Moral Rationalism and Rational Amoralism*

Mark van Roojen

Metaethical rationalism can be roughly characterized as the idea that the requirements of ethics are requirements of practical reason. The idea is attractive, in part because it can explain the plausibility of certain versions of motivational internalism about moral judgments. Since rationalism entails that right action is a species of rational action it appears rational people must be motivated to do what is right, something many internalists believe.

But rationalism's attractions are often not well enough appreciated because the very feature that makes it attractive also generates a prima facie objection. Rationalism seems to require that those who refuse to acknowledge correct moral demands therefore be irrational. Yet such people don't always seem irrational to us. People sufficiently removed from ourselves in time, place, and culture often have a divergent conception of what morality requires. If we are right about what morality requires, then they are wrong. Yet it seems unfair to accuse them of irrationality as opposed to some other sort of mistake; nothing in their

experience prepared them to see things in the way morality requires. Aristotle might be an example. Still other apparently rational individuals remain unmoved by what they believe right. If this means they are not moved by what they believe is most reasonable to do, the lack of motivation would appear to count as a species of irrationality. But that is not always how the cases strike us. Huck Finn's refusal to turn in his friend Jim was not a failure of rationality. It is partly on the basis of such examples that many theorists conclude that the requirements of ethics cannot be the requirements of practical reason.[1]

In this essay I will defend rationalism against these worries. But I hope to do more than that. I intend to show how the rationality of people like those described above is compatible with two plausible versions of internalism. Second, I will show how properly formulated rationalism serves to explain these plausible internalist theses along with the plausible cases of rational amorality and immorality which they allow. The result will be that a plausible internalism and a well-formulated rationalism are mutually supporting theories.
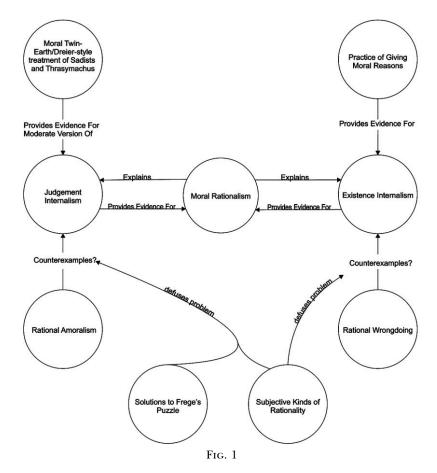
Two sets of ideas are critical to my argument. One turns on the recognition that reasoning is a process and that what it is rational for a person to think or do can depend on features of her history, circumstances, or information. This holds for both a posteriori and a priori reasoning. I embed the relevant points within a framework distinguishing various senses or kinds of rationality, each of which can be defined relative to distinct features of the agent's history, psychology, and epistemic or practical situation. When fully worked out, this package helps us to explain how rational people can have the wrong moral views.

A second set of ideas interacts with those just described to handle a different sort of counterexample—that of rational persons who may or may not have the right moral views but who are unmoved by their moral beliefs. Here I invoke considerations familiar from the literature about Frege's Puzzle and related issues in the philosophy of mind and language to show that rationalism does not rule out such examples. And I argue that the resulting view is still powerful enough to defend a moderate internalist thesis connecting morals and motives, one which has real bite but does not render the counterexamples impossible.

Together these ideas allow rationalists to defend two plausible and moderate versions of internalism about moral judgments, one connecting the truth of moral judgments with rational motivation in certain conditions and the other connecting belief in a moral judgment with

---

1. See Michael Stocker, "Desiring the Bad," *Journal of Philosophy* 76 (1979): 738–53; David Brink, "Externalist Moral Realism," *Southern Journal of Philosophy* 24, suppl. (1989): 23–40; and Mark Twain, *The Adventures of Huckleberry Finn* (New York: Charles Webster, 1885).

Fɪɢ. 1

motivation. They allow rationalists to argue in favor of their overall position as the best explanation of the sorts of internalism that are most plausible. Since the argument is complex, the diagram in figure 1 may help readers to understand the relations between the key claims.

A TERMINOLOGICAL NOTE

Since I am going to use the word 'rational' and its relatives quite a bit, I should clarify how I am using the term. I intend to use 'rational' as a term for a very general normative property, roughly the property an action, intention, or belief has if and only if (iff) it makes sense. An action is rational iff it makes sense to do it. A belief is rational iff it makes sense to accept it. An intention is rational iff it makes sense to

adopt it. To show that a belief, desire, or action is rational is to justify holding the belief or desire or doing the action.[2]

WHAT SORTS OF INTERNALISM NEED EXPLAINING?

Two different sorts of internalism are favored by arguments independent of any particular metaethical theory. (1) It is plausible that having a moral obligation to do something is necessarily a reason to do it or, put another way, that true moral propositions give us reasons to act in the ways they commend. And (2) it is plausible that there is a necessary connection between believing something right and being motivated to do it. The first sort, morality/reasons existence internalism, connects true propositions (whether believed or not) with reasons for action. The second sort, morality/motives judgment internalism, connects moral judgments with motivation on the part of those who accept them.[3]

Both sorts of internalism are controversial, so I will briefly defend each and the particular versions of each employed in my argument. Existence internalism connecting moral truths with reasons is supported by the role moral arguments play in the justification of actions. When someone asks for a reason to do something, it is appropriate and not obtuse to explain that the action in question is morally right and to offer an explanation of why it is right. No further answer to the why question would normally appear to be needed.[4] Critics of internalism are correct in pointing out that some agents may be unsatisfied with this answer. A person can doubt that she has a reason to do the action in question even in the face of such an explanation. But this does not by itself show that there is no such reason. What it shows is that not everyone accepts internalism. If the doubters are rational, one might, however, think that this result would show that existence internalism is false. After all, how could rational people ignore reasons they have? My account of rationality and the way it is tied to morality is partly aimed at diffusing this worry. Once the account is in place and further nuanced to accommodate Frege Puzzles, we will have reason to think such doubts compatible with the claim that moral judgments are essentially connected to reasons. This is a promissory note—I need much of the rest of the argument in the essay to make good on this claim. If the objection can be defeated as promised, the prima facie case rooted in our use of

2. Allan Gibbard, *Wise Choices, Apt Feelings* (Cambridge, MA: Harvard University Press, 1990), 6–7, aims to elucidate a similar use.

3. Stephen Darwall, "Reasons, Motives, and the Demands of Morality: An Introduction," in *Moral Discourse and Practice*, ed. Stephen Darwall, Allan Gibbard, and Peter Railton (Oxford: Oxford University Press, 1993), 305–12.

4. H. A. Prichard, "Does Moral Philosophy Rest on a Mistake," *Mind* 21 (1912): 21–37; W. D. Falk, "Ought and Motivation," *Proceedings of the Aristotelian Society* 48 (1947–48): 492–510; and Michael Smith, *The Moral Problem* (Oxford: Blackwell, 1994).

moral arguments to justify and give reasons for action is sufficient to support the moderate existence internalism which rationalism entails.

What defense do I offer of judgment internalism connecting moral judgments and motives, and what sort of necessary connection do I defend? My favored formulation is one suggested by Jamie Dreier which gathers support from thought experiments discussed by Hare, Dreier himself, and Horgan and Timmons.[5] Roughly stated, the formulation claims that a rational person who believes an available action right will normally be motivated to do that action.[6] Whenever we find a rational person who sincerely expresses a judgment that an action is morally right and who yet remains unmoved, it will be a case in which that person is abnormal in some way and one where other normal people appropriately related to that person would be moved by such a judgment. Following Dreier, I will call this 'moderate internalism' or 'moderate judgment internalism'. Thought experiments which support this version of internalism all suggest that our willingness to translate a foreign term with a moral term of our own depends upon the use of that term by normal members of a community to express action-guiding judgments.

Let me illustrate by discussing some examples from the literature. Note that moderate internalism is consistent with the possibility of Brink's amoralist, a person who sincerely avows that some action or other is right and yet claims to have no motivation whatsoever to do the action in question even when she is in a position to easily do so. Suppose we ask ourselves why we are inclined to take the amoralist's avowals to express the belief that the action in question is right. It is not just because her term 'right' is the same as our term 'right' which we use for that purpose. If we think of the amoralist in isolation, uttering the same sentence and showing no motivation to do what she calls "right," there is no reason to attribute a thought about rightness to her. When we think of the amoralist as expressing thoughts about rightness, we imagine her as someone like Thrasymachus,[7] as a member of a speech

---

5. See R. M. Hare, *The Language of Morals* (Oxford: Oxford University Press, 1952); James Dreier, "Internalism and Speaker Relativism," *Ethics* 101 (1990): 1–26; Terrence Horgan and Mark Timmons, "Troubles for New Wave Moral Semantics: The 'Open Question Argument' Revived," *Philosophical Papers* 21 (1992): 153–75; and Mark van Roojen, "Knowing Enough to Disagree," in *Oxford Studies in Metaethics*, ed. R. Shafer-Landau (Oxford: Oxford University Press, 2006), 1:161–93.

6. Dreier, "Internalism and Speaker Relativism," defends the principle that necessarily in normal contexts a person will have some motivation to promote what he believes to be good. I substitute rightness for goodness and adopt the limitation to rational people suggested in Christine M. Korsgaard, "Skepticism about Practical Reason," *Journal of Philosophy* 83 (1986): 5–25; and Smith, *The Moral Problem*.

7. Brink, "Externalist Moral Realism," 30.

community. We attribute thoughts about rightness to the amoralist because she is part of a speech community that uses the term to predicate rightness of actions. This then raises a question about why we are confident her community is using the term to predicate rightness. The answer is that community members use the term in such a way as to guide action. Simply stated, they treat the believed rightness of an action as sufficient to rationalize a motive to do that action. The upshot is that we can imagine people who believe an action right while remaining unmoved, but only against a background in which this is not the normal case.[8] Moderate internalism is thus vindicated by careful consideration of Brink's purported counterexample.

HOW RATIONALISM EXPLAINS EXISTENCE INTERNALISM AND MODERATE JUDGMENT INTERNALISM

Existence internalism postulates a necessary connection between having a moral obligation and having a reason. By reducing facts of morality to a (possibly improper) subset of facts about rationality, rationalism entails this sort of internalism. For example rationalism might say that to have an all things considered moral reason to φ, is just to have an all things considered reason (perhaps with the right sort of ground) to φ. For it to be right to φ is for it to be rationally required to φ (again perhaps on certain grounds).[9] And so on. The analysis entails that true moral claims imply that an appropriately situated agent has reason to do what they commend. This is existence internalism. This much is simple. The complication lies in explaining how rational people can be unmoved by what they have moral reason to do—a task I will move to shortly.

It is not as simple to show that judgment internalism falls out of rationalism. Judgment internalism entails that even false moral judgments are necessarily connected with motivation in those who believe them. Brink's amoralist is normally considered a problem for judgment internalism, but I have already explained how moderate judgment internalism makes room for the amoralist. What remains to be explained is the moderate internalist claim: necessarily, normally rational agents will be motivated by the moral judgments they accept.

8. Dreier's "Internalism and Speaker Relativism" presents essentially this argument in discussing Gideon Rosen's sadists, a group who are motivated to violate the moral norms of their society. Dreier suggests he can't define what it takes to be normal, but I think the rationalist account that follows will help fill out a story.

9. I'm being intentionally vague about the exact nature of the claimed identity. Though I can't argue for it here, I think 'is rational' and 'is right' contribute the same property to the literal meaning of sentences embedding them. This view has drawbacks, including that intuitively not every reason seems to be a moral reason, hence the parenthetical remark.

Rationalism suggests a very tight necessary connection between sincere moral judgment in rational people and motivation. It explains this connection in much the same way that we might explain why rational persons will do what they believe they have most reason to do. For, according to a rationalist, to a first approximation the belief that φ-ing is right for one is equivalent to the belief that one has overriding reason to φ. And it looks like it is a requirement of rationality that one be motivated to do what one believes one has overriding reason to do.[10] The problem for rationalism is thus explaining the weaker version of the theory suggested by moderate internalism. If there is a requirement of rationality to do what you believe you have most reason to do, doesn't rationalism make those who are unmoved by their sincere moral judgments irrational? As I'll explain below, there are reasons to think that even rational people can be unmotivated by what they regard as true moral judgments. This is why we will need to include Dreier's 'normally' in the correct statement of internalism even if we have already limited ourselves to quantifying over only rational people.

My explanation of the details—that is, how abnormal but rational failure to be moved is consistent with rationalism—requires both a sketch of a theory of rationality and some mode of theoretical response to Frege's Puzzle. Here I just note that the problems for squaring rationalism with each kind of internalism are similar. Each involves explaining how the postulated necessary connection between moral judgments and motives can be rationally disregarded consistent with moral requirements remaining rational requirements.

ADDITIONAL RESOURCES FOR THE THEORY: HUMAN RATIONALITY

A moral rationalist should work with rationality for human beings.[11] And because humans are limited in various ways, the appropriate conception of rationality should take those limits into account. We have limits on how much we can know, on what kinds of investigation we can pursue, on what we can do, and on what we can perceive. Partly as a result of such limitations human rationality is something both more and less

10. The exact nature of the requirement depends in part on whether the rationalist thesis identifies morally right action with rational action simpliciter, or with a proper subset of rational actions, perhaps those based on regard for others. The latter sort of rationalism generates a weaker internalist requirement; it requires even all things considered moral judgments to motivate rational agents defeasibly. The former sort would require all things considered moral judgments to motivate indefeasibly. I most naturally put my points in such terms, but analogous versions of each point should be available on the proper subset view.

11. This section of this essay draws heavily on Mark van Roojen, "Motivational Internalism: A Somewhat Less Idealized Account," *Philosophical Quarterly* 50 (2000): 233–41.

demanding than having full information. It is more demanding because a person with full information is not thereby rational; it is less demanding because one can be more or less rational despite lacking information. To be sure rationality requires true beliefs about a substantial number of things, but few of these are directly and universally specifiable by their content. Which true beliefs a rational person should have depends in large part on the evidence to which that person has been exposed. In this way, what rationality specifically requires of a person often depends upon that person having had certain experiences or having gone through certain thought processes.

Perhaps not every rationally required belief is like this. Perhaps some are just straightforwardly required by rationality; Descartes' *cogito* may be an instance. I will not quarrel over such examples although I will insist that they are less common than one might suppose. There is a temptation to think of all a priori knowledge as rationally required without regard to the knowing agent's circumstances or history. Perhaps this is because of a tendency to think that the conceptual nature of a priori knowledge means that grasping the concepts needed to express them is sufficient to justify a person in believing their truth.

This is misleading. Even if a priori knowledge is conceptual knowledge, knowing the conceptual truth can often involve a great deal more than understanding the concepts involved. Take some of our best candidates for such knowledge, say knowledge of arithmetic or logical truths. It would be absurd to fault a person's rationality for lacking any moderately complicated bit of mathematical or logical knowledge even where she understands all of the concepts involved. The reason is that knowledge of many such propositions requires proof—a process of justification of these propositions starting with better established claims. If I am rationally required to believe one of the less immediately obvious propositions a priori, that requirement rests on my having gone through an appropriate process of reasoning. The *a priority* of a conclusion thus does not exempt the rationality of believing it from dependence on historical facts concerning the person who is required to accept it.[12]

This brings us to another kind of requirement and limitation to human rationality. Rationality concerns not only what to believe in what circumstances but also what efforts we should make to collect evidence and what reasoning we should do. Hence there are rational requirements to pursue evidence and also to reason our way through to various conclusions in appropriate circumstances. For us humans there are limits to our abilities in these regards, and rationality for creatures such as us reflects that. We are not required to accept every consequence of

12. See Gilbert Harman, *Change in View* (Cambridge, MA: MIT Press, 1986).

everything we believe nor to collect all the evidence relevant to our possible beliefs.

So far this should be uncontroversial. But I've left out a large part of rationality, that part regulating intentions, actions, and the interrelation of belief and desire. The details of this part of rationality, practical rationality, are controversial, but the claim that rationality is concerned with practical matters is not. There are obviously rational requirements relating means to ends. Some will believe these sorts of norms to be all there is to practical rationality. Others will supplement them with norms governing the choice of ends themselves. Whatever the exact content of such principles, they are important, and importantly different than requirements governing only belief. Failure of rationality in this respect is not reducible to failure of rationality with respect to belief.

Practical rationality also respects human limits. It includes principles that have application just because there is only so much one person can do. We each have many more ends than we can bring about in one life. Thus it cannot be a requirement of rationality that we do whatever is necessary to bring about every end. There will be conflicts. We need to choose between ends. There will be controversy over what norms should govern such choices, but this does not undermine the idea that there must be some such norms. Even if there are many different correct ways to choose, some choices will be irrational and some choices will be more rational than others.

Human limitations have another role to play here analogous to the role they played in determining the rationality of belief. Since we have limited time to think through our options, and since the best way to trade off will not always be clear, the rationality of choosing one way versus another will be partly a function of the opportunities a person has to think things through. Flipping a coin to determine a trivial matter where time is tight makes sense; relying on a coin toss where the matter at hand is important and time is ample does not. Furthermore, the rationality of a practical decision is also in part a function of the actual process of thinking it through. While it might be rational to embark on a given course of action if one has not thought through its consequences, it may no longer be rational once one anticipates certain bad effects.

Thus rationality involves multiple kinds of requirements, the application of which depends on factors regulated by other requirements. Each of these is subject to human limitations, generating still further principles for dealing with such limitations. The upshot is that what even full rationality requires of a person will depend on a variety of factors, including the situation the person is in and the opportunities that situation gives the person both for investigation and for action, as well as on the actual history of deliberation that the person has engaged

in. Furthermore there are a variety of ways that people can depart from full rationality, even holding fixed such background conditions. These departures too can ground reasons.

We can idealize along each of the relevant dimensions when we consider what rationality requires of a person. We can idealize people's epistemic positions to a greater or lesser extent while holding fixed that they are fully rational and thereby vary what we think a fully rational person would know or do. We can hold fixed what the person believes and desires, or the person's evidence and opportunity to deliberate, and rank courses of action for rationality.[13] We can hold fixed even an irrational feature of the agent's psychology and rank different courses of action for rationality given that fixed psychological feature. These rankings will have to take into account a variety of factors—all of the factors I have been characterizing as distinctive of human rationality. One important upshot is that many rational requirements would not exist for creatures who were more ideal than us in various respects. We have reason to seek additional information because we don't know certain things. We have reason to choose an outcome with a high expected benefit because we don't know and don't have time to find out which option will actually have the highest actual benefit. We have reason to avoid temptation because we know that we can be successfully tempted. And so on.

The idea here is related to a standard proposal for drawing a contrast between what a person should rationally do objectively speaking versus what they should do subjectively speaking. Roughly put, the standard proposal is that an option is objectively rational if, given the actual situation, it would make sense for an agent to choose it. It is subjectively rational if given what the agent believes it makes sense for her to choose it.[14] The basic idea here is fine, but I think it is presented too simply. There is not just one determinate way to make this sort of contrast since different subjective features of an agent can determine what it makes sense for that agent to do. Thus for each such feature, what makes sense for the agent to do with the feature present differs from what it would make sense to do were it removed. Each of these features might be used

---

13. We can rank departures from rationality in terms of seriousness, even when they involve breaches of different rational norms. See van Roojen, "Motivational Internalism."

14. Philosophers have long used the terms of art 'objective' and 'subjective' to mark distinctions of roughly this sort. See, e.g., Henry Sidgwick, *The Methods of Ethics*, 7th ed. (New York: Macmillan, 1907), bk. 3, chap. 1, sec. 3; H. A. Prichard, *Duty and Ignorance of Fact*, Henrietta Hertz Trust Lecture (London: Humphrey Milford, 1932), 18–39; A. C. Ewing, *Ethics* (New York: Macmillan, 1953), 63; and Derek Parfit, *Reasons and Persons* (Oxford: Oxford University Press, 1984), 25. Not all of these formulations are equivalent. For discussion, see Mark Schroeder, "Having Reasons," *Philosophical Studies* 139 (2008): 57–71.

to generate something like the contrast between subjective and objective rationality.[15]

To illustrate, an agent may only have probabilistic information about the results of her actions. Given that limitation on her actions, it may be rational for her to choose the action which maximizes expected value even if objectively speaking that action will not bring about the most desirable result. Here we can say that objectively she should not have done what she did, but subjectively speaking she did do the right thing. Alternatively we might be interested in what is rational relative to information she could have had had she done some investigation prior to deciding what to do. This sort of rationality is also subjective in one good sense though it is not the same as the first sort. It is also objective in a good sense insofar as it is relative to information not currently subjectively available to her.

The examples above involve no failure of rationality since a lack of information or time to think is not a rational failure. But some reason-grounding limitations differ from these insofar as they constitutively involve irrationality. One example of this type has already been mentioned: I might be weak willed and hence have a reason to avoid certain sorts of temptation. This means that an otherwise rational person who knows she has this sort of disposition will avoid temptation when it is within her control to do so, other things equal. In one good sense the person has a reason to avoid (say) stocking the freezer with ice cream. In another sense there is no objective reason for her to do this, since were she fully rational it would be convenient if the freezer contained ice cream and a rational person would not be weak willed.[16]

FREGE'S PUZZLE AND STRATEGIES TO CAPTURE IT

Philosophers are fond of offering analyses that postulate identities between items that seem to be distinct. And they often argue for these identities by suggesting that, if true, they would help explain why com-

15. Judith Jarvis Thomson, in *Rights, Restitution, and Risk* (Cambridge, MA: Harvard University Press, 1986), 179, is skeptical of any such distinction. The important point for the present view is not about the meanings of the words. The point is that what makes sense to do depends on features of the agent's situation, including facts about the agent's information, self-control, time, etc. When people are thinking about what it is rational for an agent to do they allow these sorts of features to affect their verdicts about what is rational. Whether that is because the word 'rational' genuinely has several meanings or whether features of the context of utterance or evaluation allow us to use a term with one univocal meaning to convey different information in different contexts does not deeply affect the most important issues here. A number of different semantic accounts would deliver what the argument needs.

16. We could use the terms 'subjective' and 'objective' to draw the relevant contrasts in several different ways. I don't think there is anything of substance that turns on how we choose to talk about this, so long as we are clear in what we are saying.

petent cognizers act in various ways that would be explained by knowledge of those identities.[17] But various examples presented by Frege suggest we must all recognize that facts about identities, including facts about the identity of properties, are not always cognitively available to people, not even to fully rational people. Even though Venus is the Morning Star, it appears one can know that one is looking at the Morning Star without knowing that one is looking at Venus. This complicates the above sort of argumentative strategy for establishing a philosophical analysis. The explanation of the relevant cognizers' behavior seems to require not only that the entities are identical but also that the cognizers are in a position to know that the things in question are one and the same.

Rationalism is or entails one such identity thesis. It claims that one property of actions—rightness—is identical with another—that of being rational to do. And theorists often argue for it in roughly the way indicated above, by showing how its truth would explain the actions of competent cognizers employing moral concepts. Frege Puzzles complicate this strategy of argument for rationalism, just as they complicate similar arguments for other philosophical analyses. If such cognizers can rationally doubt the identity, and indeed they can, the explanation of what they do must be more complex.

Furthermore it doesn't really matter whether or not the identity is knowable a priori or only a posteriori. One important upshot of the previous section of this essay is precisely that lack of knowledge of a priori matters is not always, or even often, a rational failing. Even if a fact can be known a priori, knowledge of it may still depend on having gone through the relevant process of reasoning to show that it is true. Thus, as someone who wishes to argue for rationalism using data about what competent speakers say and think about morality, it is incumbent on me to add the needed complexity to account for Frege's Puzzle. I'll do this with a discussion of the two main approaches to the puzzle, Fregeanism and Millianism. I'll offer some suggestions about how each approach should treat judgments about moral properties such as rightness and wrongness. Many of the general points will be familiar; the innovations lie in my explanations of how to connect the views up with internalism.[18]

17. Michael Smith's *The Moral Problem* argues for a particular analysis of moral judgments precisely by suggesting that it would make sense of rational changes in motivation when people change their moral beliefs. I think that argument is unsuccessful for moving too quickly over the complications I'm trying to emphasize in this essay.

18. The general phenomenon that competent speakers and thinkers can be ignorant of identities is widely noted in discussions of the Open Question Argument. See David Lewis, "Dispositional Theories of Value," *Proceedings of the Aristotelian Society* 63, suppl. (1989): 113–37, 125ff.; Smith, *The Moral Problem*, 35–39; and Mark Kalderon, "Open Ques-

It should be uncontroversial that two terms can designate the very same thing, even while a competent speaker is unaware that they do. A competent speaker can use 'Cicero' and 'Tully' and yet not know that they designate the same person. What is controversial is how to accommodate this fact within a theory of meaning for the relevant terms. Millians about names will want to treat the terms 'Cicero' and 'Tully' as having the same meaning or semantic value which, along with some auxiliary assumptions about the contents of beliefs, will lead to the result that a competent speaker can without irrationality believe contradictory propositions. The general Millian picture is this. The sentences 'Cicero was a Roman' and 'Tully was a Roman' express the very same content because each of their constituents have the same meanings and contribute the same semantic values to what is asserted by the whole. Thus, if propositions are whatever assertive utterances express, the two sentences express the same proposition. Furthermore, if what I say when I say "Cicero was a Roman," is just what I believe when I believe that Cicero was a Roman, and similarly for other such beliefs, this belief will have the same content as the belief that Tully was a Roman. Thus someone who believes the former but also believes that Tully was not a Roman has contradictory beliefs.[19]

Fregeans, on the other hand, will want to postulate senses or modes of presentation as constituents of the meanings of the relevant designating expressions so as to explain how a rational person can treat the sentences in question as differing in truth value. Roughly speaking, a sense is a way of picking out what the term designates, and the same thing can often be picked out in multiple ways. If different terms are associated with different senses, the terms will contribute different constituents to what is expressed in using them, and the corresponding beliefs will be different even though the objects designated by the terms may be the same. On this way of proceeding the beliefs of someone who accepts a judgment expressed using one term and disbelieves what is expressed by substituting a co-designating expression for that term need not be contradictory. This is because the propositional attitudes will be toward different propositional objects or involve different constituents provided by the senses corresponding to the terms. Even so,

tions and the Manifest Image," *Philosophy and Phenomenological Research* 68 (2004): 251–89. Related to internalism, Sigrun Svavarsdóttir's "Moral Cognitivism and Motivation," *Philosophical Review* 108 (1999): 161–219, postulates motivation connected senses but attacks internalism. Ralph Wedgwood, in "The Metaethicists' Mistake," *Philosophical Perspectives* 18 (2004): 405–26, postulates Fregean senses that are essentially action guiding; and van Roojen, "Knowing Enough to Disagree," gestures at a more Millian story.

19. See Kalderon, "Open Questions"; Scott Soames, *Beyond Rigidity* (Oxford: Oxford University Press, 2002); Michael Thau, *Consciousness and Cognition* (Oxford: Oxford University Press, 2002); and Nathan Salmon, *Frege's Puzzle* (Atascadero, CA: Ridgeview, 1991).

given that the terms designate the same object, it will be true that the two beliefs cannot in fact differ in truth value.

Each of these treatments can be extended from names to predicative expressions. Two predicates might designate the same property, and yet competent speakers may be unaware of this. And so it might be for our target moral predicates. We can say of a certain action that it "is right." We can say of an action that it "is rational." When we say the former we are predicating a property, rightness, of the action. When we say the latter we are predicating the property of being rational of the action. The rationalist proposal is that these are in fact the very same property and that the term 'right' designates the same property as 'rational' or something very similar. Given this analysis, it is open to the rationalist to treat either or both predicates as the Millian treats names, that is, as contributing just the property to what is expressed, or as the Fregean treats them, as contributing a mode of presentation of the property to the proposition expressed.

While my inclination is to think the Millian view correct, at least for the term 'right',[20] my main point relies only on accepting what the two views share in common—that competent speakers can be unaware that co-designative property terms pick out the same property. I think this possibility is open even where the fact that the two terms are co-designative is secured by a priori philosophical argument. For we can find examples fitting the pattern exemplified by the Cicero and Tully example even when what seem to be two properties are necessarily coextensive and even identical, and where this can be shown a priori. Just as one can be ignorant of an empirically confirmed identity because one has not made the necessary investigations, one can be ignorant of an a priori accessible identity if one has not gone through the relevant reasoning processes. Earlier in this essay I argued that ignorance of a fact need not by itself constitute irrationality. For uncovering the fact might depend on some process of discovery that one is not irrational for not having undertaken. This remains true when the matter to be discovered is the identity of seemingly distinct objects or properties and even when the process is one that leads to a priori arguments for those identities. Fregeans should want to capture this by postulating distinct senses, whereas Millians will allow that beliefs with contradictory contents need involve no rational failure on the part of their possessor.

That the conflicting beliefs involve no irrationality has conse-

20. My worry about Fregean treatments of these terms is partly that there seems to be no specific way of thinking about the property capable of uniquely determining the designatum which is also systematically associated with the same words across different speakers. Yet we seem justified in attributing the relevant beliefs based on sincere avowals using the terms, other things equal.

quences for the rational assessment of action as well. How it is rational to act depends on what one believes. The belief that a number is not prime may, for example, rationalize trying to factor it or rationalize asking the nearest mathematician what its factors are. The belief that a number is prime would normally rationalize not doing any such thing. Where a person has beliefs of both sorts, either in the sense-mediated way the Fregean postulates, or in the way Millians favor, we would not be surprised to find a rational person attempting to find factors for the number picked out by an Arabic numeral but refusing to do so when it is picked out with 'the 2001st prime'. Something analogous can be the case for the property that is designated by 'right' and also (if the analysis is correct) by 'rational'. This point will be important later on to explaining one way in which rational people can remain unmotivated by their moral beliefs.

Given these considerations advocates of rationalism can go on in either of two ways. Fregean rationalists can hold that a thinker or speaker may employ distinct senses or modes of presentation when thinking of rightness. These distinct modes can explain how a thinker can believe that something is the right thing to do, while doubting that it is the rational thing to do (or vice versa) even while the properties rightness and rationalness are identical. Alternatively, Millian rationalists should say that people can rationally believe two thoughts which are strictly speaking inconsistent. On the one hand, they can believe that an action is right while, on the other hand, believing that it is not rational. Because the 'right' and 'rational' contribute the same semantic values to the thoughts expressed using the terms, the speakers will thereby be thinking a thought and its negation. But because competent speakers may not be in a position to know this they may nonetheless rationally accept both claims. Neither way of proceeding will commit the theorist to ruling such thinkers and speakers irrational.

RATIONALISM AND THE TWO INTERNALISMS IN LIGHT OF THESE COMPLICATIONS

With these materials, the multiple relativized notions of rationality and either of the methods of accommodating the account of moral property terms to Frege's Puzzle, we can begin our explanation of internalism friendly rational amoralism. Our first task is to clarify the rationalist thesis given the multiplicity of kinds of rationality. Which of these kinds should a rationalist use when she reduces moral facts to facts about rationality? My claim is that she should use all or almost all of them.

A rationalist should say that morally right actions are those actions which a rational person would choose in a given circumstance. But a rationalist should not have to choose between identifying rightness with what a fully rational person with full information would do and what a

fully rational person with the agent's limited information would do. A rationalist can have it both ways so long as she is clear about what she intends to say. And similarly for the other limitations that an agent might be under. She can identify one sort of objectively right action with what would be chosen by a fully rational person under conditions of full information, and she can identify different sorts of subjectively right choices with what should be chosen in conditions that depart from the ideal. Since there are several ways that one's situation might be limited (information, time, etc.), there might be several subjectively right options corresponding to different limitations.

One might doubt that the most objective notion of rationality is something that a rationalist is entitled to employ. There are two concerns. One is that a theory which adds full information to the requirements of rationality to generate moral obligations is not really entitled to bill itself as a form of rationalism; one can be rational and not have full information, as I have emphasized. The other concern has to do not with the label but with the content of the rationalist analysis. If a rationalist reduces objective moral obligations to what a rational agent would do if she had full information, there is a good chance that the analysis will be circular if full information requires information about right and wrong. Those are just the concepts the rationalist was trying to analyze.

The first worry strikes me as merely terminological. No one ever thought that empirical information was irrelevant to morality. That includes rationalists. Even Kant, who to my mind underappreciated the relevance of empirical information, would admit the need for empirical information in determining which particular action is right or wrong. More importantly, nothing of philosophical interest can turn on a terminological objection like this. Give up the term, and the substantive issue—whether we can reduce moral facts to truths about what makes sense to do given certain sorts of information—will remain. 'Rationalism' is the term currently in use for the thesis that we can, but that is not a philosophical claim.

The worry about circularity can be handled by constructing the rationalist analysis to avoid it. When we first explain that the right thing is what a rational person would do given certain information we can be careful not to include information about what is right and wrong in that information. Or at least we can start that way and build up from there. We may need to proceed in stages because there may be second-order truths about what is right to do given that some other thing is right or wrong to do. For example, it may be right to discourage others from doing wrong. Thus if capital punishment is wrong, working to end it would be right. If we think there is a norm of rationality requiring opposition to what is wrong, we think that a rational person who knows

that capital punishment is wrong would oppose it. So long as the initial judgment that capital punishment is wrong does not require independent knowledge of the moral fact that it should be opposed, we introduce no circularity by allowing the rationality of further actions to depend on knowledge of moral status of capital punishment.

In any case the philosophical point is that by relativizing the rationality of various actions and attitudes to various limits, we end up with both a fully objective sense of rationality and with many differing subjective senses of rationality. And we can use each one. If a rationalist equates rightness with the property had by those things that are rational to choose, she can generate different notions of rightness, objective and subjective, mirroring each of the senses of rationality we might have an interest in.[21] One advantage to this approach is that it fits many of our actual judgments about what is right and what is wrong. Sometimes we say it is right to do some action where that claim can only make sense relative to some feature of the agent that is less than ideal. At other times we make judgments about rightness that can only be interpreted as a claim about the ideal. The approach allows both and can allow features of the context to disambiguate which is meant in that context.[22]

21. Someone may worry this generates too many different relativized notions of rightness and this would make it hard to explain how we can say what we want to say and know what we are saying. But there is no more problem here than there is for the corresponding notions of rationality which are similarly relativized. In ordinary talk we have little trouble figuring out what claim is at stake in a given conversational context. If, in a context where time is short, I say it is rational for us to just arbitrarily pick, the salience of the shortness of time plus the fact that given more time we would be able to make an optimal choice makes it obvious that the sort of rationality I have in mind is one relativized to the time we've got. If I say that it is right to pick arbitrarily, the same features of the conversational context narrow the choice of interpretations in the same way.

22. Talk about different senses of a term can be sloppy. I want to retain that sloppiness here because I think that much of what I say remains true no matter how we fill out the details. Consistent with the idea that 'is rational' can capture any one of the more subjective or objective "senses" of rationality there are a number of different semantic proposals for explaining exactly how we can get across which sort of rationality is at stake. On one approach the terms in question have one core sense which is their literal meaning, either on the fully objective or the fully subjective end of the spectrum. Still, various features of the context could allow us to imply things we don't literally say by employing Gricean mechanisms to make clear that we are trying to communicate related facts about one of the other notions of rationality. Another approach would suggest that 'is rational' has a literal meaning which contributes an incomplete relational property to an utterance. The information so conveyed is then completed (again via various Gricean mechanisms) by the context. For example, if due to features of the context it is unlikely that we meant to convey that an action is rational given the actual facts, listeners will interpret the intended claim as relative only to what the agent could take account of. Still further views might treat the term like an indexical with features of the context determining which of several candidate contents is literally expressed. For relevant discussion regarding options in other domains of discourse, see Kent Bach, "Conversational Impliciture," *Mind and Language* 9

This is all to the good. But a second reason to like the proposal is what it enables us to say about some of the puzzling cases of immorality and amorality that began the essay.

EXPLAINING EXISTENCE INTERNALISM CONSISTENT WITH RATIONAL WRONGDOING

I've already explained that existence internalism is a consequence of rationalism. If the truths of morality just are facts about what we are rationally required to do, then we will necessarily have a reason to do what is morally right. That's part and parcel of the rationalist project of reducing moral truths to truths of rationality. Having introduced multiple relativized notions of rationality and equated them with corresponding notions of rightness, the same line of reasoning vindicates existence internalism for each of these kinds of rightness. If (for example) relative to an agent's evidence it is right for her to intervene in a dispute, then relative to that same evidence it is also rational for her to intervene in that dispute.

The fact that we can make such judgments relative to different features of the agent's situation can now be invoked to explain how it might be that a person might, without irrationality, do what is in fact wrong. The general idea is to account for various kinds of rational immorality by noting that judgments of irrationality are usually or often made relative to one of the subjective senses of rationality. People who do what is objectively wrong will not be counted as irrational in one good sense so long as what they did made sense relative to the information that they have. Thus there is a sense in which those who do what is objectively wrong can still be rational though in one of the subjective senses.

One sort of rational immorality which a rationalist should have no trouble admitting involves actions which are rational because the agent lacks certain empirical information which would, if available, have changed what made sense to do. Clearly such agents are not subjectively irrational; they are doing what makes sense given the evidence they have. But this result is compatible with the chosen action being irrational relative to fuller information that the agent might have possessed. By equating what is objectively morally right with what is objectively rational in light of full information, we can truly say of such cases that the agent did something objectively morally wrong, but rational given what she knew.

This response might seem useless for other sorts of rational im-

---

(1994): 124–62; Jason Stanley and Zoltan Szabó, "On Quantifier Domain Restriction," *Mind and Language* 15 (2000): 219–61; and Paul Grice, *Studies in the Way of Words* (Cambridge, MA: Harvard University Press, 1989).

morality. Can't we imagine a person with full nonmoral empirical information who still does something morally wrong but is not irrational for doing so? Doesn't history provide us with just such examples? In answer to this we can extend the previous answer. Our access to even a priori information can depend on our actual history of reasoning. If we have not gone through the relevant reasoning processes, we can lack that knowledge and yet be rational. Even so, such knowledge can be rationally privileged insofar as someone who had gone through the right sort of reasoning would have that knowledge and because, in certain circumstances, that process of reasoning would itself be rationally required. This idea can be used to explain several additional sorts of seeming rational immorality. One sort of case is simply where the reasoning leading to the conclusion that some action is right is complex and time consuming. One need not be irrational if one does not do an action that would take one a long time to determine one should do.

This strategy of response may cover those such as Aristotle, who are in societies with abhorrent moral views. The process of reasoning which would lead to the rejection of these views requires, among other things, consideration of alternative ways of life which might not be obvious to people without appropriate acquaintance with other cultures. Even when there is an a priori argument in favor of a certain sort of hypothesis, it might take some experience before one is likely to consider the hypothesis and look for reasons for it. And if moral hypotheses can be justified by inference to the best explanation, even if a priori argument shows a competing moral hypothesis to be a better explanation, our consideration of that hypothesis might itself require imagination, luck, or experience. If those in the societies in question hold their views for reasons such as this, a rationalist need not regard their immorality as irrational.

EXPLAINING JUDGMENT INTERNALISM CONSISTENT WITH THE AMORALIST

Judgment internalism is a substantive claim over and above the rationalist thesis, so its explanation will be more complicated than the explanation of existence internalism. I argued earlier that the particular connection to be explained is that necessarily normally a rational person who believes an action all things considered right will be motivated to intend or try to do it.

An overly simple but instructive way to see the beginnings of an explanation is to think about whether it could be rational to act in a way one thinks irrational. It seems not, and this provides the core of our explanation. On the rationalist analysis the belief that an action is morally right is a priori equivalent to the belief that the action is rational. Or, in the words I used earlier to explain what I meant by 'rational', it

is a priori equivalent to the belief that the action makes sense to do. To say it is the only right action and hence morally required entails that it is the only action that in those circumstances makes sense to do. So it looks like someone who acts contrary to her moral judgments must be irrational in just the way that someone who acts contrary to her judgments of rationality is irrational. Both Frege's Puzzle and the relativization of rightness and rationality argued for above complicate this simple explanation. And that complication is necessary to explain how someone like Brink's amoralist or Huckleberry Finn is possible.

FACTORING IN RELATIVIZED RIGHTNESS

The belief that an action is rational to do might have the content that the action is objectively or ideally rational—that it is what a fully rational agent with no rational flaws, all relevant evidence, sufficient time to deliberate, and so on, would do. Or it might have the content that the action is rational in one of the various subjective senses relativized to evidence, limited time to deliberate, one's actual rational limits, and so on. It is only irrational to act in a way that is irrational given one's actual condition; it is not irrational to act in a way that would be irrational if one were in some other set of circumstances. Thus the belief that an action is objectively right, in a sense not relative to the actual features of the agent's situation, need not have the tight link with motivation that the simple explanation above relies upon. Only more subjective senses must have such a tight connection.

This may seem puzzling if you think only about the rationality of belief. Normally if you think belief in some claim justified, you think that if your evidence were more ideal than it currently is it would remain justified. That is because you think your current evidence justifies the claim that the thing you believe is true and that if the claim is true ideal evidence relevant to its truth will show that it is. If one believes it is rational given one's evidence to hold some belief (say that quantum mechanics is true) one should also believe that one has good evidence to think that one would hold that belief given full evidence. Thus one should only believe it rational relative to one's current evidence to accept the truth of quantum mechanics if one also believes it would be rational in a fully objective sense to do so.

The rationality of actions and intentions does not sustain as tight a connection between what is rational in actual circumstances and what would be rational under ideal conditions. One sort of example involves the interaction of epistemic limitations with one's goals; another involves practical limitations of agency. The epistemic cases involve actions under uncertainty. If I don't have conclusive evidence for some claim it may not be rational to act as if the claim is true, even while I think that if I had all the relevant evidence I should act that way. It may be rational

to spread my investments around rather than put them all in one place, even when I think that under conditions of ideal information I would know that Firm X would be the best investment and that in those circumstances it would be rational to invest all of my assets in Firm X. Under conditions of uncertainty it may be better to maximize one's expected benefit or to avoid risk.[23]

Practical nonepistemic limits on a person's agency may also underwrite divergence between what would be rational in ideal circumstances and what is rational in one's actual circumstances. For example, if an agent has a certain disposition to irrationality it can be rational for that agent to avoid situations in which that disposition will manifest itself even though it would not be rational for a fully rational agent without such a disposition to act in the same way. And we can recognize this from a first-person perspective. Given what I know about myself it is wise for me not to stock the freezer with ice cream, though were I more rational and therefore more resolute it would actually be economical and hence rational.[24]

These kinds of cases demonstrate that what an ideally rational or informed agent would do and what it is rational for an actual agent in actual circumstances to do can be very different. This complication is in fact an advantage for a rationalist. We can use it to accommodate certain exceptions to the idea that moral judgments must motivate rational agents. It will allow us to claim that an agent can rationally choose to act in ways that a perfectly moral agent would not choose and to show how this is compatible with reducing morality to rationality. This in turn will explain another way in which a person can sincerely admit some action right while failing to be moved to do it. A person can know that an action is objectively right—that if she were completely rational she would in fact do it. But she might also know that she has some sort of rational failing, such as weakness of will or a bad temper that she should take into account in deciding what to do in her actual circumstances. The upshot is that it is not always subjectively irrational to be unmotivated by one's judgments that an action is objectively right in the sense of objective rightness corresponding to full objective rationality.

On the other hand, it would show insincerity or irrationality to claim that in my actual situation, as I actually am, it makes most sense to apologize and yet not try to apologize. The same holds with respect to the corresponding notion of rightness (ignoring Frege's Puzzle for

23. Frank Jackson, "Decision-Theoretic Consequentialism and the Nearest and Dearest Objection," *Ethics* 101 (1991): 461–82, 462–63.

24. This ice cream example is parallel to that of the ill-tempered squash player (borrowed from Watson) in Michael Smith, "Internal Reason," *Philosophy and Phenomenological Research* 55 (1995): 109–31.

the moment). If a person claims it is right to apologize in this more subjective relativized sense and yet is not moved to apologize, we might rightly question either her sincerity or her rationality, or both.

There should be no worry that this undermines one of the main motivations for rationalism—that rationalism is well placed to secure the rational authority of morality and to explain why we rightly criticize those who act immorally. Rationalism continues to identify objective moral rightness with objective rationality and to recommend this as a rational end for agents, except in special cases. It just recognizes that sensible ways to aim at that goal are partly a function of facts about the agent's subjective situation. An agent is only rarely in a position that requires choosing between doing what makes sense to do in light of her evidence, time, and so on and doing what she thinks would make sense to do if she had full information. In normal cases these questions collapse into one another from the first-person agential perspective; the agent is trying to do what she thinks is objectively right *by* doing what her subjective situation suggests it makes sense to do given that goal and her evidence, time, and so on.[25]

FACTORING IN FREGE'S PUZZLE

The considerations I rehearsed earlier about Frege's Puzzle generate a further way that a person could remain unmoved, even if rationalism is true and even if she sincerely judged an action right for a person like her in her actual circumstances. While this complicates the kind of judgment internalism rationalism will underwrite, the complication will again be welcome. It will enable us to explain the sorts of amoralists highlighted by Brink and Rosen. Since these sorts of amoralists were the primary motivation for accepting the normalized internalist claim embodied by moderate internalism, the complication will cause our rationalist explanation of judgment internalism to match what needs to be explained.

If rationalism is correct, a rational person can have a thought that something is right and yet remain unmoved, but only provided that she is not in a position to recognize the identity of rightness with the prop-

25. So while the abnormal cases are important because they block us from saying that it is always subjectively rational to intend to do what is objectively right, they are unusual and exceptional. In an overwhelming majority of situations subjectively rational agents will and should intend to do what is objectively right. And this gives objective rightness its rational authority and grounds rational criticism of agents who don't meet this requirement. We should not want more than this because we want to give the exceptions their rational due as well. We morally criticize agents who risk catastrophic outcomes for small gains that they believe to be objectively available, and we therefore need to carve out just the space the theory predicts in the subjective rational authority of objective rightness.

erty of being rationally required. If she does or should recognize the identity then she also should rationally be motivated. While some amoralists may be irrational either because they recognize the identity but fail to act accordingly, or because they should but don't recognize the identity, some may fail to recognize it through no rational fault of their own.

In fact the most plausible examples of such people do seem to be ignorant in just that way. Thrasymachus is most plausibly thought of as challenging the rationality of moral action.[26] He denies the identity that would otherwise rationalize acting on the relevant moral beliefs. Whether he was rational to be in this state of mind depends on his evidence, but we should allow that a person could have evidence for views like his. Fregeans and Millians will describe Thrasymachus's state of mind somewhat differently. Fregeans will say that the thought that an action is morally right and the thought that an action is practically rational have different contents even if these turn out to predicate the very same property of the actions. For according to Fregeans, the contents of these thoughts will include the ways in which the property is picked out, and this way will be different when a person thinks of the property as moral rightness as opposed to the way involved in thinking of it as practical rationality. Millians will think that the two thoughts have the same contents but that Thrasymachus does not know that they do.[27] Despite having different ways of characterizing the contents of

26. You might think that there is an important difference between the rational status of ignorance of an identity and the rational status of denying an identity. But it is also true that the rationality of a state of mind depends on one's total evidence. So if it can be rational to be ignorant of an identity, and if there can be evidence that the referents of two terms are not identical, we should be able to construct situations where doubt and denial are rational by adding this evidence to a case in which ignorance is rational.

27. Jeffrey King's *The Nature and Structure of Content* (Oxford: Oxford University Press, 2007) suggests that philosophical analyses can be captured by biconditionals relating structured propositions. His account is motivated partly by its ability to avoid the paradox of analysis. Biconditionals can be informative because the propositions represented on either side of the 'iff' either have different structures or different constituents at different points within those structures. It can be informative to find out that something is a vixen just in case it is a female fox because the proposition that something is a female fox is a structured proposition with a component of that structure representing a complex whereas the proposition that the same thing is a vixen has a simpler structure with vixenhood as the corresponding component. If the right version of rationalism reduces the property of being morally right to the property of being rationally required based on certain grounds, rationalism can just take King's ideas on board. But there may be no principled way to divide moral from nonmoral grounds so that the best overall account will reduce moral rightness to rational rightness simpliciter. We would then still have Frege Puzzles but no difference in structure or the components to do the work. Millians have to treat such cases in whatever way they treat the analogous story about groundhogs and woodchucks.

Thrasymachus's thought, both accounts describe states of mind that he could subjectively rationally be in depending on his evidence.

Why then think that we have any explanation of judgment internalism? Why doesn't the account allow no connection between sincere moral judgments and motivation rather than the normal connection postulated by moderate internalism? Couldn't everyone be unaware of the identity in question and hence not moved by their judgments of what is right? No they could not, for if no one was so moved their judgments would not be judgments about what is right. The Fregean version of this story will suggest that you can have a moral thought with this content only in a way that depends (perhaps through deference) on people who take it to be action guiding. The Millian can say that thinking a thought with this particular content depends on being in touch with others who take the rightness of an option to rationally permit or require one to act rightly. Thus on either approach when a thinker thinks an action right, this depends on her being appropriately related to normal cases in which the rightness of the action motivates rational agents. You might, however, want to know a few more details about the resulting internalist requirement. A full answer will require me to say more about what the Fregean Rationalist and the Millian Rationalist each have to say. I'll start with an analogy that both sorts of rationalist will want to endorse before moving on to discuss each position separately.

AN ANALOGY

Burge tells us that 'arthritis' refers to a distinctively painful disease of the joints that could not be had in a location that is not a joint. This seems to be purely a matter of the practices of the relevant experts regarding arthritis—medical doctors. They could have decided to apply the term 'arthritis' also to similar pains occurring elsewhere but they did not. When we find speakers who express their beliefs by saying, "I have arthritis in my thigh," we attribute to each the belief that they have arthritis in their thigh. At least we do if their community is one in which the relevant experts have defined the term in the way that ours have. Had these speakers instead been members of a community that had allowed the term to extend to pains not in the joints, we would have attributed a different thought to them, not about arthritis but about some other similar ailment that extends also to the thighs.[28]

28. Tyler Burge, "Individualism and the Mental," *Midwest Studies in Philosophy* 4 (1979): 73–121. Certain disanalogies make the example imperfect; these include the role of medical expertise in the example and the fact that it is hard to imagine a society without a moral term, whereas arthritis terms seem optional. Since the example is the most well known of the sort, it has compensating expository advantages.

I think there is a marked similarity between this example and the amoralist example as a rationalist must conceive of it. In each case we are willing to attribute a thought the truth conditions of which would seem to entail that the speaker is expressing something ruled out by the correct analysis of the terms used to express the belief. And in each case we are willing to do so only against a background in which the most competent speakers would not avow those attitudes, and in which those other people are members of the first speaker's language community. If the amoralist were isolated from communities in which the term 'right' was used to commend we would not have attributed a thought about rightness to her, just as we would not have attributed a thought about arthritis to the medically ignorant patient in a community where doctors did not use the term 'arthritis' to pick out exclusively a disease of the joints.

Because of the role of experts in the arthritis example the analogy is not perfect. Medical doctors are uncommon, so while they are normal in the sense that their beliefs are normative for the extension of arthritis, they aren't statistically normal. But there is no principled reason why for some terms the people who determine the meanings of an expression could not be lay people, or most rational people, or some other group in a position to have a thought about a thing or property due to their experience of the referent. For example, it seems plausible to me that this is just how color terms operate. The term 'red' picks out what it does because of how it is used by normal sighted people. Yet color-blind people can use the term and refer to the same color, even though they can only do so in a speech community where those who determine what the word means are not color-blind. I think that a person could believe that chartreuse was a shade of red but that this could not be the normal case. The referent of color terms is determined by the practices of most sighted people in the language community. Competent speakers can run afoul of those practices and yet still possess the relevant concepts and attitudes constituted by them, but only in virtue of the background conditions involving the normal speakers.

So here we have a certain sort of necessary connection between the attitudes of normal speakers in a community but of a sort that does not require that all members of that community share the attitudes. The explanation is that the designatum of a speaker's terms can depend on the practices of the community in which she is a member, and the content of her thoughts expressed using those terms can depend on the same facts about the same community. She may flout the norms of her community and yet harbor thoughts which are partly constituted by the very norms she flouts.

Millians and Fregeans can agree that the actions and practices of most normal speakers in treating rightness as sufficient for rationalizing

and justifying an action make it the case that 'rightness' designates the same property as 'practically rationally permitted'. It is these actions by normal speakers in the community that make it the case that the attitude expressed by a speaker of that community by calling an action "right" attributes a property which a thing has only if it makes the most sense to do of the available options.

But Millians and Fregeans will disagree on the contents of the thoughts involved. Fregeans want the senses, reflecting the different ways in which speakers think of the referents of their terms, to both determine the referents of the relevant terms and to be constituents of the contents of the thoughts in question. Thus the natural move for them is to build deference to the normal speakers in the community into the contents of the thoughts in question.[29] They take themselves to be referring to what the normal speakers in the community are referring to, and the norms of these community members get to determine the referents of their thoughts because they are so referenced in their contents. In this way a Fregean rationalist can explain the necessity of a normal connection between thinking a thing right and being motivated while allowing that some rational individuals may not be so motivated.

Millian rationalists have both an easier and a harder time. Their task is easier insofar as they need not suggest that any particular way of thinking of the property in question is necessary to entertaining thoughts about either rationality or morality. Thus they can explain a lack of motivation. But it then becomes more difficult to explain why the view generates an explanation of internalism. To begin with the easier task, Millians will say that the two thoughts predicate precisely the same property of the action in question and that the property itself just is the semantic value of the relevant terms. A person can have two thoughts with identical content without recognizing that this is so. Thus entertaining a content which when viewed in the right way would rationalize a certain response need not rationalize that response if one does not recognize this content as one capable of doing so. Motivational responses are no different from other rational responses in this respect.

For the harder task of generating any substantial internalist commitment, Millian rationalists can explain the normal case internalism offered by moderate internalism. It is part of the Millian framework that speakers can designate the same item with a term as others in their

29. This is not quite what Ralph Wedgwood proposes in "The Meaning of 'Ought'," in *Oxford Studies in Metaethics*, vol. 1, ed. R. Shafer-Landau (Oxford: Oxford University Press, 2006). Wedgwood requires that the term play the right action-guiding role in an individual's psychology. That makes it hard to explain the moderate internalism supported by the sadists example.

community in virtue of standing in the right sort of relation to use of a term by others in their community without knowing everything there is to know about that term. So if rightness is a property that actions have when they make sense to do, a community member can use 'right' to designate that property if they interact with others in their community who use it to express that knowledge, even when they themselves lack it.

In more detail, normal speakers in a community have practical concerns that cause them to think about what it makes sense to do. They wind up using 'right' to express their knowledge that certain things make sense. Given their practical orientation these normal speakers are generally motivated by these judgments. For them it would be irrational to think that some action was morally right to do and to remain un-motivated by that thought unless it was equally rational to remain un-motivated by the corresponding judgment of rationality because of its place on the scale from subjective to objective rationality.[30] Those in the community who don't use 'right' with this action-guiding purport de-pend on those who do insofar as they use the term to refer to this property.[31] Millian and Fregean rationalists can thus explain moderate internalism, and this means they can accept plausible varieties of rational amoralism.

WHAT WERE THEY THINKING?

A reader might be forgiven for wanting an answer to the question of what (consistent with the analysis) rational thinkers could be thinking while doubting their reasons to do right. So I should say something to make the thoughts seem plausible from the inside even though I suspect there is no general account of rational amoralist thoughts. Perhaps I can sketch the situation of a person who has such thoughts in such a way that those thoughts seem rational. One possibility is that a speaker might know rightness as a property which the members of his or her community took as reason-giving without believing that his or her peers were in fact right for doing so. In a society where people had a tendency to go in for reprehensible things, a person might believe that the prop-erty people labeled "right" supervenes on reprehensible features of ac-

30. Normal speakers are those who at least tacitly know the truth of the reduction. They are in a state of mind that would most naturally be expressed by asserting the sentence capturing the rationalist reduction.

31. Millians can add a further quasi-internalist constraint since they need not deny descriptive constraints on competence. Millians can require that all speakers who can think thoughts with moral content know there is some sort of connection with reasons. Insofar as the use of moral terms requires interaction with speakers who use these terms to make action-guiding judgments it is not implausible that competent speakers know about those so guided.

tions such that this property did not in fact have much to commend choice. If everyone around thinks support for slavery right and acts accordingly, a person might reasonably think these others don't know much about what is right. But a person might instead just as reasonably think their neighbors are right about the extension of rightness but wrong to do what is right.

A person in that community can use the communal words 'right' and 'wrong' to express the thoughts of Huckleberry Finn. Huck thought it was wrong to work to free slaves in general and Jim in particular. He believed himself generally uneducated and those around him to be better judges on all sorts of matters, including matters of morality. He believed he was going to pay a high cost for helping Jim to flee. But thoughts about morality and even worries about his own long-term post-death self-interest were insufficient to motivate him to act in accordance with his beliefs about the wrongness of helping slaves. It seems fair to say that he did not see them as providing sufficient reason to act in ways incompatible with his friendship with Jim. And it seems fair to say he saw his friendship, and perhaps also Jim's humanity, as reasons not to turn him in.

Huck's thoughts here seem perfectly rational. It was not irrational for him to be ignorant of the fact that those around him were no experts about morality or metaphysics. And given his deference on these matters, it was also not irrational for him to accept his neighbor's views about the extension of moral obligation. At the same time it is most reasonable to think that if morality requires returning slaves it does not make sense to do what morality requires. On this basis Huck could reasonably conclude that it was right to return Jim and that he had no sufficient reason to do so. Rationalism need not say otherwise. Millians and Fregeans will capture the exact content of Huck's thought that it makes sense not to do what morality requires in somewhat different ways, but both can accept this basic picture.

Huck's case is rich enough to be read differently by different people.[32] The important point for the present argument is that there are

---

32. How best to describe Huck's case is a complicated matter and bound to be controversial (especially given his beliefs about Hell which give him prudential reasons not to protect his friend). Nomy Arpaly, *Unprincipled Virtue* (Oxford: Oxford University Press, 2003), 76, for example, denies that Huck in any sense believes that it's right not to turn Jim in, and suggests that he only dimly recognizes himself as acting for reasons. My Millian sympathies disincline me to accept the first denial, at least to the extent that he sees himself as acting for reasons. But we both see him as acting for reasons, insofar as Jim's humanity and friendship are reasons to protect him. And these reasons justify him in acting as he does. Huck is giving them appropriate weight when he acts, even though he thinks that he is thereby violating the demands of morality (and self-interest) when he does so.

ways of constructing stories like this in which a person, in virtue of seeing the strong reasons against doing what his or her peers take morality to require, could act rationally to violate those norms even while reasonably thinking that his or her peers are right about the requirements of morality. Thus, if you are convinced of the possibility of rational amoralism or even immoralism in whichever examples seem most plausible to you, this should be no impediment to accepting a rationalist reduction, provided you can tell the kind of story I want to tell about Huck in the cases that move you.

What of Thrasymachus and Rosen's sadists? The examples are not filled in enough for me to have a clear view of the reasonableness of their views. With the sadists we aren't told much about the background morality they reject, so it is hard to say whether their rejection is rational. Conventional Greek morality in Thrasymachus's time wasn't the most attractive social institution. It might well be reasonable to have thought that if moral terms applied to what most Greeks thought they did, morality might be more accurately thought of as a hoax than the upshot of rational thought about what makes sense to go in for. Of course, Thrasymachus probably had less reason to defer to his peers about the content of morality than Huck. So he might have made a rational error in deferring to them. But again, the point is that if there are plausible ways to argue that his view was rational, we can explain that upshot by identifying the relevant sort of rationality as a species of subjective rationality and by using either Millian or Fregean accounts of the contents of his thoughts to explain why their possibility remains consistent with the truth of rationalism.

A WORRY AND A RESPONSE TO IT

You might worry that the reconciliation of rational amoralism with rationalism I propose has undermined one initial rationale for accepting rationalism. This rationale, in a nutshell, was that rationalism is the best explanation of the action-guiding role of morality. My proposal may seem to muck up the tight connection between moral judgment and action and perhaps make it harder to see why even the core group of "normal" people should be motivated by their moral judgments. The reconciliation may seem to make rational amoralism too common. And relatedly, as one reader suggests, "the more reasonable it turns out to be to fail to believe rationalism the less reasonable rationalism itself may seem."[33]

This worry is motivated by the thought that rationalism makes morality important and that if morality is important one would be subjectively irrational for having negligently failed to figure out that right

33. These words come from a referee for *Ethics.*

action is rationally required. The theory here actually agrees with one way of understanding this thought. In one of the more salient and useful subjective senses of rationality, one is subjectively irrational if one doesn't take the time one has to figure out what one morally ought to do when the stakes are high. But one can fulfill this obligation without accepting the sentence that expresses the analysis. I can assure myself that the action I take has the property that actions have when they make most sense to do without knowing that the words 'morally right' pick out that property.[34]

In answering the worry that the proffered account makes rational amoralism too common, it is important to recall where we began. The initial argument from action-guidingness to rationalism was already under some challenge to just the extent that morals/motives judgment internalism was under pressure from plausible examples of rational amoralism. Reasonable people can disagree over those they think plausible. I've chosen Huck Finn as my paradigm example because I think he's more plausible than the rest, but readers may differ. The resources I employ to make sense of Huck are general in one sense; limitations of time and background information and imagination can be used to make many errors rational. But whether in any given case they succeed in rationalizing an agent's behavior and judgments is going to depend on a more substantive conception of how these things affect the rationality of certain judgments and courses of conduct and then applying this conception to particular cases. It will be no surprise if one's pretheoretical judgments about the rationality of particular amoralists dovetail with one's ideas about where the general factors I'm using can be used to rationalize amoralism. Nothing in the proposal requires finding amoralism rational in cases where an agent has good reason to accept the action-guiding nature of morality. Rather it shows that the rationalist analysis can accommodate those instances of amoralism that meet plausible rational standards.

The present account does not make rational amoralism even as prevalent as doubts about rationalism. One can rationally disbelieve a true identity and yet believe that something close to it is true. And that state of mind will still rationalize acting as though the identity is true for a wide range of cases. For example, one can deny or doubt that being morally required is identical with making most sense to do, even while believing that generally what is morally required is what makes most sense to do. And the latter thought will rationalize much the same behavior as accepting the identity for most situations. So we should

34. Similarly, I can know that an old female fox is a vixen without knowing the analysis of 'vixen' because I don't know if very young foxes are vixens.

expect rational failure to act on moral judgments to be less widespread than the rational denial of rationalism, even while the theory is true.

Furthermore, in such cases the truth of the analysis and its recognition by some can explain why those others believe there is a close connection between morality and what it makes sense to do, even if they don't go so far as to accept the analysis. If the core group of people, those who consciously use moral terms to pick out rationalizing features of reality, have sensible views about what it makes sense to do, others will notice that the extension of their moral talk closely corresponds to what it makes sense to do. That recognition by these others counts as knowing something about morality even if it isn't knowing the analysis. So when they are acting in light of that knowledge they too will be for the most part guided by their moral judgments, whether they accept the analysis or not. And this means that the truth of rationalism can genuinely be part of the best explanation of such people being guided by morality, even while they don't accept rationalism.

Finally, it is not part of this story that one can come to accept rationalism only by complex philosophical argument. A good number of people find it perfectly natural to use moral terms to express their judgments about what makes sense to do. I think it is fair to say that they accept the identity whether or not they are philosophically sophisticated or have explicit thoughts about identity or analysis or composition. In fact since it seems all agents actually do have thoughts about what it makes sense to do, and since words like 'right' and 'wrong' are apt to express those thoughts, people like this are very common.

CONCLUSION

We have reason independent of any commitment to moral rationalism to distinguish various more or less subjective sorts of rationality. And we have independent reason to allow the phenomena highlighted by Frege's Puzzle to shape our account of the practical upshot of normative belief. When these insights are properly combined with a rationalist account of morality identifying moral requirements with a possibly improper subset of rational requirements, we can construct an account that explains two sorts of independently motivated internalist principle—one suggesting that moral requirements entail reasons to act as they suggest, and the other suggesting that genuine moral judgments must be such that they normally motivate rational agents who accept them. Furthermore the resulting view leaves room for rational agents to have mistaken views about morality and also for some such agents to be unmotivated by the judgments they accept. Rational amoralism and even rational immorality is consistent with moral rationalism.