

Papers from the 2003 meeting of the Canadian Association for Reductionist Philosophy
“Rationality” by Jan Narveson, University of Waterloo

Rationality

Why the familiar Aristotle/Hume account will maybe do, and anyhow can hardly be wrong...

by

Jan Narveson

University of Waterloo

copyright © 2004 by Jan Narveson. All rights reserved.

Abstract of “Rationality” by Jan Narveson
--

This paper defends the standard account of instrumental rationality derived from the work of Aristotle and Hume.
--

**Papers from the 2003 meeting of the Canadian Association for Reductionist Philosophy
“Rationality” by Jan Narveson, University of Waterloo**

Appeals to rational action are, we thought, supposed to *help*. We suppose that decision theory will tell us something, possibly something useful, about some reasonably important matters. And we suppose that these methods will help because they are solid. In particular, they are, we suppose, independent of particular sectarian theses about morals and politics.

But, notoriously, some deny this. Peter Vallentyne complains that if we try to base morals on practical reason, then we have merely “traded in intuitions about morals for intuitions about what’s rational.”¹ It has always seemed to me that he is wrong about this. Indeed, it still does. The present enterprise is hoped to add grist to at least one of our mills, preferably mine.

What is practical rationality, then?

To say that someone or something is rational is to say that it can reason, and usually that it typically *does* so as well. Either it does or it doesn’t, is the idea - either its action is rational or it isn’t. But sometimes when we say this sort of thing, we mean that it is *rather* rational, or *quite* rational, etc. - implying that rationality is a matter of degree (and, perhaps, a virtue). An important question is what these two uses have to do with each other. Is it quite a different sense of the term altogether when we *classify* X as a “rational being” or not, than when we estimate the *degree* of rationality of some scheme or action or decision? I’m going to suppose (and suggest) that to be rational in the first sense is to be *amenable to appraisals* along the lines employed in the second sense. So to raise the question of whether someone was acting “rationally” we appropriately answer it by showing that it employed reasoning in the process. If so, then the question arises how well it reasoned, which things if any went how wrong, and the like.

Of course, we are talking here of *practical* reason. One interesting question is whether there is a significant difference between it and something else, “theoretical” reason. In principle, it would seem so. Could nonpractical beings reason? There seems to be a fairly clear sense in which computers do something that might be called “reasoning.” They “do” logic, don’t they? If we say that they do, however, we will soon need to invoke the distinction between doing logic in the sense of thinking logically, especially *because we want to* and doing it because “we” are a machine programmed to do this sort of thing. The former is what brings up the subject of practical reason - not the latter.

Further, it seems possible for a practical being to do logic (and suchlike) “for its own sake” - just for fun, so to say. Indeed, not only is it possible, but it happens. (Some of my best friends seem to do this....!) There is nothing non-practical about doing things just for fun or just for interest, and any accounting of rationality that confined it to “instrumental” reasoning in any sense that precludes this would clearly be too narrow.

In this paper, I want to address two questions:

(1) What is the *essential form* of practical reason? What has to be present before we can speak of rationality, and such that if it is present, then we can and do so speak?

(2) Is there possibly anything interesting about (1)? In particular, what is the relation (if any) between practical reason and the more general sort, which we’ll call (following Aristotle?) ‘theoretical’ reason? For that matter, can we hope to extract something of ethical or political interest from it?

My familiar answer to (1) comes from Aristotle and Hume and people like that. For practical reason, what we need are:

1. Some “*values*”: desires, wants, that sort of thing. The modern term of choice for this is ‘preference’: so, the rational agent is concerned to achieve what he *most prefers*. (This, as is familiar,

¹ Said in private discussion. Sorry, no better reference.

Papers from the 2003 meeting of the Canadian Association for Reductionist Philosophy “Rationality” by Jan Narveson, University of Waterloo

has plenty of problems. So one major question is whether any plausible doctrine of this type survives that can either finesse, side-track, or generalize around those problems.)

2. Some *information*, specifically concerning (a) one’s repertoire of known unilaterally available options, and (b) their likely impact on the furtherance of the values in (1).

And it has to be the case that those two components do stimulate action. The practically rational being, wanting x and knowing that to get x it must do y, *does* y and does it *for that reason*. The object of this exercise is to find, and then do, the best result as dictated by (1) given (2).

Note that this characterization uses the term ‘best.’ Since our idea here is to invoke rationality as the basis for various normative conclusions, the ‘bestness’ in question had better not be the very thing we are supposed to be explaining. But it isn’t: the claim that x is the “best means” to y (say) is one whose terms are defined by the practical premises being employed. What constitutes a ‘best result’ is itself a function of your value scheme as in (1).

It is frequently, indeed always, the case that specifications of values for entry (1) above are *incomplete*. Doing y is the best means to x, which I want; but I want lots of other things and perhaps y interferes with some of those. Reason may then call upon me not to do y after all. I intentionally slide over the huge problem of how to integrate a set of different values so that you can make choices. I do so on the ground that it’s your problem, and part of the problem, as it were, for you, but not a problem that we theorists have to solve in the abstract. All that we can say in the abstract is that an individual is considerably defined by his preferred solutions to that problem. Different people will have different selections and weightings, and they can all be acting quite rationally in so doing.

Its being “your” problem is an important category for moral theory. The area of moral philosophy we are especially interested in, though, is that of *interaction, among miscellaneous persons*. The question is what A must do in relation to B. Prominent among the things that A need *not* do, at least normally, is to solve B’s problems for him.

Thus consider those who say there are “incommensurables” in value-schemes. Are they saying that no solution to any practical problem can be rational? Presumably not. Are they saying that insofar as we have incommensurability that’s so? That depends on your answer to meta-proposals such as “choose at random!”, or (if this is different), “when faced with a choice between incommensurables, pray for approximately 30 seconds, then choose the first option that comes into your head.” At any rate, I am going to airily assume that incommensurables too are, as I put it, “your [i.e., the agent’s] problem.” The idea of incommensurability is such as to make certain problems insoluble: choice between x and y when the two are genuinely incommensurable, say. But nothing may hang on this for other persons; and if it does, then there will be an extra fact about x and y that may dictate choice of x regardless of their otherwise incommensurable nature.

“Reasoning” concerns (especially) the relation between (2) and (1). It is not a third premise or premise-like constituent. Reason is transparent. Its results show but it itself lurks in the dark regions of the soul, or one’s neurological system or such. That, I think, is what’s so cool about it. Given our preferences, and given our supply of information, rationality is a matter of putting them together, of extracting the implication for immediate action from the premises in question. Aristotle and (I take it) Hume take this to be a matter of “deduction”: I want things of type F (all things considered, more than anything else, just now); this is a possible action for me that will get me a thing of type F (if done at time t); so, I do it (at t).

The Nature of Ratiocination

One interesting point about reasoning, per se, is that apparently a lot of it - what we might call “calculation” especially - seems to go on *sotto voce* or better, *sub voce*. Suppose I am walking along a

**Papers from the 2003 meeting of the Canadian Association for Reductionist Philosophy
“Rationality” by Jan Narveson, University of Waterloo**

street, become hungry and spy two suitable eating establishments across the street, one of which is a bit farther than the other, but there’s a traffic light by the farther one whereas I have to jaywalk to get to the nearer. At a glance, I proceed to one or the other. Presumably something in there has done my homework and ground out the report that I’ll get to x sooner than y. But none of this was upfront. Or again, the famous example of the professional ballplayer who is off at the crack of the bat and just catches the long flyball - it seems, again, so plausible as to be virtually inescapable that he has a calculator whirring away in there guiding him very precisely to the desired spot, even though the player himself, consciously, couldn’t have made those calculations in a lifetime. How much of this sort of thing do we include under the heading of rationality?

My suggestion is: for present purposes, in a sense, *none*. Rather, we just take it for granted that all these subterranean operations happen and are available for further work at just about any time. Calculative reason, in that sense, is *presupposed* by rationality. But still, it is we, that is, our conscious selves, who make the decisions. Now, miscalculation is always possible, and could make a huge difference, and so the point of excluding it from our theory is only this: *qua* practical, we are only focussing on those aspects of reasoning that are *special to practice* insofar as it is or can be rational. The rest is necessary- but for another inquiry.

Note that it is possible for the agent to be aware of how good or bad his calculators are, and to act accordingly. If, knowing I am no good at calculating H, I nevertheless rely on doing it well, I act less than rationally. In that way, the point of saying that calculational capability is “presupposed” is not that it is assumed to be unlimited or beyond reach, but that information about it is part of our relevant information-base for acting (sometimes).

There is also the fact that the calculative abilities in question are hardly special to us: squirrels, bats, and assorted other remarkable creatures do some of that sort of thing better than we do. Is it only bias or effrontery that leads us to classify animals as less than fully rational? It may, we should surely allow, be a matter of degree. Again I sweep this under the carpet, or at least behind the sofa. Is there a theory of practical reasoning for apes? My suspicion is that there is, and that this is it, but that that fact doesn't really matter all that much - *pave* Descartes, say.

But this perhaps brings up a question: in what sense *do* we “make” decisions? There have been notorious recent experiments (esp. by Libet²) purporting to show that when we think we are making up our minds, we are in fact merely reporting decisions that have already been made *sub rosa*. There is, to be sure, controversy about this, but the idea is intriguing and suggestive. Some think this shows that we do not, after all, have free will. But the more general point is that there are indeed neural mechanisms - hardly an astonishing bit of news by now, one would think. The startling claim would be that our consciously made decisions aren’t in fact consciously made at all. And that would be startling, all right; but I shall take it that any such conclusion would have been supported only if defined in some peculiar way. Presumably we think that in some sense of ‘is’ the mind *is* the brain; and so, a decision made by the mind is a decision made “by” the brain. When our conscious mind is said to be “in control,” then, what is in control is *us*, with a *readout into* consciousness. If the linkage is as advertised, then the fact that our readouts come just after the subconscious machinery has done its work doesn’t really matter. What does matter is that the person in question is doing something intentionally, has assembled some information with a view to accomplishing something or other, and acts in light of the information.

What we are concerned with here, anyway, is the conceptual structure of decision-making: what makes a decision a “rational” decision. The old-fashioned answer is that the agent addresses

² Benjamin Libet, “Unconscious Cerebral Initiative and the Role of Conscious Will in Voluntary Action.”

Papers from the 2003 meeting of the Canadian Association for Reductionist Philosophy “Rationality” by Jan Narveson, University of Waterloo

the question what to do, assembles his values, assesses his facts, and chooses the option which will do best, given the facts, by his values. And this still seems to me to be, essentially, the right answer; with the qualifications and interpretations about to be noted.

The trouble is that, if it is so, it is not so as an account of a step-by-step procedure that we can be thought to be *aware of as we go through it*. Insofar as this is right, it's analytical, in the spirit of reconstruction. If I do x, having deliberated and decided to do it, then I *must* somehow, what amounts to having done all those things - all those things are relevant and must somehow have been done. They are there to assess. In analyzing a decision and asking, Why did he do it?, we have to suppose that the Agent thought he could do it, wanted to do it enough so as to be moved actually to do it, and so must have supposed that this would be “the” thing to do in the sense that if he envisaged any alternatives - as how could he not? - he must have supposed this one somehow *better*. In the limiting case where there are said to be “no alternatives” - and I am unsure that this makes any sense - then one can hardly impute rationality *or irrationality*, unless one does so farther back up the line, e.g. in stupidly getting into this situation in the first place, perhaps.

Assessing Rationality

Now, we sometimes *assess* the rationality of a given decision. Indeed, we very often do in the sense that if we or someone calls into question what we have done, we are ready either to point to the various factors that would support a verdict of “well(-enough) done”; or we are ready to say that we goofed it up. Well -- we *suppose* we can. And yet it can be argued that we *always* act rationally; that's just the way we're built. How, then, can some actions be *more* rational than others?

The easiest way to accommodate these two is, of course, to identify different but related senses. All behavior is rational in the sense that it is liable to rational criticism - done under the aegis of having reason and acting as such. But some, then, will be better than others in respects that are central to rationality.

Here it might be helpful to consider what might go wrong.

a. One's *information* may be mistaken.

a1. Sometimes the mistake is one's *fault*: you didn't do your homework (or, enough homework). Needed: a general theory of rational homework, covering when there is normative clout to be had.

b. You *misread your values*: you thought you wanted x, but it turns out that you really wanted y.

c. There was *incoherence*: some of your wants contradict others, in the sense that fulfilling one frustrates the other

d. *Limited responsibility*: you “weren't thinking”: drugged, say, or too tired to think

e. there was a calculational or computational error: your conclusion didn't follow from your premises

Things can be said about each of these.

(a) isn't interesting, I'll assume - though it's obviously important. Wrong information doesn't make for irrationality unless it is *one's fault* that it is wrong, and thus we have

(a1), which *is* interesting. But we presume that sometimes the “fault” is quite excusable, for the amount of time and effort required to get it right may not have been worth expending - which is another exercise in practical reasoning and indeed not necessarily a fault. In fact, it is often a virtue. It is more rational, not less, to act with limited information expansion of which is too costly for the purposes at hand. (This all to be part of the “general theory of homework” alluded to above.)

(b) is terribly interesting, to be sure - but also puzzling. How could one *have one's values wrong* if one's values are simply *what one wants*? But there are things we can say. For example, one thing we can say is that immediately after doing x we realize that the consequences are unpleasant, and y

Papers from the 2003 meeting of the Canadian Association for Reductionist Philosophy “Rationality” by Jan Narveson, University of Waterloo

would have been better. We learn from experience, which is hardly surprising. But we will say more on this below. And next there is

(c) - incoherence - which is a nuisance indeed: one part of you turns out to be incompatible with another, and then what? That we can fail in this way seems to me not surprising, considering the general cussedness of life and the complexity of our selves, organic and otherwise. Does it cut against rationality? It would if we, *realizing* the incoherence, nevertheless *persist* in trying to have it both ways when we know we can't. A rethink is in order.

But whether we can do that depends, I think, on the possibility, if it is one, of a further thing: you figured out what the rational thing to do would be, *but did something else*. Perhaps, even, *deliberately*. ... Is *that* possible? Does it make sense? The problem is that you *decided* to ignore your first result; but that calls in questions its status. Someone says, “why did you decide to act irrationally? You say: “I dunno, I just felt like it.” But that looks like a *change in preference*, which is surely allowable.

(d), alas, is also boring; of course we may be in no condition to be trying to sort out some thorny problem, and we'd rather just sleep, only we can't because, dammit, we have to do it *now*. I don't think this points to unreason or irrationality; it does point to the fact that life, and we, ain't perfect. And as to

(e), I am not sure whether this is boring or not. We may think: how could you *possibly* think that p and if p, then q, and yet not q?? There is, I think, something to the idea that this is, literally, *impossible*: someone who professes to think that must have *misrepresented* what he thinks he thinks. But of course this misrepresentation must, then, be possible and so it is not clear how much it helps to affirm impossibility here. What reinstates possibility, at any rate, is the recognition that calculative reason is mechanical, and machines break down, even if they are organic, neural mechanisms rather than nuts-and-bolts mechanisms.

What this all adds up to so far is that there are many ways in which things can go wrong and some of these would earn the epithet ‘irrational.’ But do any of them suggest that the general idea is wrong, and that rationality does not consist in aligning one's actions with one's values? I don't, as yet, see this.

But the most promising candidate here is no doubt (b): acting on one picture of what we want, and then deciding it was the “wrong” such picture.

Critiquing Values

Let's turn to (b), then.

b1. Now, on one (mis)understanding of this view, values are *beyond criticism*: they simply *are*, and Reason has nothing to say about them. But in order to turn that into a plausible doctrine, it would be necessary to be speaking, in the first place, of “fundamental” values: values that didn't in their turn stem from anything further, didn't inherit their valuableness from considerations lying in some way beyond the immediate boundaries of the value in question.

Let's take some examples of the sort of values that *do* inherit their value in some way from relations to other things. The most obvious, of course, are those in which the sole reason for doing x is that it is a means to something quite else, y, $x \neq y$ (*decidedly* not identical, come to that). Thus, I go to the dentist to get my teeth fixed. No fix, no go. But while that is true, it may also be true that I find some periods in the dentist chair rather pleasant, and of course it's nice to look at his pretty dental assistant now and again, and so it goes!

There are lots of mixed cases, and indeed I think that virtually all are so. Almost always, a valued experience is so both in part for its character at the time and in part by virtue of its bearing on many other things.

**Papers from the 2003 meeting of the Canadian Association for Reductionist Philosophy
“Rationality” by Jan Narveson, University of Waterloo**

b2. In any case, many values *aren't* “means” in the same sense. If x is a necessary condition for doing y, then if we want y, we'd better attend to x. My favorite example is Liberty. If I want to do x, whatever x may be, then I may have to cope with the possibility that someone will prevent my doing it, or try to. But liberty isn't a “means” in the most obvious sense of the word. It is merely a necessary condition for doing what one wishes to do - which of course still means that it is plenty important.

b3. We often hear philosophers and others asserting that some value or other is “fundamental”. But in few such cases is it plausible to think that it is a value which stands independently of anything else.

Sometimes we say we have “basic commitments.” And they might be basic in the sense that one can't readily think of anything more important that might take precedence over them or stand higher in general importance. But over time things have a way of changing, as we know. Few commitments are impervious to change, and the change is often *reasonable* - which means that the commitment in question couldn't have had quite the rock-bottom character claimed for it.

Probably the best candidates for ultimate or fundamental values are things like the taste of this bordeaux now, the sound of that cadence in a Haydn quartet, and many others of the sort. We simply like the items in question, and can think of no further reason why we do. But everything I have already said before applies to these, too.

All in all, therefore, it seems implausible to think that the “rational life” suspends from immovable, unshakable, uncriticizable values. The search for such is bound to be futile.

b4. On the other hand, we can get an appearance of such by clever semantic artifices. Is it possible to deny that happiness is good, for example? Would we not all want to be happy, if we could? But unless happiness is construed as, simply, the sum of all good, whatever that may be, any more specific ideas about it are bound to be uncertain and variable. And if it is construed that way, then it can give us no guidance. Of course we want what is good, but what is good? The idea of happiness that makes it definitionally true that happiness is good makes the advice, “seek happiness!” empty.

b5. There is much dispute, both ancient and modern on the matter of whether values are things you can “know” or not, and if they are, of course, how you would find them out. But for rationality theory, we must make a distinction. Suppose there are truths about values about which we can be wrong. Still, if A in fact believes, accepts V even though V' (\neq V) is the truth, can he be rational to act on V'? I would point out that it does not clearly make any *sense* to say that he “acts on V”. One *can* only act on *one's* values. That they are yours in the sense that they, or rather, your thoughts of them, get you into action is what makes them *yours*. But this does not in the least mean that we must accept “subjectivism”: that is, the view that if George holds that V, then, by gum, V is *the right value* for George. It seems to me that we all have the right to criticize - including George, who can change his mind, and do so in light of what he takes to be reasons.

So, why not say that George in a longer view can agree that he himself was *mistaken* earlier on? But he cannot, at that very time earlier on, say that himself. That is what makes no sense. “I am hereby acting on the wrong values” implies a contradiction. So much, it seems to me, is credit to Socrates.

We can say, “I am not *sure* this is best, but will do it anyway.” But that's very different from saying, “I *am* sure that this is *not* best, but will do it anyway.” (That seems to me a fundamental flaw

**Papers from the 2003 meeting of the Canadian Association for Reductionist Philosophy
“Rationality” by Jan Narveson, University of Waterloo**

in some formulations of the “satisficing” view of rationality; as argued in my paper of some years ago when we went round about that.³)

b6. A long-popular theory is that we should act so as to “realize our nature.” This has the obvious problem that we are always acting and in so doing always realizing our nature, and have no choice in the matter. But perhaps something can be preserved of this hoary theory. Namely, we can suppose that what we do works or not, and does so by virtue of the fact that we are put together the way we are. Our values emerge from our “natures” and when we think about them, we check out whether this or that feature that we know to be in our make-up is such as to make this or that action plausible. There’s a sort of organic satisfaction available when things work, and the reverse when they don’t (this in degrees, of course.) The suggestion, then, is that there’s no reason why we shouldn’t accept “objective value” and yet this doesn’t imply the negation of the Aristotelian/Humean idea.

That idea, to repeat in slightly fancier form, is this:

The Practical Syllogism

Major Premise: asserts a desire, interest, want, or “value” of the agent

Minor Premise: asserts a factual claim to the effect that doing something, x, which is within the unilateral power of the agent, would in fact promote the value affirmed in the major premise. And, if matters of degree are in question, x would promote it *more* than known alternatives.

Conclusion: the appropriate (or, *an* appropriate?) *action*.

The nifty thing about Aristotle’s idea was just this: that we can reason our way into action. (Note: The conclusion can also be a decision or resolution to perform some action in the future. The connection with actual action is that of a switching/storage device. We load into ourselves now a direction to do x at the appropriate time. Obviously this is subject to change of mind, erosion, or corruption; but still, what is installed is a disposition, such that in the absence of change, the action will be forthcoming.)

Now: the thesis here is that *motivation* is built into premise (1), but *not*, of course, premise (2) by itself. We take facts as such to have no motivational force. For this purpose, there is a clear distinction of facts and values. And yet, both are necessary. If we have no idea what to do to realize a value, then the value’s motivational capability cannot get us into action.

Some insist that this sharp distinction is unsustainable. But it is not very clear just what that could mean here. Either one is motivated to do something, or one isn’t. The fact that a certain action that is possible for this agent *would* promote this or that result is, of itself, without motivational force. Only a desire or the equivalent can do that.

Perhaps one can think of a desire as the *representation of a certain possibility with a motivational component*: Agent Wants that (p) = Agent envisages p, and in doing so is disposed to bring about p, somehow and if possible.

We can agree too that some objects of desire are themselves actions. In this case, the “result” is achieved simply by doing it. It should not be thought that the existence of this category contradicts the general hypothesis. It is simply one sort of case, a fairly common one.

We then turn to the “minor” premise. Here we have two distinguishable components: first, that some possible act will achieve p, and second, that this possible act is genuinely possible, or

³ This has been published: “Maximizing: Life on a Budget; or, If You Would Maximize, Then Satisfice!” pp. 59-70 in *Satisficing and Maximizing: Moral Theorists on Practical Reason* ed. Michael Byron, Cambridge University Press, 2004.

Papers from the 2003 meeting of the Canadian Association for Reductionist Philosophy “Rationality” by Jan Narveson, University of Waterloo

available, to the agent, so that he has but to choose it to do it. The idea is that in the circumstances, choosing to do x is a sufficient condition for achieving p .

But the idea is too strong. More plausible would be to see it as a sufficient condition for something a lot less: namely, raising the probability that p by some amount, possibly indiscernibly. Familiarly, when this is all we can do (and it's virtually always all we can do), we must discount the value of p by the probability of attaining it. Often that will motivate rejection of the option in question. It is arguably irrational not to do this, but if we say this, we must be careful. Suppose that Agent A *believes* that this is not so, and expects p as if its probability were 1.0. Does he act rationally when he acts on this? Given the belief, his act is rational. But the estimate may be crazy. We cannot divorce practical from theoretical reason.

We can no doubt install a modification regarding information. The extreme version is Brandt's: he proposes that full rationality calls for full information. But that is an impossible ideal; we could never tell whether someone acted rationally or not, on that view. So what's to do? Perhaps, to say that we are rational when we act on the information we have, but we can be irrational or less than rational in acquiring the information, and less than rational about how important it is to acquire it.

If we do the latter, of course, we will also have to bring in the agent's longer-term interests. But we have already allowed for that: the agent identifies with, embraces, his interests, and those may and usually do include rather long-term ones. But this identification is also variable. I may decide on occasion to ignore my long-term interests. In doing so, do I act irrationally? To say so is to assume the hegemony of those longer-term interests, but that is what I would have just renounced.

Well, here we have another way of criticizing a course of practical reasoning that is cogent and relevant. At some point, it doesn't very much matter whether we say that the agent “acted rationally.” Enough, perhaps, that he acted in ways susceptible of these (and perhaps other) criticisms, stemming from the foregoing. He acts well, rationally speaking, if he acts in such a way that none of those criticisms seriously apply. But I don't offer a threshold, other than that egregious failure on all counts get us into something less than practically reasoned action.

Reason and Morals

Did the 9/11ers act rationally? In one obvious sense, they did: their attack was carefully planned, they each did their part, they ran smack dab into the World Trade Center towers, just as planned, and did a whale of a lot of damage to a great many people - just as planned. Is that enough? Does my account imply that it is enough? I'm still ready to say, No. We can act in such a way as to pay no attention to the sort of considerations we tend to recognize as “moral” considerations. Doing so might be part of a plan, for that matter. But are there some moral-type considerations which can claim to the status of nonignorability by rational agents?

I still think there are. The rules of morals are generated by reflection on iteration of game-theoretically trackable situations, notably PD but not limited to that. If we can see, as we sometimes can, that under iteration a strategy has negative returns, then we have the basis for a strong social rule against it. Offenders ignore something that makes possible a social situation in which we can all flourish, and which indeed enables us to take advantage of our fellows who helped bring it about on the expectation that they would be treated better than *this*. Criticism of persons who act thus is rational. It is not the same, quite, as the criticism due to people who add 3 to 7 and get 13, but for all that, it is relevant. It is rational to be concerned about such behavior, and irrational to pay no attention to considerations of this kind.

**Papers from the 2003 meeting of the Canadian Association for Reductionist Philosophy
“Rationality” by Jan Narveson, University of Waterloo**

That is a longer story, to be sure. But I don't think it's a story about “intuitions.” The gains from general moral behavior are demonstrable. The gains from immoral behavior are sometimes so too, but only, as Hobbes says, on the hypothesis that one's fellows are asleep at the wheel.

General Upshot

Do we act rationally, or don't we? Or is it that sometimes we do and sometimes not? And is it a matter of degree? Our findings confirm the second option and the third. However, the view that rationality is a matter of degree segues very quickly into out-and-out discussion of value and values. The fully rational action, we may say, is the best action the agent could do under the circumstances, and in light of available information which he could and should have accessed. But we can depart indefinitely far from the goal of being sufficiently informed, and our values themselves can sometimes benefit from a re-think, and we sometimes fail in that. And thus the claim that we “always” act rationally could use the notion of rationality only in a minimal sense: a sense that below that level, we just aren't acting rationally, while above it our behavior is subject to an indefinite array of rational criticisms, but is at least not wholly wanting.