Oxford University Press.

Alvarez, M. 2016. "Reasons for Action: Justification, Motivation, Explanation." In *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2017/entries/reasons-just-vs-expl/>.

Anscombe, G. E. [1957] 2000. *Intention*. Harvard University Press.

Aristotle. 2000. *Nicomachean Ethics.* Translated and edited by Roger Crisp. Cambridge University Press.

Arpaly, N. 2002. *Unprincipled Virtue: An Inquiry into Moral Agency*. Oxford University Press.

Dancy, J. 2000. *Practical reality*. Oxford University Press.

Kant, I., and Schneewind, J. B. [1785] 2002. *Groundwork for the Metaphysics of Morals*. Yale University Press.

[3] I am grateful to Artūrs Logins and Euan Metz who kindly provided me with comments and feedback on a previous version of this old piece.

Laitinen, A. 2018. "On Some Counter-Examples to the Guise of the Good-Thesis: Intelligibility without Desirability." *Ethical Theory and Moral Practice* 21: 21-36.

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

Miller, C. 2017. *The Character Gap: How Good Are We?*. Oxford University Press.

Müller, A. W. 2004. "Acting Well." *Royal Institute of Philosophy Supplements* 54: 15-46.

Parfit, D. 1997. "Reasons and Motivation." *Proceedings of the Aristotelian Society*, supplementary volumes 71: 99-130.

Parfit, D. 2001. "Rationality and Reasons." In *Exploring Practical Philosophy: From Action to Values*, edited by Dan Egonsson, et al., 17-39. Routledge.

Portmore, D. W. 2022. "A Comprehensive Account of Blame: Self-Blame, Non-Moral Blame, and Blame for the Non-Voluntary." In *Self-Blame and Moral Responsibility*, edited by A. Carlsson, 48-76. Cambridge University Press.

Quinn, W. 1993. *Morality and Action*. Cambridge University Press.

Scanlon, T. 2000. *What We Owe to Each Other*. Harvard University Press.

Scanlon, T. M. 2014. *Being Realistic about Reasons*. Oxford University Press.

Shoemaker, D., & Vargas, M. 2021. "Moral Torch Fishing: A Signaling Theory of Blame." *Noûs 55*(3): 581-602.

Smart, J. J. C., & Williams, B. 1973. *Utilitarianism: For and Against*. Cambridge University Press.

Srinivasan, A. 2018. "The Aptness of Anger." *Journal of Political Philosophy* 26(2): 123-144.

Srinivasan, A. 2020. "Radical Externalism." *Philosophical Review* 129(3): 395-431.

Sylvan, K. 2015. "What Apparent Reasons Appear to Be." *Philosophical Studies* 172(3): 587-606.

Tenenbaum, S. 2020. *Rational Powers in Action: Instrumental Rationality and Extended Agency*. Oxford University Press.

Williams, B. 1995. "Internal Reasons and the Obscurity of Blame." In *Making Sense of Humanity*, 35-45. Cambridge University Press.