Teleology and Degrees of Freedom[1]

[illegible]

There is a debate in philosophy of mind about the nature of reason explanations of action, and this volume is testament to a resurgence of interest in non-causal accounts. In [illegible][2] I have proposed a non-causal account according to which common-sense reason explanations of action are irreducibly [illegible] in form. I claim that we explain behavior by citing the state of affairs towards which the agent was directing her behavior, i. e., by citing the [illegible] or [illegible] of the behavior. I will not be defending that account of action explanation here, but will be assuming it and applying it to the free will debate. I will argue that the teleological account of action explanation leads to a view of free will with some interesting and attractive features.

Philosophical accounts of free will typically propose some sort of criterion for

both high level general principles and considered judgments about particular cases. By going back and forth between these levels, we hope to come up with a theory that both seems plausible in the abstract but also accounts for our strongly held intuitions about cases.[3]

I will approach the free will debate in something like this way. In section 1, I begin at the bottom level by discussing some cases and intuitive judgments, particularly cases involving weakness of will and addiction. Our intuitions and practice suggest a view of free will that seems to be largely overlooked in the literature, namely, that freedom comes in degrees. In many cases, we are neither wholly free nor wholly unfree, but in some grey region in between. In section 2, I turn to higher level principles. I begin with an overview of the teleological account of action explanation, and I will argue that this account suggests a certain positive view of free will. Roughly, the view is that behaviors are free to the extent that the agent is rationalizable when performing them. In section 3, I return to diagnosing particular cases and intuitions with this positive view in hand. I argue that the view aligns well with the intuitive idea that freedom comes in degrees, and, in particular, that the view harmonizes with our intuitions and practice regarding cases of weakness of will and addiction. Moreover, I will suggest that the teleological account of free will goes further than merely agreeing with clear intuitions; in addition, in the grey areas, the view helps us to think the cases through further, both agreeing with our strong intuitions and clarifying our muddier ones. Finally, in section 4, I consider some objections. Here too, I engage in the reflective equilibrium process, for at first glance it might appear that my view of freedom gets certain clear cases quite wrong. I will fend off this misunderstanding, thereby both defending and clarifying the view.

[illegible]

The free will debate is informed by a large array of stock examples, and a prominent part is played by cases of weakness of will, compulsion, and addiction. Typically, though not always, philosophers assume that when we behave akratically we are still free, but that when compelled or addicted we are not free.[4] Either way, philosophers usually assume that in these test cases the answer is all or nothing: the agent is either free or not free. I suggest instead, that with cases like this, we should see freedom as impaired in varying degrees.

---

3 For Rawls's version in the context of political philosophy, see: John Rawls, [illegible], Harvard 1971.

4 See Jeanette Kennett/Michael Smith, »Philosophy and Commonsense: The Case of Weakness of Will,« in: Michaelis Michael/John O'Leary Hawthorne (eds.), [illegible], Boston 1994, 141–157.

In each case, some students were willing to say that the behavior was completely free, and very few were willing to say it was completely unfree – which already seems to go against most philosophical treatments which assume that cases of addiction are unfree. But my main point is that in both of these hard cases, nearly 3/4 of the students gave an answer in the middle, somewhere between simply free or unfree.

There is other evidence that we think of freedom as coming in degrees, namely, how we think of the agency of young children. I hold my 10 year old daughter largely responsible for most things she does, good and bad; in other words, I assume that she is a free and responsible agent, although even here there will be times when I think of her freedom and corresponding responsibility as somewhat attenuated. When she was a newborn infant, of course I did not take her to be a free and responsible agent. Even basic motions of her limbs did not at first seem under her control, and it would take an awfully callous parent to blame a newborn for crying or soiling her diaper. As they proceed from newborn to normal

blame or praise we attribute. If a recovering alcoholic loses her resolve not to drink and goes on a binge drinking episode, we will typically regard her as responsible for that action. However, if we find out that, through no fault of her own, she was placed in circumstances in which everyone around her was drinking and pressuring her to join, then we are less inclined to hold her as fully responsible.

I would not claim that any of the above constitutes a knockdown argument

example: a theory according to which the entire universe simply popped into existence five minutes ago, with everything in place just as we naively think things were at that time.[9] We rule out such craziness, and all manner of less obviously crazy theories, by appealing to theoretical norms that go beyond consistency with the data. In particular, I take it that both in common sense and natural science we assume something like the following general principle, labeled »(S)« for »simplicity«:

(S) Given two theories, it is unreasonable to believe one that leaves significantly more unexplained mysteries.

The five-minute theory, along with other faulty theories, fails precisely because it leaves so many coincidences or mysteries utterly unexplained.

When we are interpreting [illegible], we likewise aim to be consistent with observational data, and we construct theories in accord with (S). But, I claim, we also do something different. Loosely following Davidson's views of interpretation (but not his endorsement of the causal theory of action), I suggest that we arrive at teleological explanations as part of an overall attempt to construct a theory of an agent, and part of our aim is to produce a theory according to which the agent is as rational as possible. In general terms, I suggest that our theorizing about agents is constrained by something like the following principle.

(R) Given two theories of an agent, it is unreasonable to believe one according to which the agent is significantly less rational.

Rationality can be assessed in various different ways, of course, but two aspects are particularly relevant here. First, we assume that,

(R_1) Agents act in ways that are appropriate for achieving their goals, given the agent's circumstances, epistemic situation, and intentional states.

But not just any state of affairs can count as an intelligible goal for an agent. We assume that,

(R_2) Agents have goals that are of value, given the agent's circumstances, epistemic situation, and intentional states.

So, roughly put, there are two axes on which a candidate explanation is judged: the degree to which it makes the behavior appropriate for achieving the goal, and the degree to which the goal is of value.

---

9 Bertrand Russell, [illegible], London 1921.

In simple and straightforward cases, application of these principles is almost entirely automatic. Recall Josephine who went upstairs, and suppose that the circumstances were these: another ten year old friend had shown up at the door and asks Josephine wants to come outside to play; Josephine starts to run excitedly outside, but her father says, »wait! you need shoes! they're upstairs«. And off she goes up the stairs. And here we can see that it is quite obvious that going upstairs would be appropriate to achieving the goal of getting her shoes, and we can easily understand the value that this would have for Josephine. We might also believe various counterfactual conditionals that point in the same direction: If Josephine had believed that her shoes were in the kitchen, she would have gone there instead of going upstairs; had she believed that she already had shoes on, which perhaps she did prior to her father's admonition, she would have simply gone on outside; etc. The general point: our theory of Josephine is constructed so as to make the most rational sense we can out of her behavior in the actual and in nearby counterfactual circumstances.

Of course, further information might lead us to reject the initial hypothesis. Josephine might return downstairs entirely shoeless but carrying a favorite toy, and when her father reminds her that he told her to get her shoes, she might say that she didn't hear him. So we revise our original explanation and conclude that she went to get the toy rather than her shoes. But this is still in accord with the rationalizing principles – it's merely that we have gained further knowledge of her intentional states, and we are giving a rationalizing theory about a broader set of data. Naturally, we must also make our theory of Josephine conform with the simplicity principle, (S). Apart from our allegiance to (S), nothing would stop us from concluding that Josephine . . get her shoes upstairs, but that they magically transformed into the toy on the way downstairs.

However, if we were . . concerned with observational consistency and simplicity, and if we were not concerned with making rational sense of the behavior, all sorts of explanations would be possible and fully consistent with the data and with (S): e. g., that Josephine went up the stairs to get to France, or that she went upstairs hoping thereby to become Pope. If we are willing to attribute crazy enough beliefs and desires, any interpretation becomes possible. This is, I take it, exactly why Davidson thinks that Quine's radical interpreter can get nowhere at all without assuming a principle of charity. We rule out such interpretations precisely because they fail in our common sense psychological aim of making . . . of the person with whom we are dealing.

Naturally, these rationalizing principles do not constrain our theorizing about the behavior of inanimate things like rocks or planets. Or, to put it the other way around, on any theory according to which a rock was an agent, the rock would either come out as quite irrational, or would have too impoverished a set of goals to count as a genuine agent. If we attribute to the rock one and only one desire, the desire to follow the laws of physics, then of course the rock comes out as

the principles guarantee that we will be able to find either a perfectly rational or perfectly simple story. Since teleological explicability thus comes in degrees, the degree to which a behavior counts as goal-directed or on purpose comes in de-

behavior is determined by antecedent causes is one question; whether it is teleologically explicable is a different question. We answer the teleological question by coming up with the best theory of the agent that we can manage, where the theory must be consistent with the data but where it also must make coherent sense of the agent. This project is, on its face, different from the project of identifying causes and determining their nature. So if I am right about the irreducibility claim, then there is no obvious way that the incompatibilist can argue that determinism somehow shows that, really, all behaviors fail to be in the free action category. Teleological realism makes determinism irrelevant to agency and freedom.

Now that we have the outlines of the positive account in hand, let's return to diagnosing particular cases, especially those where I've suggested that freedom seems to come in degrees. But I'll start with a case where the agent's freedom is not typically thought to be impaired, but which nonetheless nicely illustrates the view I am defending. Martin Luther was famously brought before the Diet of Worms in 1521, and was asked to recant certain propositions he had written, propositions that had been condemned by Pope Leo X.[11] After contemplating this request, Luther is reported to have refused: »To go against conscience is neither right nor safe. I cannot, and I will not recant. Here I stand. I can do no other«.

If we take Luther's words seriously, then we would conclude that he could not have done otherwise than he did. On certain incompatibilist views of freedom, this would show that Luther was not free, which strikes most of us as counterintuitive, a point emphasized by Daniel Dennett.[12] Far from being a case of weakness of will, Luther was acting with full conviction and in accord with his best judgment. If he was being compelled, he was being compelled by what he at least thought was right reason. On the teleological account, his behavior seems eminently rationalizable, and hence it counts as a goal-directed action, a freely performed behavior for which he is responsible.

We don't know for sure whether Luther actually said, »I can do no other,« and, in any event, we might take that claim to have been hyperbole. But the general point is familiar enough. Suppose that I am walking into the voting booth, having firmly made up my mind to vote for the Democrat; I believe that she is eminently qualified, her positions on the issues are similar to mine, and, besides, I think

11 See »Diet of Worms,« 2008. 2 May 2008, http://search.eb.com/eb/article-9077510.

12 See Daniel Dennett, , Cambridge MA 1984.

that the Republican candidate is dangerously wrong on most of the issues, and would be a disaster in office. All of my reasons point towards voting for the Democrat, so that's what I do. There is a strong sense in which I can do no other. Voting for the Republican would make no rational sense, given my set of beliefs and desires. Of course, I could imagine my arm twitching involuntarily and thereby marking a vote for the Republican, and I can even (dimly) imagine suddenly reevaluating my beliefs on the spot. But, keeping my beliefs and desires the same as they were at that moment, it is hard to see myself [illegible] voting for the Republican. (Nor would I [illegible] it to be the case that I could have done otherwise than I did; i.e., it is hard for me to see the value of freedom as the libertarian defines it.) On the other hand, I take full responsibility for my vote and see it as a paradigm case of a free action, precisely because it is so clearly something I did on purpose – a goal-directed behavior. So, again, this sort of case fits neatly with teleological account of free will, but it is much more problematic on a view that identifies freedom with some sort of ability to do otherwise.

By contrast with the case of Luther, the characteristic feature of weakness of will is that we act against our own best judgment, though this can be complicated by the fact that our judgment itself might change at the moment of temptation. Indeed, I think it is useful to distinguish between what we might call our [illegible] rationality versus our [illegible] rationality. Our hot rationality is what seems rational to us in the heat of the moment, at the moment where some urge (whether for sweets, alcohol, nicotine, sex, etc.) is felt very strongly and when a fairly simple momentary action could put us directly on the path of satisfying it. Our cool rationality is what we would want ourselves to choose, when judged from the vantage point of circumstances in which the specific urge is not felt so strongly or where, even if it is felt reasonably strongly, we don't have the path of immediate gratification available to us.

So at one extreme, we can imagine the drug addict who, literally shaking with desire for another dose, succumbs to the temptation, despite having told herself many times in cool moments that she must quit. At the other end, we can imagine a fairly healthy person who is slightly overweight, and would like to lose 10 pounds or so, and thus has decided that he will forego desserts. But at a special dinner, he is presented with an exquisitely prepared fruit tart. Like the drug addict, he might give in to the temptation, but the case is quite different. It is a special occasion, the stakes are not that high, and the dessert is not that unhealthy; that is to say, even if he were to judge the situation from the vantage point of his cool rational self, it might be a close call. So while it still might be a case of weak willed action, it has little of the tinge of desperation had by the drug case. In between, there are all manner of other sorts of cases. Many of us are subject to cravings of various sorts and degrees. In the heat of the moment, all such desires can impair our judgment, as compared with what we would think in cooler moments.

So are cases of weakness of will rationalizable? Yes, to varying degrees. In each of these cases, when the agent does act and succumbs to whatever temptation is in play, the action is done in order to satisfy the urge in question. Satisfying an urge of one sort or another will typically be an intelligible goal. Thus the behavior will be teleologically explicable, and thus count as within the realm of free actions. However, depending on the nature of the action, while the agent's behavior will be intelligible, it may be far from perfectly reasonable. We may understand that the agent acted in order to satisfy the urge, but it will also be true that there was another action available to the agent that would have been more rational overall, an action serving a goal of greater value, even given the agent's own intentional states and epistemic circumstances. Moreover, if the agent's beliefs and desires themselves are far from rational, then this too will diminish the agent's overall rationalizability.

In minor cases of weakness of will, like the dessert case, resisting the temptation would be more in line with the agent's desires and values, but it might be a close call. The action of taking the offered fruit tart is easily rationalizable, even if there was an alternative that would have been somewhat more in line with what the agent valued. On the other hand, the morbidly obese person with the defective leptin receptors knows that it is very much in her best interests to leave the bag of cookies alone, and she knows she will regret it later, but she acts instead to satisfy the ravenous hunger she feels at the moment. Moreover, given the hunger she feels, she would have eaten the cookies almost irrespective of how much she had eaten that day already and how bad the cookies were for her. So, although satisfying hunger uuosflready agand oulu/Wehe

counterfactual circumstances, the agent would be ultimately far better served by other actions. Accordingly, on the present account of freedom, we conclude that the addicted agent is only minimally free.

What if we are dealing with a willing drug addict? For example a cigarette smoker with extremely strong nicotine cravings, but who positively affirms her smoking habit and has no desire to give it up? On standard sorts of incompatibilist views, the unwilling drug addict and the willing addict are in exactly the same position, and it all depends on whether or not it was physically possible for them to do otherwise. Given the incompatibilist position, if determinism is true then we are all unfree, including the addicts. If determinism is not true, then whether this particular action was free depends on the physical possibility of doing otherwise. [illegible] we can conclude that it is physically impossible for either addict to act otherwise, and conclude this because in cases of severe addiction it [illegible] like one has no choice. But it is a real question whether this feeling is indicative of a genuine physical impossibility. In any event, the incompatibilist will probably put the willing and unwilling addicts into the same category, and that would probably be the unfree category. Some compatibilists, on the other hand, might distinguish between the two cases. Harry Frankfurt[13] says that the unwilling addict is not free, because his actions are not in accord with his second order desire to stop taking the drug. The willing addict, however, is different, for his second order desire is to continue using the drug, and Frankfurt asserts that he does act of his own free will.

On the teleological account of freedom, there is no need to shoehorn the addicts into the »free« box or the »unfree« box. The unwilling addict, as suggested just above, is only minimally free, and is less free to the extent that her drug use goes against her interests and violates her own judgment about what is best for her to do. What of the willing addict? Obviously, she is not acting contrary to her own judgment of what she should do, and thus her values and beliefs are at least, to that extent, more internally coherent and thus more rationalizable than those of the unwilling addict. So she is more free than the comparable unwilling addict, and this seems right: we will generally hold someone more responsible if she affirms and relishes her drug use than if she despises it and struggles against it. But the drug user does not become fully free simply by having the conviction that she likes using drugs. Where she falls along the spectrum of freedom and unfreedom will depend on other details. For example, some coffee drinkers report feeling quite addicted to coffee in the morning and will go to some trouble to obtain coffee if they find themselves out of it at home, or if they are staying in the home of a non-coffee drinker. Nonetheless, they might fully endorse their habit.

---

13 See Harry Frankfurt, »Freedom of the Will and the Concept of a Person,« in: [illegible] 68 (1971), 5–20.

At the other extreme, there could be a heroin addict whose life revolves around getting her next dose, who has lost her friends, family, and job, and whose life is on the brink of collapse; but she might nonetheless endorse her heroin use. While the behavior of the willing heroin addict is somewhat more rationalizable than that of the unwilling heroin addict in analogous circumstances, the willing heroin addict is still very different from the willing caffeine addict. Her heroin use is at odds with all sorts of other things that she does to value in life, even if she is currently blind to seeing those values. While we can understand the attraction of the euphoric high said to go with use of the drug, we conclude, with reason, that the heroin addict's life is far from ideal, and we suspect that the addict herself would see this too if she had some appropriate distance from her use

furt[15] type views, for example, whether our will is free depends on whether we act in accord with our second order volition – i. e., whether the desire we wish to act on is the one we in fact act on. By definition, we exhibit weakness of will when we act contrary to our own best judgment, so Frankfurt should count all such actions as simply unfree. Similarly, on Gary Watson's view, we are free when our desires and our values are in harmony, which is precisely what is not happening in cases of weakness of will; thus, again, cases of weakness of will are automatically unfree.[16] Surely this is a counterintuitive result, especially in minor instances like occasional overeating or watching mindless television rather than washing the dishes. On the other hand, I find it equally counterintuitive to claim that severe cases of weakness of will are simply and straightforwardly free actions.reunti. On il

the note he says, »I felt good and all washed clean of sin for the first time I had ever felt so in my life.« But as he reflects on the time he has spent with Jim, he reconsiders:

> It was a close place. I took [the note] up, and held it in my hand. I was a trembling, because I'd got to decide, forever, betwixt two things, and I knowed it. I studied it a minute, sort of holding my breath, and then says to myself:
> »All right, then, I'll [illegible] to hell« – and tore it up.[17]

Huck acts against his own best judgment, believing that he will even suffer eternal damnation for his action. But in fact it was the right thing to do, and we have no trouble in seeing the value in his action, even if Huck himself is convinced otherwise, and thinks that he has just been incredibly weak and has committed a grave sin.[18] Huck has been taught and is operating within a system of values and beliefs according to which black people are mere property; Huck appears incapable of consciously thinking his way out of that system; he can't bring himself to consciously affirm that it is some of the values he has been taught that are gravely mistaken. But on another level, he knows this, and when he acts he goes with his emotionally laden instincts rather than his conscious judgment. And this choice, besides being the admirable and right one, ends up being at least as free as the alternative of turning Jim in. Huck would not have been more rationalizable as an agent had he taken a Spock-like approach, ignored his emotions and thought only about what seemed most logical.

In some ways, Huck seems like a special case, since the issues are so weighty, and Twain is so adept at portraying the inner torment of the boy whose emotional perceptions fly in the face of societally imposed values. But the point applies to more trivial contexts as well. One might impulsively decide to stop working on a philosophy paper, grab the kids and head to the beach; or one might spontaneously put the housecleaning on hold and go off to a bar with a friend. Such actions are quite rationalizable. Indeed, even if these precipitate actions are contrary to the agent's carefully considered plans for the day, the actions might nonetheless be at least as valuable as the more deliberate alternative. It might even be important to the value of the actions that the agent [illegible] as if she was playing hookey. Being rational, in the broad sense I have in mind, does not necessarily mean always carefully planning and weighing options logically. It means instead pursuing courses of action that are of value from the agent's perspective, and actions with this feature are not necessarily coextensive with actions that are carefully planned. Perhaps if we were Godlike in our ability to deliberate and plan, things would be different, for if we were always perfectly rational thinkers and

17 Mark Twain, [illegible], 1884, chapter XXXI.

18 Susan Wolf makes a similar point about this example (see [illegible], New York 1990).

planners, then it would presumably be irrational ever to spontaneously discard our plans and act on whim. But since we are not perfectly rational when deliberately making plans, it can happen that our emotionally informed impulses or whims are, on occasion, more trustworthy and more reasonable overall than our more consciously employed rational faculties.

Antonio Damasio describes a famous case that also illustrates this point.[19] Phineas Gage, a railroad worker in the mid 19th century, suffered a terrible accident in which a large metal bar was driven completely through the front part of his skull. Astonishingly, Gage survived, and his speech, memory, and intelligence seemed intact. However, Gage was far from himself. His ability to interact socially, and his ability to make intelligent decisions, were both greatly impaired. Damasio also describes a contemporary case of a man, whom he calls Elliot, with somewhat similar symptoms. Elliot had had a large brain tumor removed, and there had been damage to frontal lobe tissue. After the surgery, Elliot's intellectual abilities, as measured by a large variety of tests, were still intact. Nonetheless, Elliot could no longer hold down a job, and was constantly making decisions with detrimental consequences. Damasio observes that Elliot seemed virtually devoid of emotional reactions, either when telling his own sad history, or when being shown images of natural disasters and the like. Damasio says, »We might summarize Elliot's predicament as [illegible]«[20]. Damasio concludes that Elliot's defect in decision making and social behavior was connected to this emotional deficit:

> I was certain that in Elliot the defect was accompanied by a reduction in emotional reactivity and feeling […] I began to think that the cold-bloodedness of Elliot's reasoning prevented him from assigning different values to different options, and made whenu5sh6.6on makig pduT568 Tw 0 -

Once Jake has made his decision and fails to install the devices, the second explanation is, of course, ruled out. But comparing (1) and (2) becomes relevant when we rate Jake's overall rationality. If Jake [illegible] more rational if he were such that (2) was true, then this means that the truth of (1) makes him less rational than he otherwise might have been.

To illustrate, compare this with a case of weakness of will. Suppose that there is a different agent, and one of two explanations might become true of him tonight:

(3) Scott spent the evening watching television and eating potato chips for the pleasure of salt, fat, and mindless entertainment.
(4) Scott spent the evening writing in order to finish his free will paper

If in fact Scott spent the evening watching television, then (4) cannot be the correct explanation of his behavior. But in rating the rationality of Scott's action, the alternative of (4) becomes relevant: Scott would have been more rational had he taken that course. But note that, if (4) were true, Scott would be more rational, in part, because his action would match his own best judgment about what he should do. As seen above, it is not always irrational to act contrary to one's best judgment, but such actions automatically involve [illegible] tension; acting against one's best judgment at least shows some failure of coherence between behavior and values. That is part of why akratic actions are less rationalizable.

But the factory owner, Jake, is not acting against his own best judgment. We are supposing that Jake made the calm, cool judgment that he didn't really care about long-term environmental consequences or about the fishing families living downstream; he just wanted to maximize his own already substantial profits. To those who know Jake, this comes as no surprise, for he has always been rather self-centered and disdainful of the interests of others around him. If Jake had suddenly chosen to pay for the anti-pollution devices, then, at least in one respect, he would have been [illegible] rational, for he would have been acting contrary to his own beliefs, desires, and judgments. In other words, there are different ways in which an action can be of value. In terms of objective, overall value, it would be better if Jake put the anti-pollution devices on. But, as Jake sees things, it is better for him to make more money. Recall that we are trying to tell as rational a story as we can about the agent, given the agent's circumstances and intentional states. This last proviso means that self-consistency and coherence make for a more rationalizable agent.

This is not to say that one's intentional states can rationalize any behavior or that Jake is perfectly rational so long as he is acting in accord with what [illegible] values. Jake's rationality is still diminished by his mistaken judgment that he should maximize his own profits at the expense of the fishing families trying to eke out

a reasonably consistent and coherent whole. Still, the very fact that we believe him to be making a [illegible] in his value judgments does mean that there is a sense in which we believe that he could be yet more rational, and in this same sense he fails to be perfectly rationalizable.

But before we conclude that Jake is less than a free agent, we should consider a couple of things. First, if we demand a status of perfect or near perfect rationality before we count an agent as free, then who among us will achieve that status? While the view is that freedom comes in degrees, and this suggests a spectrum from zero to a perfect 100 %, we needn't be mindlessly rigid and require 100 % rationalizability before we simply count an agent as free. There is an analogy here to degrees of confidence in our beliefs. We are more confident of some of our beliefs than of others, even if we would be hard pressed to attach an actual percentage degree confidence to individual beliefs. If we are clear headed and even remotely impressed with skeptical arguments, we should admit that there are few, if any, of our beliefs concerning which we should claim absolute 100 % confidence; but we still straightforwardly and correctly say that we believe certain propositions even if we are not completely and utterly certain of their truth. Similarly with freedom: we can straightforwardly and correctly say that an agent is free even if the agent is not perfectly and completely rationalizable.

The second point is that it may be misguided or incoherent to talk of a simple linear scale of degrees of rationality or rationalizability. There are various differing facets of rationality: the consistency and coherence of our beliefs, of our values, the correctness of our beliefs and values, and the degree to which our behaviors are in accord with what we value. And there are arguably distinct kinds of value: short-term prudential, long-term prudential, concern about friends and loved ones, moral values, etc. If there is a way of translating all of these factors into a linear scale of rationality, I will certainly not be attempting it. But there can still be a clear sense in which some actions and agents are more rationalizable than others, even if the standards are defeasible and open-textured, and even if there are epistemological problems in ascertaining correct answers in certain cases. So, even though Jake may not be perfectly rational, he is still free, and the very fact that he chooses the wrong action does not exonerate him from responsibility. Of course, the point generalizes: when we make poor choices, then this does show that we fail to be perfectly rational, but it does not thereby sink us into degrees of unfreedom that will begin to exonerate us from moral responsibility.

[illegible]

Descartes's claims in the [illegible] notwithstanding, philosophical argument rarely proceeds by ironclad argument from supposedly indubitable premises to now-indisputable conclusions. In this paper, I have suggested that a plausible view