To the participants in the Triangle Ethics Group:

I meant to give you a paper on an issue that lies more squarely within ethics. Alas, it is not yet in the appropriate state. I hope this piece is of interest to you. I would very much benefit from hearing your reactions, since I plan to convert it into an introductory chapter of a book provisionally titled Practical Rationality, Justification, and Values.

EVALUATIONS OF RATIONALITY

Sigrún Svavarsdóttir
Ohio State University

1. We, humans, are subject to evaluation in terms of rationality: we can be more or less rational or irrational. It is customary, at least among philosophers, to draw a distinction between theoretical and practical rationality. Theoretical rationality is displayed in regulating one’s beliefs and other cognitive states, while practical rationality is displayed in regulating one’s intentions, plans, and actions. I am mostly interested in practical rationality: how is it best to construe that type of evaluation? However, I will approach this question through more general reflections on assessments of rationality—indeed, paying initially more attention to theoretical than practical rationality.¹

2. It seems uncontroversial that only rational as opposed to arational beings are subject to evaluation in terms of rationality or irrationality. (In order to avoid confusion, I will use small capitals when using ‘rational’ to contrast with ‘arational’.) Now, a natural thought is that rationality is the excellence of a being qua rational being. This will be my guiding idea. Notice that this approach makes rationality out to be a virtue—a virtue of rational beings. Of course, no one might instantiate this virtue perfectly. People are rational to a greater or lesser extent. At some point they are not merely lacking in rationality, but are downright irrational.

---

¹ Since the Spindel Conference in October 2002, I have made some changes at the end of the paper. However, these changes were not made in response to Russ Shafer-Landau’s excellent comments and in no way undermine his comments. The footnotes have, for the most part, been added since the conference and this should be born in mind when reading Russ Shafer-Landau’s comments. Some of his minor points have been indirectly addressed in the footnotes, but I have not addressed his main critical point. That is something I will have to do in the context of paying the promissory note that I issue at the end of the paper. I would like to take this opportunity to thank Russ for fair and challenging comments. I also thank the audience for engaging questions, which have, indeed, inspired some of the footnotes. The paper was written while I was a Fellow in residence at the National Humanities Center.

excellence of a being qua RATIONAL being—consists in using one’s cognitive capacities as well as possible given one’s epistemic limitations.

3. Something seems still amiss. It is irrational to clutter one’s mind and waste one’s time by teasing out all the trivial and uninteresting implications of one’s beliefs or to weed out any inconsistency between them. It seems also irrational to go endlessly off tangents and explore at length insignificant side issues rather than continue a main line of inquiry, no matter how exquisitely one handles the tidbits. In these cases, one need not be making avoidable errors in representation or reasoning. One might, indeed, be using one’s cognitive capacities superbly, at least in the sense that the reasoning is sound and the representation both subtle and correct. So why is it a shortcoming of me qua RATIONAL being to spend my days teasing out trivial implications of my beliefs? A first stab at an answer focuses on my limitations: I have a finite mind and time. Spending them in this way diminishes my opportunity to grapple with matters more important than trivial implications of my beliefs. This raises the question whether the excellence of a being qua RATIONAL being is in part constituted by not letting an inquiry of less important matters detract from one’s inquiry of more important matters. If so, we could even save the suggestion that rationality—the excellence of a being qua RATIONAL being—consists in using one’s cognitive capacities as well as possible given one’s limitations. We would only have to add that it is possible to use one’s cognitive capacities poorly, not only by falling into avoidable error, but also by using them on one subject matter at the cost of another, more important, subject matter.

I do not dispute that some subject matters are more important than others and that we may be using our cognitive capacities poorly if we concentrate on the less important at the cost of the more important subject matter. But it does not seem like this in itself amounts to irrationality. It seems more important to find the cure for cancer than design household equipments. Moreover, it seems a shame if a genius decides to use his formidable mental powers on designing household equipments rather than on finding the cure for cancer. Perhaps, it is also a moral failing. Certainly, it is not the best use to which the genius can put his cognitive capacities. But this hardly suffices for making the genius irrational. Imagine encountering the genius in his lab excitedly explaining to you the newest intricate design of the perfect bottle opener. You think of what he could have accomplished as a medical researcher. Alas, he was never able to develop interest in anything but complex design of silly household equipments. Do you accuse him of irrationality for designing bottle openers rather than developing

new drugs? That does not seem like the appropriate kind of criticism in this case.

Now imagine that the genius starts obsessively to work out all the implications of his beliefs. He gets to the point of tediously listing all the disjunctions that follow from his belief that snow is white, an endless task indeed. There are half-finished designs waiting on his drawing table, finished designs that need to be forwarded to the production department, unpaid bills, missed appointments. He loses touch with friends, loses sleep and most meals. Calling him irrational is at best an understatement. This is a descent into madness. But consider a slightly milder case of this disorder: the genius has a tendency to break off his work on his intricate designs by starting to work out the implications of his various beliefs, even if they do not bear on what he is doing and even if they are of no interest to him beyond being implications of what he already believes. These obsessive reflections seriously hold up his projects from time to time, although they never take over completely. Alternatively, imagine that the genius has a tendency to go off tangents. Frequently, he explores thoroughly possible addendum to the equipments he is designing, though it should be clear to him from the start that they have no good use or do not fit with the overall design of the object. These become little side projects that are abandoned as fruitless after holding up the main project for hours on end, even when much rides on finishing the design quickly. These are, I submit, clear cases of irrationality. Being a genius, the designer’s reasoning and representations are beyond reproach even when he goes off tangents or obsessively works out all the implications of his beliefs. So, what is his shortcoming qua RATIONAL being?

4. The key to understanding this form of irrationality is to appreciate that evaluations of rationality apply only if the subject is using his cognitive capacities purposefully. Imagine a machine that at least closely simulates our powers of representation and reasoning. Like us, it can fall into error and we may even imagine that it can in some cases detect and correct its mistakes much like we do. Say that this machine replaces field-workers who gather and analyze data about various features of the physical environment in remote regions of this world. Now, compare this machine to a field worker whom the machine has replaced. The representation and reasoning they use may be very similar. We may even assume that the machine and the field worker can be misled in a similar way by an input from the environment, but they both have the ability to catch the mistake and correct it. When the machine fails to correct the mistake, it is malfunctioning, whereas the field worker is being sloppy, dense, or irrational. Also, imagine that some input triggers the machine to spend inordinate time on recording data of little interest given the main line of research. It goes off a tangent so to speak. If a similar thing happens to the field worker, we deem her irrational. But such an evaluation does not apply to the

---

3 Of course, the notion of an agent’s epistemic limitations needs to be clarified. I leave that for the future.

4 This observation is borrowed from Gilbert Harman. See his Reasoning, Meaning and Mind, 18-19.
5 Am I imagining something that is at least metaphysically impossible? I just might. Perhaps an adequate theory of our representational and reasoning capacities will reveal that nothing like them could exist outside of a nexus of propositional attitudes. However, that would strengthen rather than weaken my ultimate point.

6 What if an agent is relying on the cognitive capacities of a machine (or another agent) in a purposeful manner? Is he being irrational when the machine (or the other agent) does not exercise its cognitive capacities as well as it could? I think that the agent is being irrational only if he should be able to foresee, given his cognitive capacities, that it is not worth the risk of going awry to rely on the machine (or the other agent) in the task he is undertaking.

7 I am assuming that making an assumption involves being or entering into a cognitive state and, hence, that one may display theoretical rationality or irrationality in making an assumption.
sometimes those you least suspect can turn abusive in one way or another. You also believe that the ratio of wrongful accusations to rightful accusations is low in sexual harassment cases. However, the information you get about the interactions between your friend and the alleged victim is highly ambiguous. You have some grounds for distrusting the alleged victim, but at the same time you worry about the effects of self-deception on your trusted friend. You really want to know what happened and have tried your best to figure it out. But you are completely unsure about what to think. The case is settled, by the relevant authorities, in favor of your friend. The evidence is deemed insufficient for establishing that sexual harassment has taken place. You still feel completely unsure about what actually happened. Nevertheless, you decide to set this behind you and relate to your friend on the assumption that he is a victim of a wrongful accusation. After all, you firmly believe that people should be presumed innocent until proven guilty beyond reasonable doubt. Moreover, this friendship is important to you and you care about your friend. It proves psychologically easier than you expected to relate to your friend on this assumption. It is, indeed, as if you believe him innocent. However, on the few occasion that you revisit the matter, you have to admit to yourself that you are completely unsure about what happened.

Are you rational or irrational in adopting your working assumption? Whatever the answer to that question may be, it seems clear that the rationality at stake is practical rationality so long as you are adopting the assumption, not in order to gather more evidence to settle the case, but rather in order to maintain a valued friendship. In so far as your theoretical rationality is at stake, it has to do with whether your uncertainty is really warranted given the information you have about the case. Given your relation to one of the parties to the dispute, you might be subject to bias in your assessment of the evidence. Perhaps you are guilty of wishful thinking: Hoping for your friend’s innocence, you are not using your cognitive capacities as well as you could when assessing the reliability of his testimony. Or, perhaps you are perversely biased against your friend as a result of overcompensating for an initial bias in his favor. Theoretical rationality has much to do with neutralizing the impact on inquiry of concerns that are likely to undermine one’s pursuit of truth and understanding, provided that one is in a position both to appreciate the danger and combat it. This is because assessments of theoretical rationality take place against the background of only such cognitive goals as figuring out what is the case or how to understand something. In contrast, there is no such restriction on the types of goals against which an agent’s practical rationality is assessed. Practical rationality is assessed against the background of all the agent’s goals, whatever they are. This is, I submit, the crucial difference between assessments of theoretical and practical rationality. This calls for a sharpening of my guiding idea that rationality is the excellence of a being qua rational being. Theoretical rationality is the excellence of an agent qua rational being engaged in inquiry, while practical rationality is the excellence of an agent qua rational being, no matter what he is up to.

6. You must suspect by now that I am going to propose that evaluations of rationality, both theoretical and practical, are relativized to the agents’ ends. Whereas theoretical rationality is relativized solely to the agent’s cognitive goals, practical rationality is relativized to all his ends, taken as a whole. This suggestion is, of course, neo-Humean in spirit. However, I am not proposing that rationality is a matter of maximally advancing one’s ends like neo-Humeans are wont to characterize at least practical rationality. It need not be irrational to act such as to undermine overall one’s ends. For the agent may not be in a position to anticipate this outcome. Such an agent is unfortunate rather than irrational. He is not falling short qua rational being. Evaluations of rationality need to be relativized not only to the agent’s ends but also to his epistemic position. Indeed, they are best seen as evaluations of how well the agent has used his cognitive capacities, given his epistemic limitations, in the service of his ends. Practical rationality consists in an excellent use of one’s cognitive capacities, given one’s epistemic limitations, in the service of one’s ends. This is how I propose to develop, for now, my guiding idea that rationality is the excellence of an agent qua rational being or, more specifically, that practical rationality is the excellence of an agent qua rational being, no matter what he is up to. The idea that theoretical rationality is the excellence of an agent qua rational being engaged in an inquiry can, in a similar vein, be developed into a conception of theoretical rationality according to which it consists in an excellent use of one’s cognitive capacities, given one’s epistemic limitations, in the service of one’s cognitive goals.

---

8 If your evidence indicates that some of your concerns or biases enable you to understand your subject matter better or figure out things that you would have otherwise overlooked, it is not theoretically irrational of you to leave your inquiry open to influence from these concerns or biases. What if you do not have any evidence for this, but this is still the case? In that case, your theoretical rationality is, I submit, undermined, even if you might end up knowing and understanding more than your more theoretically rational colleague. You are being lucky rather than rational.
The observation that evaluations of rationality apply only to those who use their cognitive capacities purposefully does not, on its own, imply a relativization of such evaluations to the purposes of the agent evaluated. Take, for example, the recognition view of rationality, sometimes associated with Aristotle\textsuperscript{11} or with Plato.\textsuperscript{12} This view construes the excellence of an agent qua rational being as a matter both of using one’s cognitive capacities well in discovering reasons for belief and action and of basing one’s beliefs and actions on these reasons. This view makes perfect sense of why evaluations of rationality apply only to agents engaged in purposeful activity. However, it does not relativize these assessments to the purposes of the agent. For an essential ingredient of the recognition view is that reasons for action and belief are not fixed by the agent’s purposes. Rather, whether a consideration is a reason for belief or action is determined by whether and how it bears on the truth of the belief or the value of the action (or its outcome), when both truth and value are objective in the sense of not being fixed by anyone’s mental stance on the matter at hand.

It is the commitment to objective values that has deterred many from embracing the recognition view of practical rationality. Metaphysical and epistemological worries are frequently cited as motivations for rejecting the view.\textsuperscript{13} These are not my motivations. By relativizing assessments of rationality to the purposes of the agent evaluated, I am not attempting a naturalistic reduction of rationality. In my characterization of rationality so far, I have relied on the notion of excellent use of one’s cognitive capacities in the service of one thing or another, and I have not made any promise that this notion can be unpacked in naturalistic terms. Let me add now that such excellent use requires two things: first, that the agent represent and reason as well as feasible, given his epistemic limitations, when figuring out whatever is relevant to the purpose for which he reasons; and, second, that the agent put his conclusions to an appropriate use in fulfilling his purpose.\textsuperscript{14} Now, it is in line with my conception of rationality to understand appropriate use in fulfilling one’s purpose as use conducive, as far as one can tell, to fulfilling one’s purpose. Still, this does not amount to a naturalistic reduction of rationality, given the first condition on excellent use. Moreover, I suspect that the second condition, when more carefully formulated, will appeal to some idealized conditions that cannot be easily captured in naturalistic terms.\textsuperscript{15} In any case, my proposal is not motivated by reductive ambitions, even if it might take us a few steps towards an understanding of rationality that can be incorporated into a naturalistic conception of ourselves and our surroundings.

It is not that I am a fan of extravagant ontologies or a fiend of naturalism. Rather, the relativization issue is not a metaphysical issue and does not have any obvious metaphysical implications. This can be best appreciated by seeing that even emotivists about evaluative discourse can acknowledge both relativized and non-relativized evaluations. According to them, evaluations that are relativized to the ends of the evaluated agent express attitudes, on part of the evaluating agent, conditioned on the assumption that the evaluated agent has certain ends, whereas evaluations that are not thus relativized express attitudes towards the evaluated agent that are in no way conditioned by what sort of ends he has. I am not advocating the emotivist account of what is going on when we make evaluations that are or are not relativized to the ends of the evaluated agent. What I am highlighting is that this contrast between relativized and non-relativized evaluations pertains, first and foremost, to the nature of the mental act of evaluating or the concepts employed in that act. It is a mistake to run that issue together with the metaphysical issue concerning whether the value of an object obtains independently of our evaluation of the object.\textsuperscript{16} Of course, once we start to ask about the nature of the act of evaluating, metaphysical issues about the mind-independence of values may arise. I would have to be convinced that there is no mental faculty or more abstract normative authority—called reason or rationality—that dictates how to think, feel, or behave. I conceive of practical rationality as a virtue. In failing to use our cognitive capacities in the service of our ends, we fail to display this virtue rather than violate a command or a law.


\textsuperscript{13} G. Cullity and B. Gaut, Ethics and Practical Reason, 15-18; O’Neill, Bounds of Justice, 13-14.

\textsuperscript{14} These two conditions should call to mind the two types of shortcomings for which there must be room for evaluations of rationality to apply: first, avoidable inaccuracy in representation and reasoning; second, inappropriate use given the purpose for which the cognitive capacities are enlisted (see p. \textsuperscript{9} above). Notice that given the second condition, rationality requires something more than using one’s cognitive capacities well in figuring out how to fulfill one’s purposes. It also requires that one put this conclusion to use in fulfilling one’s purposes. Thus, “true irrationality” in Christine Korsgaard’s sense is possible on my neo-Humean conception of practical rationality. See her “Skepticism about Practical Reason” reprinted in her Creating the Kingdom of Ends (Cambridge: Cambridge University Press, 1996), 318-325. I believe this opens up a way to understand the irrationality of the incontinent person as well as of the depressive, although I will not spell out the account here. When I speak of requirements of rationality, I am simply speaking of necessary conditions for counting as rational. I reject imperative and legal models for understanding such talk.

\textsuperscript{15} Unlike many contemporary neo-Humeans, I think that the relativization has to be to the agent’s actual ends or ends that bear a close relation to the agent’s actual ends rather than to highly idealized ends (e.g., ends that the agent would have under full information). However, some idealization might be warranted. Also, the epistemic condition in the requirement has undoubtedly to involve some idealization.


There is no mental faculty or more abstract normative authority—called reason or rationality—that dictates how to think, feel, or behave. I conceive of practical rationality as a virtue. In failing to use our cognitive capacities in the service of our ends, we fail to display this virtue rather than violate a command or a law.

Unlike many contemporary neo-Humeans, I think that the relativization has to be to the agent’s actual ends or ends that bear a close relation to the agent’s actual ends rather than to highly idealized ends (e.g., ends that the agent would have under full information). However, some idealization might be warranted. Also, the epistemic condition in the requirement has undoubtedly to involve some idealization.

that these issues are more pressing or less tractable if the evaluations are non-relativized rather than relativized.

I reject the recognitional view because it does not relativize evaluations of rationality to the purposes of the agent evaluated irrespective of whether this will lead us into a metaphysical quagmire. Earlier I considered and rejected the proposal that the excellence of a being qua rational being is in part constituted by not letting an inquiry of less important matters detract from one’s inquiry of more important matters. This was not driven by metaphysical worries about objective measures of importance or value, but rather by an intuition about when charges of irrationality are appropriate. Moreover, we were able to understand the theoretical irrationality of teasing out the trivial consequences of one’s beliefs and going off tangents only against the background of the agent’s cognitive goals. Typically, our cognitive goals are relatively specific: our goal is to discover or understand some specific phenomenon rather than to discover any old truth or understand whatever there is to be understood. It is only because of this that it is irrational to waste one’s energies on discovering or understanding certain facts at the expense of discovering and understanding some other facts. Thus, my motivations for relativizing evaluations of theoretical rationality to the purposes of the evaluated agent has rested on intuitions about when charges of irrationality are appropriate as well as assessment of how these intuitions are best explained. I have been hoping that this makes my readers more receptive to a controversial idea that I happen to find intuitively compelling, namely that assessments of practical rationality are anchored in the purposes of the agent evaluated.

Motivating my proposal by appealing to intuitions may seem dubious. Even intuitions to which I have already appealed may be turned against me. Imagine an agent whose sole aim in life is to have a deductively closed system of beliefs. My proposal has the consequent that it is rational for such an agent to work out all the implications of his beliefs. This seems to go against the intuition about irrationality on which I have been relying, namely the intuition that it is irrational to spend one’s days teasing out all the implications of one’s beliefs. Nevertheless, I accept this consequence of my view. I am not sure that the intuition, on which I have been relying, is triggered by the purported counterexample when properly understood and, in any case, I would not want to rely on an intuition triggered by this imagined scenario. It is difficult to get an imaginative fix on an agent whose only aim is to have a deductively closed system of beliefs. You have to imagine an agent who does not aim to understand anything but the implications of his current beliefs, and who does not even aim to stay alive and well. Otherwise it is doubtful that we have an agent whose ends are overall well served by working out all the implications of his beliefs. Well, perhaps we can imagine such a guy. Conjecture up the image of an independently rich guy sitting in his study teasing out all the implications of his beliefs as a servant brings him sustenance. He passes his days in euphoria, going through one trivial implication after another until sleep overtakes him. Do we have an irrational agent or simply a nut case? I would not let my understanding of rationality hinge on the intuition triggered by this case or similarly farfetched cases.

I would be satisfied if my conception of rationality matched and made sense of intuitions about relatively realistic cases. However, the prospects for this are grim, especially when it comes to intuitions about practical rationality. Our intuitions conflict. For example, many find it counterintuitive to think that an agent with awfully silly endeavors be rational. Imagine a person who sets out to discover how many Johns live in Durham, North Carolina. Admittedly, I would suspect him of irrationality, both theoretical and practical. For it is difficult to believe that his life is so impoverished that spending his intellectual resources in this way does not detract from cognitive goals and other endeavors closer to his heart. But if I am wrong on this score, I don’t see that his shortcoming is that of irrationality, especially if he pursues the matter in the most ingenuous way. The man is just silly and shallow—an awfully uninteresting person, indeed. My intuition here is strong and plays a crucial role in motivating me to advance a neo-Humean conception of rationality. But apparently the intuition is not widely enough shared to provide solid grounds for defending my conception of rationality. What is needed, I believe, are theoretical considerations in favor of understanding evaluations of rationality in the way I have proposed. I will not have time to develop the theoretical considerations here, but let me explain how I conceive of that project.

There are two main questions that need to be addressed: First, are there genuine evaluations that pertain to how well people use, given their epistemic position, their cognitive capacities in the service of their ends? Second, assuming that these are genuine evaluations, why is rationality best understood as the virtue epitomized in this type of evaluation? Both questions need some clarification. It is far from obvious what would constitute adequate answers to them. The first question, I submit, requires that we be given a good sense of why it matters that people use their cognitive capacities well in the service of their ends—matters in

---

17 Perhaps the most difficult cases for me involve self-destructive agents. Consider a severely depressed agent whose only aim is to find a way to end his life, even if he has been informed that with proper medical treatment he will get out of the depressive state and enjoy life again. Is he rational when he uses his cognitive capacities, in the most ingenuous way, to fool the person who stands a suicide guard over him and to find the means to end his own life? If he fails to do this, is he failing to display the virtue of rationality? My intuition is shaky about this case and I am weary of resting my conception of rationality on the intuition that it triggers (in me or others). I would like to come back to this case, once I have provided a deeper and more theoretical understanding of why the virtue of rationality is best regarded as an excellence in the use of one’s cognitive capacities in the service of one’s ends—the sort of understanding for which I call below. –Thanks are due to Walter Sinnott-Armstrong for pressing the case of the suicidal depressive on me.
such a way that this can be considered a virtue. The second question, I submit, requires that we be given a good sense of why an excellence in the use of one’s cognitive capacities in the service of one’s ends is the virtue of agents qua RATIONAL beings. Moreover, it must be elucidated why the verdict that someone uses his cognitive capacities well in this way has a justificatory force, while the opposite verdict has a critical force—the sort of justificatory and critical force that verdicts of rationality and irrationality are commonly understood as having.

It is far from trivial to meet these requirements. This can be best appreciated by imagining an evil or a silly person who uses his cognitive capacities superbly in advancing his evil or silly ends. Can we really think of him as justified in his evil or foolish ways just because they result from an excellent use of his cognitive capacities in the service of his ends? Don’t his ends have to be worthy as well? Can we really think of the excellence of this person’s use of his cognitive capacities as mattering in the way that virtues matter? Can we think of it as the virtue of agents qua RATIONAL beings? Does that virtue not also consist in having worthy ends—or least ends which are such that there is a reliable connection between using one’s cognitive capacities well in the service of them and ending up having true beliefs and doing valuable things. Isn’t that required for verdicts of rationality to have the justificatory or critical force that they are commonly understood as having?

These challenges assume that the justificatory force of verdicts of rationality and the status of rationality as a virtue cannot be elucidated unless rationality is shown to serve truth and value. This is an assumption that I want to challenge, even if I believe that as a matter of fact rationality tends to serve truth and value. Understanding why rationality is a virtue and, in particular, why verdicts of rationality have justificatory force requires that we probe deeper into the nature of the type of agents—RATIONAL agents—that are subject to evaluation of rationality. What sort of agents can be justified in or criticized for forming a belief or an intention to act? It seems that the ability to use one’s cognitive capacities purposeful does not suffice. What is required is that one has the ability to reflectively assess and revise one’s beliefs and intentions. Such reflections can only take place against the background of other beliefs and concerns. There is no way of stepping outside of one’s mental stance to figure out what to believe or do. However, one may use one’s cognitive capacities better or worse in the course of reassessing and regulating one’s beliefs and intentions. Here evaluative notions such as responsibility, conscientiousness, and integrity find their natural niche and, I want to argue, the notion that an agent can be more or less justified in forming a belief or an intention. This is, I contend, the key for understanding why evaluations of rationality have a credible justificatory or critical force, even if they are relativized to the agent’s mental states in such a way that there is no guarantee that rationality will serve truth and value. It will also cast light on why rationality, as understood here, matters in such a way that full justice can be done to my guiding idea: namely, that rationality is the virtue of agents qua RATIONAL beings. At the same time, it will require that my conception of practical rationality be further developed. An agent can display practical rationality not only in acting one way rather than another, but also in forming and revising intentions, plans, and other commitments. Moreover, evaluations of practical rationality need to take account not only of the agent’s goals, but also of his various attitudes regardless of whether they are reflected in his current aims. Alas, I have issued a lot of promissory notes. Let me stop before I sink further into debt.

---