On Three Alleged Theories of Rational Behavior

STUART RACHELS
University of Alabama

What behavior is rational? It's rational to act ethically, some think. Others endorse instrumentalism – it is rational to pursue one's goals. Still others say that acting rationally always involves promoting one's self-interest. Many philosophers have given each of these answers. But these answers don't really conflict; they aren't vying to describe some shared concept or to solve some mutually acknowledged problem. In so far as this is debated, it is a pseudo-debate. The different uses of 'rational action' differ merely in meaning. I shall defend the following claims: 'rational behavior' is used in ethical, prudential, and instrumental ways (section I); these uses of 'rational behavior' are distinct (section II); they do not represent competing theories of rational behavior (section III); we should stop using 'rational behavior' ethically and prudentially, but we may continue its instrumental use (section IV).

I. THE THREE MAIN USES OF 'RATIONAL BEHAVIOR'

In general, our use of 'rational' is unwieldy.[1] But I won't examine every way we use the term; I'll just discuss 'rational behavior' and 'rational action'. I won't discuss 'rational' as applied to beliefs, desires, emotions, deliberation, dispositions, contracts, persons, and so on. I'll use 'rational behavior' and 'rational action' interchangeably, since they differ trivially in meaning, if at all. I won't discuss the phrase 'rational choice', though it often means the same thing.[2]

In this section, I'll describe the three main ways philosophers use 'rational behavior'.[3] For now I wish to remain neutral on whether these uses reflect mere semantic variation or genuine disagreement about rationality. That is an issue for section III. However, biased language might creep in. For example, I might refer to an 'account' of rational behavior, which suggests that it competes against other accounts. In this section, ignore such insinuations, which are irrelevant to its taxonomic purpose.

[1] For an illuminating discussion of why philosophers overuse 'rational', see Jonas Olson, 'Reasons and the New Non-Naturalism' (unpublished).

[2] Actions can be conceived as resulting from choices, so 'rational action' does not always mean 'rational choice'. But even then, an act is considered rational just in case it is rationally chosen, so action and choice remain closely related.

[3] Ordinary speakers use 'rational behavior' in other ways. For example, acting rationally is often contrasted with acting from emotion or impulse. Philosophers talk this way too, but the three main uses get more currency.

© Cambridge University Press 2009
Utilitas Vol. 21, No. 4, December 2009
doi:10.1017/S0953820809990252

https://doi.org/10.1017/S0953820809990252 Published online by Cambridge University Press

Rational behavior as ethical behavior

'Rational' is closely linked in etymology and everyday use to 'reason'. Indeed, 'rational behavior' is sometimes used interchangeably with 'reasonable behavior'.[4] Hence, rational behavior is sometimes thought of as behavior best backed by reasons.[5] Typically those reasons are thought of as *subjective*, as propositions the agent ought to regard as providing practical reasons, on her evidence. Thus I may rationally opt for surgery if my evidence suggests I have a malignant tumor, even if I really don't.

On this use of 'rational', the rational action is 'the best thing to do' or 'what it makes sense to do' or 'what one ought to do', as Gibbard tells us.[6] These phrases can also describe ethical behavior. Thus, rational behavior is sometimes identified with ethical behavior.[7] This use of 'rational' is uncommon in economics,[8] but philosophers such as Sidgwick, Smart, Brandt, Frankena, Hare, Gauthier, Gibbard, Quinn, Sumner, and Scanlon have used 'rational' along these lines, sometimes tailoring it to their own ethical views.[9]

[4] But not always; see Alan Gewirth, 'Rationality vs. Reasonableness', *Encyclopedia of Ethics*, 2nd edn., vol. 3, ed. L. C. Becker and C. B. Becker (New York and London, 2001), pp. 1451–4. Also, Parfit's recent work treats the two terms differently (see D. Parfit, 'Climbing the Mountain', unpublished).

[5] Rational behavior is less often thought of as behavior permitted by reasons, such that one might be permitted to act with less than full rationality. In general, the literature is about fully, not partly, rational behavior.

[6] A. Gibbard, *Wise Choices, Apt Feelings* (Cambridge, Mass., 1990), p. 49.

[7] As Brandt says, 'Sometimes, "it is rational for X to" seems to be identified with "X ought to"...' (Richard B. Brandt, 'The Concept of Rational Action', *Rationality in Action*, ed. P. K. Moser (Cambridge, 1990), p. 398). Others might say that the sphere of the rational includes the ethical but is broader. Making a valid deduction, for example, might be considered rational without being thought ethical.

[8] According to Sen, 'there are two predominant methods of defining rationality of behavior in mainline economic theory. One is to see rationality as internal consistency of choice, and the other is to identify rationality with maximization of self-interest' (A. Sen, *Ethics & Economics* (Oxford, 1987), p. 12).

[9] In attributing the ethical view to Sidgwick, I rely on the impressions of Parfit, Frankena and Richardson (D. Parfit, *Reasons and Persons* (Oxford, 1984), p. 129; William K. Frankena, 'Sidgwick and the History of Ethical Dualism', *Essays on Henry Sidgwick*, ed. B. Schultz (Cambridge, 1992), pp. 175–98, at p. 194; and Henry S. Richardson, 'Commensurability as a Prerequisite of Rational Choice', *History of Philosophy Quarterly*, 8.2 (April 1991), pp. 181–97, at pp. 186–7). Unfortunately I know of no passage in which Sidgwick obviously uses 'rational' interchangeably with 'ethical'. But Sidgwick does define 'unreasonable action' as 'voluntary action contrary to a man's deliberate judgment as to what is right or best for him to do' (Henry Sidgwick, 'Unreasonable Action', *Practical Ethics: A Collection of Addresses and Essays* (Oxford, 1998), p. 132). And Sidgwick also calls this phenomenon 'subjective irrationality' (p. 134). See J. J. C. Smart (with Bernard Williams), 'An Outline of a System of Utilitarian Ethics', *Utilitarianism For and Against* (Cambridge, 1973), pp. 3–74, at pp. 46–7; R. Brandt, *A Theory of the Good and the Right* (Oxford, 1979), p. 10 and Brandt, 'The Concept of Rational Action', pp. 413–14; W. K. Frankena, *Thinking About Morality* (Ann Arbor, 1980), p. 85; R. M. Hare, *Moral Thinking* (Oxford, 1981), pp. 104–5 and p. 214; David Gauthier, 'Deterrence,

https://doi.org/10.1017/S0953820809990252 Published online by Cambridge University Press

Rational behavior as self-interested behavior

The word 'rational', Hare says, 'is sometimes used more or less synonymously with "prudent"'.[10] To act rationally, in this sense, is to act in one's self-interest. Or rather, it's to act in what seems like one's self-interest, given one's evidence. This use of 'rational' has deep roots. Sidgwick cites over a dozen historical figures who, like him, sympathize with identifying rational and self-interested behavior.[11] There is 'preponderant assent', Sidgwick says, to this view in 'the common sense of mankind' and in 'the history of ethical thought in England'.[12] And according to Parfit: 'It has been assumed, for more than two millennia, that it is irrational for anyone to do what he knows will be worse for himself.'[13] 'Rational' is also closely associated with self-interest among economists and social theorists. According to Sen, 'the self-interest interpretation of rationality . . . has been one of the central features of mainline economic theorizing for several centuries'.[14] And '[t]he concept of rationality familiar in social theory', Gauthier says, 'identifies rationality with the maximization of individual utility'.[15]

Maximization, and Rationality', *The Security Gamble*, ed. Douglas MacLean (Totowa, 1984), p. 115 and elsewhere; Gibbard as already cited in the text and on p. 49 of *Wise Choices*; Warren Quinn, *Morality and Action* (Cambridge, 1993), p. 254 (in Essay 12, 'Putting Rationality in its Place'); L. W. Sumner, 'Welfare, Preference, and Rationality', *Value, Welfare, and Morality*, ed. R. G. Frey and Christopher Morris (Cambridge, 1993), pp. 88–9; T. M. Scanlon, *What We Owe to Each Other* (Cambridge, Mass., 1998), pp. 30–2 (though Scanlon uses 'rational' and 'reasonable' in various subtly different ways).

[10] Hare, *Moral Thinking*, p. 190. For simplicity, I ignore Frankena's distinction between the rationality of pursuing one's happiness and the rationality of pursuing one's perfection. (See William K. Frankena, 'Concepts of Rational Action in the History of Ethics', *Social Theory and Practice* 9.2–3 (Summer–Fall 1983), pp. 165–97, at p. 168.)

[11] For a taste of Sidgwick's view, see *The Methods of Ethics*, 7th edn. (Indianapolis, 1906; repr. 1981), p. 498. Sidgwick's list of philosophers who sympathize with identifying rational and self-interested action includes the ancients, Spinoza, Hobbes, Cumberland, Shaftesbury, Samuel Clark, Butler, Berkeley, Hume, Kant, Paley, Stewart, Reid, Bentham, and Comte. On close inspection, Shaver disagrees with putting Butler, Hume, Kant, and Bentham on this list but adds John Clark (R. Shaver, *Rational egoism* (Cambridge, 1999), ch. 4). Sidgwick on ancient Greece: *The Methods of Ethics*, pp. 91–2; Spinoza: *The Methods of Ethics*, p. 89; Hobbes: *The Methods of Ethics*, pp. xix, 86, 89; Cumberland: 'Hedonism and Ultimate Good', *Mind* 2.5 (1877), pp. 27–38, at p. 30; Shaftesbury: *Outlines of the History of Ethics*, 5th edn. (Indianapolis, 1988, originally 1902), pp. 184–90; Samuel Clark: *The Methods of Ethics*, pp. 119–20; Butler: *The Methods of Ethics*, pp. 119–20 and pp. 205–6; Berkeley: *The Methods of Ethics* p. 120; Hume: *Outlines of the History of Ethics*, pp. 205–6n.; Kant: *Outlines of the History of Ethics*, p. 276; Paley: *The Methods of Ethics*, p. 121; Stewart: *Outlines of the History of Ethics*, p. 232n. 1; Reid: *Outlines of the History of Ethics*, p. 228; Bentham: *The Methods of Ethics*, p. 119; Comte: *Outlines of the History of Ethics*, pp. 268–9.

[12] Henry Sidgwick, 'Some Fundamental Ethical Controversies', *Mind*, 14.56 (1889), pp. 473–87, at p. 483.

[13] Parfit, *Reasons and Persons*, p. 130. This is overstated; see Pettit and Smith's claim below.

[14] Sen, *Ethics & Economics*, p. 15.

[15] D. Gauthier, *Moral Dealing: Contract, Ethics and Reason* (Ithaca, NY, 1990), p. 152. But see below: by 'the maximization of individual utility' Gauthier might mean, not

https://doi.org/10.1017/S0953820809990252 Published online by Cambridge University Press

Instrumental rationality

Instrumentalism is roughly the idea that I act rationally just in case I intelligently pursue my ends, whatever they happen to be. Hume, Ramsey, Russell, Savage, Hempel, Foot, Williams, Harsanyi, Simon, Harman, Fumerton, and Allingham, for example, have all used 'rational' or 'reasonable' instrumentally.[16] The instrumentalist intuition is the basis of decision theory, which, according to Pettit and Smith, has been the orthodox account of rationality for the last two hundred years.[17] The decision theory literature is vast, but I'll lay out its central thrust about the nature of rational behavior.

I act with instrumental rationality just in case I maximize expected utility, under a specific interpretation. In general, to calculate the expected utility of an action, multiply the value of each possible outcome by that outcome's probability, and sum the products. The maximizing act (if there is just one) is the possible act with the highest expected utility. So, to determine the instrumentally rational act, one must determine: (i) the possible acts; (ii) the possible outcomes of those acts; (iii) the probabilities of those outcomes obtaining; and (iv) the outcomes' values. How do instrumentalists do this? They say little about (i); there is no distinctively instrumentalist way to identify possible acts.[18] (ii)–(iv), however, merit discussion. I'll take them up in reverse order.

(iv) Instrumentalists measure value or utility in terms of the agent's preference-satisfaction. Thus, I act with instrumental rationality just in case I maximize my expected preferences. Instrumentalists typically interpret 'preference' broadly, to include desires, evaluations, and

the maximization of individual welfare, but the maximization of individual preference-satisfaction, which is different.

[16] See D. Hume, *A Treatise of Human Nature* (1739), bk. 2, pt. 3, sect. 3; Frank Ramsey, 'Truth and Probability', *The Foundations of Mathematics and Others Essays* (London, 1931), pp. 156–98; B. Russell, *Human Society in Ethics and Politics* (London, 1954), p. viii; L. J. Savage, *The Foundations of Statistics* (New York, 1954) and *The Foundations of Statistics*, 2nd edn. (New York, 1972); C. G. Hempel, *Aspects of Scientific Explanation* (New York, 1965), p. 463; Philippa Foot, 'Morality as a System of Hypothetical Imperatives', *Virtues and Vices and Other Essays in Moral Philosophy* (Berkeley, 1978), pp. 157–73, p. 162 (but Foot changed her view; see *Natural Goodness* (Oxford, 2001)); Bernard Williams, 'Internal and External Reasons', *Moral Luck* (Cambridge, 1981), pp. 101–13; John C. Harsanyi, 'Morality and the Theory of Rational Behavior', *Utilitarianism and Beyond*, ed. A. Sen and B. Williams (Cambridge, 1982), pp. 39–62, at p. 42; H. Simon, *Reason in Human Affairs* (Stanford, 1983), pp. 7–8; Gilbert Harman, 'Human Flourishing, Ethics, and Liberty', *Philosophy and Public Affairs* 12 (1983), pp. 307–22, at p. 320; R. A. Fumerton, *Reason and Morality: A Defense of the Egocentric Perspective* (Ithaca, NY, 1990), p. 155; M. Allingham, *Rational Choice* (Basingstoke, 1999), pp. 1–2.

[17] Philip Pettit and Michael Smith, 'Parfit's P', *Reading Parfit*, ed. Jonathan Dancy (Oxford, 1997), pp. 71–95, at pp. 83–4.

[18] Sobel, however, does talk about this. See J. H. Sobel, *Taking Chances* (Cambridge, 1994), ch. 10 (originally in *Theoria* 49 (1983), pp. 159–83), sect. 1.

https://doi.org/10.1017/S0953820809990252 Published online by Cambridge University Press

projects; and narrowly, to include only ultimate preferences, since those are what ultimately matter.[19] An ultimate preference is a preference for the thing itself. For example, I may ultimately want money, and not just what money can buy. So, I act with instrumental rationality by maximizing the expected satisfaction of my ultimate preferences, broadly conceived. But not every ultimate preference counts; only the coherent ones do. Decision theorists spend ink and tree debating the formal nature of coherence.

(iii) Instrumentalists differ on how to determine outcome probabilities. Thus they differ on (ii) how to determine the possible outcomes, since those are just outcomes with non-zero probability. According to some, we should simply plug in numbers representing the agent's beliefs, provided those beliefs are coherent.[20] Think of rationality, on this view, as acting rationally *given the agent's beliefs and desires*. This makes acting rationally rather easy. For example, I can act rationally based on my horoscope, if I happen to believe that its predictions are reliable. One might say, 'Given that he believes in that stuff, it makes sense for him to seek new business opportunities and to prepare to hear from an old friend.' According to others, the probabilities should match *the agent's evidence*, which is what the agent should believe, on the evidence available to her.[21] On that view, my horoscope-based action would not be rational, if my evidence recommends acting differently.[22] Different accounts of evidence spawn different instrumentalist theories.

In sum, I act with instrumental rationality just in case I maximize the expected satisfaction of my ultimate, coherent preferences, calculating outcome probability in terms of either my evidence or my coherent beliefs.

[19] See Robert Audi, 'Rationality and Valuation', *Rationality in Action*, ed. P. K. Moser (Cambridge, 1990), p. 422, and see Quinn, *Morality and Action*, p. 211 (in Essay 11, 'Rationality and the Human Good'). On whether instrumentalists should interpret preference in terms of belief, liking or valuing, see David J. Velleman, 'The Story of Rational Action', *Philosophical Topics* 21.1 (Spring 1993), pp. 229–54, at pp. 232–3. See J. E. Hampton, *The Authority of Reason* (Cambridge, 1998), ch. 5, for criticisms of instrumentalist accounts of preference.

[20] See, for example, Pettit and Smith, 'Parfit's P', p. 79, and Robert Audi, 'Rationality and Valuation', p. 418. It is my own assumption that instrumentalists require coherence from beliefs, just as they require coherence from desires.

[21] See, for example, Peter Gardenfors and Nils-Eric Sahlin, 'Introduction: Bayesian Decision Theory – Foundations and Problems', *Decision, Probability and Utility*, ed. P. Gardenfors and N. Sahlin (Cambridge, 1988), pp. 1–15, at p. 4.

[22] What should we say if I blindly follow my horoscope, and if, coincidentally, my evidence entails what my horoscope says? We might say that my action is rational, but I am not. Here I follow Philip Pettit and Michael Smith, 'Practical Unreason', *Mind* 102 (1993), p. 58.

https://doi.org/10.1017/S0953820809990252 Published online by Cambridge University Press

II. THE THREE USES OF 'RATIONAL BEHAVIOR' REALLY ARE DISTINCT

I have identified three uses of 'rational behavior'. On them, rational action is ethical, self-interested, and preference-maximizing. Are these uses of 'rational' really distinct? They are, in every relevant way.

First, compare the ethical and self-interested uses. These are distinct in the sense that we use them differently; we apply them to different actions. Murdering for profit, we think, can be rational in the self-interested sense but not in the ethical sense. Moreover, these uses *should* be applied to the same actions only if ethical egoism is true, since only then would ethical behavior always be self-interested. Here I'll just assume that ethical egoism is false.

Next, consider the ethical and instrumentalist uses. Instrumentalists hold that rational action maximizes the expected satisfaction of one's coherent preferences. But the coherence requirement for preferences rules out little in terms of content. Coherence is understood, not robustly, but merely as analogous to consistency – intransitive preferences, for example, are likened to inconsistent beliefs.[23] For this reason, many coherent preferences are blatantly unethical, and much instrumentally rational behavior is manifestly deplorable.[24] Thus, the ethical and instrumentalist uses are distinct.

Finally, compare the self-interested and instrumentalist uses of 'rational behavior'. Promoting one's interests and satisfying one's desires do seem closely related. Are the instrumentalist and prudential uses really the same, in some important way? They are not. First, using 'rational' to refer to self-interested behavior does not entail using it to refer to behavior that satisfies one's desires. The desire theory of self-interest faces strong criticism,[25] and there are well-known alternatives to it.[26] But more importantly, consider this example. Suppose I would maximize my expected preference-satisfaction by giving up my life for a friend, since my altruistic desires are so strong. Such a sacrifice would be rational, on the instrumentalist view. But need it be in my interest? Obviously not. Thus, the self-interested and instrumentalist uses are distinct, and here is why. In its best form, the preference theory of self-interest does not hold that my life goes best if my preferences in

[23] See, for example, Savage, *The Foundations of Statistics*, 1st edn., pp. 19–21.

[24] See, for example, Amartya Sen, 'Rationality and Uncertainty', *Recent Developments in the Foundations of Utility and Risk Theory*, ed. L. Daboni et al. (Dordrecht, 1986), p. 4.

[25] See, for example, W. H. Shaw, *Contemporary Ethics: Taking Account of Utilitarianism* (Oxford, 1999), pp. 52–6.

[26] As Parfit notes, aside from desire-fulfillment (or preference-satisfaction) theories of self-interest, there are also hedonistic theories, objective list theories, and hybrid theories. See Appendix I of *Reasons and Persons*.

https://doi.org/10.1017/S0953820809990252 Published online by Cambridge University Press

general are maximized. Rather, it holds that my life goes best if those preferences of mine *that concern my own life* are maximized.[27] This qualification is not in instrumentalism; instrumentalism is not about satisfying just those ultimate, coherent desires *that concern one's own life*. Thus, self-interested and instrumentalist rationality are distinct.[28]

The literature often suggests that the instrumentalist use of 'rational' is equivalent to one of the others, even though it isn't. Instead of saying that rational action maximizes the agent's expected preference-satisfaction, instrumentalists often just say that it maximizes expected utility – thus seeming to identify rational action with ethical action from a utilitarian standpoint. And instrumentalists often just say that the rational action maximizes individual utility – thus seeming to identify rational behavior with self-interested behavior. Gauthier admits making this second slip.[29] And Rawls slides between instrumentalist and self-interested uses in *A Theory of Justice*. On one page, he says, 'I have assumed throughout that the persons in the original position are rational. In choosing between principles each tries as best he can to advance his interests.'[30] This suggests rationality as self-interest. But on the next page Rawls gives 'rational' an instrumentalist interpretation:

> Thus in the usual way, a rational person is thought to have a coherent set of preferences between the options open to him. He ranks these options according to how well they further his purposes; he follows the plan which will satisfy more of his desires rather than less, and which has the greater chance of being successfully executed.[31]

There is no suggestion here that the agent's purposes and desires can't be altruistic and self-denying.

III. THE THREE USES DO NOT REFLECT RIVAL THEORIES OF RATIONAL BEHAVIOR

How are the three uses of 'rational action' related? On the leading view, they reflect substantive disagreement about what rational behavior is; they express rival theories of rational action.[32] However, it is not clear

[27] See Parfit, *Reasons and Persons*, p. 494.

[28] And so Eells should not equate well-informed instrumentalist action with action that 'is truly, objectively in the agent's best interest to perform' (E. Eells, *Rational Decision and Causality* (Cambridge, 1982), p. 5).

[29] See David Gauthier, 'Rational Constraint: Some Last Words', *Contractarianism and Rational Choice: Essays on David Gauthier's 'Morals by Agreement'*, ed. P. Vallentyne (Cambridge, 1991), pp. 323–30, at p. 323.

[30] J. Rawls, *A Theory of Justice* (Cambridge, Mass., 1971), p. 142.

[31] Rawls, *A Theory of Justice*, p. 143.

[32] This idea runs through a lot of work. See, for example, Parfit, *Reasons and Persons*, pt. 2 and ch. 14 (Parfit never implies that the leading uses of 'rational action' reflect

https://doi.org/10.1017/S0953820809990252 Published online by Cambridge University Press

what this disagreement is supposed to be about. I'll consider three attempts to pinpoint the alleged dispute, all inadequate.

1. Perhaps the different uses of 'rational behavior' reflect competing views about what our (real) normative reasons for action are – and thus about how we ultimately ought to live. On this suggestion, one camp believes that our only reasons are self-interested; another believes those reasons to be ethical; and still another thinks they involve satisfying one's desires.[33]

This suggestion, however, makes the ethical view too easy a winner. The claim that we ought to live ethically is tautological; it is hard to believe that those who use 'rational' differently disagree. Instrumentalists do not believe that I should do horrible things, even if my sadistic desires are coherent. An instrumentalist might say that I should do horrible things *given my desires*, but that is not to ultimately recommend the horrible behavior. And philosophers who use 'rational' interchangeably with 'prudent' aren't usually ethical egoists, so typically they don't believe that we (really) ought always to pursue our self-interest.

2. Perhaps the different uses of 'rational behavior' reflect competing theories of rational desire: on instrumentalism, rational desires are coherent desires; on the ethical use, rational desires are ethical desires; and prudentially, rational desires are those sanctioned by self-interest. On this suggestion, the different uses of 'rational' reflect competing theories of what it is to act on rational desires.

This suggestion, however, merely passes the buck by creating the problem of pinpointing the alleged dispute between the different theories of rational desire. Are these competing theories of rational desire, or do different uses of 'rational desire' differ merely in meaning? This new problem is just as hard as the original problem of pinpointing the alleged disagreement between the different theories of rational action. Nothing has been gained.

substantive disagreement in 'Climbing the Mountain', but I did find some passages in the 2006 draft of that book to be puzzling without that background assumption); David O. Brink, 'Sidgwick's Dualism of Practical Reason', *Australasian Journal of Philosophy* 66.3 (1988), pp. 291–307; Kurt Baier, 'Egoism', *A Companion to Ethics*, ed. P. Singer (Oxford, 1991), pp. 197–204; Quinn, 'Rationality and the Human Good'; Shaver, *Rational Egoism*; and many others.

[33] Similarly, Chase Wrenn suggested to me that the different uses of 'rational behavior' might reflect competing views about which actions are supported by our best arguments. So, on instrumentalism, we cannot profitably argue about ultimate ends; on the ethical and self-interested accounts, we can, with those accounts disagreeing about which ends our arguments support. However, this can't be what is at issue, since the content of the literature on ultimate ends overlaps so little with that on rational behavior. For a review of the argument-types employed about ultimate ends, see Stuart Rachels, 'A Defense of Two Optimistic Claims in Ethical Theory', *Philosophical Studies* 112.1, January (I) 2003, pp. 1–30, sect. 2.

https://doi.org/10.1017/S0953820809990252 Published online by Cambridge University Press

3. Parfit defines 'rational action' as 'action not open to rational criticism'.[34] This must mean action not open to *good* rational criticism. For if rational action were not open to any sort of rational criticism, good or bad, then *no* action would be rational, since any action can be badly criticized. But now consider this dilemma: in the phrase 'good rational criticism', either 'rational' is redundant, or it isn't.

Suppose first that 'good rational criticism' is redundant, since 'rational' refers to reason-giving, and good criticism always involves giving reasons. So interpreted, Parfit's definition of 'rational action' is 'action not open to good criticism'. This definition yields a third attempt to pinpoint the alleged disagreement underlying the different uses of 'rational': that those uses represent conflicting beliefs about which acts are not open to good criticism.

This suggestion has the same problem as the first: it makes the ethical view too easy a winner. *Of course* the act not open to good criticism is ethically best. The ethically best action just is the action best supported by the relevant considerations. It is hard to believe that many philosophers dispute that.

Now consider the second horn of the dilemma, which Parfit prefers:[35] 'rational' is not redundant in the phrase 'good rational criticism'. What rational action is not open to is not good criticism of any sort but merely good *rational* criticism. On this view, the different uses of 'rational' reflect conflicting beliefs about which acts are not open to good rational criticism. This suggestion, however, has the same problem as the second: it passes the buck by creating the problem of characterizing rational criticism. We need to understand why philosophers variously identify action not open to rational criticism with self-interested behavior, ethical behavior, and instrumentally sound behavior. Parfit offers no account of rational criticism. Without one, this proposal doesn't explain any disagreement underlying the various uses of 'rational'.

Thus, three suggestions have failed to identify the alleged dispute; they either create an equally hard and similar problem, or they entail that the 'dispute' has such a clear winner that no large group of philosophers would ever have disagreed. In criticizing these proposals, I did not nitpick; I did not assume they should offer precise analytic

[34] Parfit actually defines *not irrational* as 'not open to rational criticism' (*Reasons and Persons*, p. 119). Thus, I have substituted 'rational' for 'not irrational'. Parfit's definition can be cleaned up a bit. As it stands, it entails that one must act irrationally when every possible action is open to some rational criticism. But one still acts rationally when one's act is the best of the lot. So, a better definition of 'rational' is 'open to the least amount of rational criticism'. Making this change in the text would not materially affect the discussion.

[35] Personal correspondence, 29 October 2001.

https://doi.org/10.1017/S0953820809990252 Published online by Cambridge University Press

conditions for rational action and then offer clever and contrived counterexamples. Rather, I showed that they do not illuminate an underlying dispute even dimly. And I know of no proposal better than these.

Moreover, appealing to paradigm cases won't help show the existence of a disagreement. Imagine some exemplars of rational behavior. One might think the dispute is over what unites them as rational: are they rational because of their relation to self-interest, to ethics, or to desire-satisfaction? But merely gesturing at examples doesn't show that some privileged set of features unites them as rational; each may be rational in three different senses.[36]

Also, for what it's worth, my intuitions suggest that there is no substantive issue. Consider, for example, an act that will harm me but is ethically right and will maximize my expected preference-satisfaction. Suppose I willingly make some large sacrifice for my child. Is my act rational? One might say: 'The act is rational on the ethical and instrumentalist understandings of "rational", is not rational on the self-interested understanding,[37] and that is the end of it.' Intuitively, this seems right to me; I don't sense a further issue about whether the action is 'really' rational.

I can't prove that no dispute divides instrumentalists, rational egoists, and rational ethicists. Perhaps there is an explanation of the dispute that no one has offered. Or perhaps the disagreement resists explanation; the disagreement is about what makes behavior rational, but this notion of rationality cannot be characterized in any illuminating way. I can't rule out these possibilities, but I know of no reason to believe them. And, Ockham's razor counsels denying that there is a genuine dispute, if we lack any justification for believing in one.

I conclude that no legitimate controversy underlies the three different uses of 'rational'. The ethical, self-interested, and instrumentalist uses reflect mere semantic difference, not substantive disagreement.

Of course, we can have substantive disagreements about whether an act is rational, if we both understand 'rational' in the same way. In such a case, we're not really arguing about rationality; we're arguing about ethics, or prudence, or coherence. Debates within one of the three major uses may also concern which analysis or precisification most deserves the commendatory force of 'rational'.

[36] Compare Alston on paradigm cases of justified belief: William Alston, 'Epistemic Desiderata', *Philosophy and Phenomenological Research* 53.3 (September 1993), pp. 527–51, at pp. 536–7.

[37] Of course, the act would be rational on the self-interested use, if on that use one is rationally permitted to forego one's greater good for moral reasons.

https://doi.org/10.1017/S0953820809990252 Published online by Cambridge University Press

We can even have substantive disagreements when our understandings differ. Consider, for example, whether it is rational to co-operate in the prisoner's dilemma. Is the issue whether co-operating would promote the agent's goals, or promote the agent's self-interest? This is rarely made clear, but in practice this doesn't affect the debate, since by stipulation the agent's goals are self-interested.

But where there is no real dispute, we should end the frivolous debate over what rational action 'really' is. Many philosophers have wondered, for example, whether rational behavior is really self-interested behavior. Kurt Baier says, 'the jury on this case is still in disarray'.[38] One distinguished juror, Derek Parfit, devotes Part Two of *Reasons and Persons* to criticizing the self-interest theory of rationality.[39] On my view, the self-interested use of 'rational' no more reflects a *theory* of rationality, competing with the ethical and instrumentalist theories, than the slope-of-a-stream use of 'bank' reflects a theory of banks, competing with the financial-institution and row-of-objects theories. The jury's disarray results from there being no genuine dispute.

Innocence by association?

'Everything you have said about *rational*', one might protest, 'can be said about the ethical *ought*: "ought" is used in different ways (for example, in Kant's deontological way, in Mill's utilitarian way, and in Ross's pluralistic way); no one has given an illuminating, theory-neutral characterization of "ought"; and we can claim of paradigm ethical actions that we ought to do them in different senses of "ought". But Kant, Mill, and Ross genuinely disagree about ethics; their disputes are not verbal or confused. So, if the age-old ethical disputes are legitimate, then why shouldn't the age-old rationality disputes be legitimate as well?'

I'll make three points in response.

First, the facts cited do count against the legitimacy of ethical debates. Such ideas legitimately motivate anti-realist views in ethics. The similar moves I have made also count (inconclusively) against the legitimacy of the rationality debates.

Second, even if we can characterize the subject matter of ethics no better than by saying 'it's about how one ought to live, all things considered', this is less problematic than characterizing the subject matter of practical rationality by saying 'it's about how one ought (rationally) to live, all things considered'. Here is why. Both the ethical 'ought' and the rational 'ought' are normative terms. But while

[38] Baier, 'Egoism', p. 204.
[39] Parfit says that this is the main thesis of Part II on p. 120.

https://doi.org/10.1017/S0953820809990252 Published online by Cambridge University Press

normative realists must grant the legitimacy of the ethical debates, they needn't grant the legitimacy of rationality debates. Normative realists must accept ethics as legitimate because the ethical 'ought' has so little content, aside from pure normative content, applied to action. But the rational 'ought' must have more content than that, since it is not the same concept as the ethical 'ought'. Thus, while a defense of the ethical debates must merely defend the normative, a defense of the rationality debates must also supply this extra content. And, as I have argued, no one has explained what that content is, in a way that makes sense of the rationality debates.

Third, and again for what it's worth, my intuitions suggest that ethical debates are legitimate. Suppose that a little girl whose mother has just died asks an atheist if he believes in Heaven. Should he lie? I have the sense that he really should; I am not satisfied by the claim, 'He ought not to lie on Kant's use of "ought"; he ought to lie on Mill's use of "ought"; and he ought to lie on some pluralistic uses of "ought" – and that's the end of it, aside from the laborious task of cataloguing the other uses of "ought" and determining whether he ought to lie on them.' By contrast, as I've said, my intuitions don't support the existence of a legitimate dispute among those who use 'rational' in the three different ways.

A note on meaning

Here again are the main ways philosophers use 'rational behavior':

1. rational behavior = what the agent ought to do, given what she credibly believes. (ethical rationality; the ethical interpretation)
2. rational behavior = what is best for the agent, given what she credibly believes. (the self-interested or prudential interpretation; rational egoism)
3. rational behavior = what maximizes the expected satisfaction of the agent's ultimate, coherent preferences, given what she credibly believes. (instrumentalism)

These interpretations, I've argued, differ semantically but not substantively. How are they related semantically? Do they represent entirely separate meanings of 'rational behavior'?

The interpretations share two important semantic elements. First, they emphasize using reason to achieve goals (be those goals ethical, self-interested, or simply the agent's). If one acts rationally, it always involves intelligently using what one knows, or using the evidence at one's disposal. Second, 'rational' is typically a term of praise.[40] Rational

[40] Gibbard thinks endorsement is central to the meaning of 'rational'. See *Wise Choices*, pp. 6, 10, 20.

https://doi.org/10.1017/S0953820809990252 Published online by Cambridge University Press

action is commended for its relation to good cognition, and perhaps also for the value of its ends (be they ethical, self-interested, or simply the agent's). So, the different meanings of 'rational behavior' can be thought of as species under a genus. The generic meaning is that 'rational' describes behavior that could result from good reasoning. The species differ as to what the reasoning is about – whether self-interest, the agent's goals, or what it would be best to do.[41]

Given how much 'rational' is used, this analysis is bound to be imperfect. Sometimes philosophers (and often economists) seem to use 'rational behavior' to refer to self-interested behavior without intending to praise it. And sometimes the cognitive element of rational action seems to disappear behind the goal of the behavior. The important point for us is that when philosophers apply 'rational' to different practical domains, they are not thereby disagreeing.

Gibbard's objection

Gibbard argues that 'rational' can't mean 'instrumentally effective'. And since instrumentalists aren't defining 'rational', they must be offering a theory of rationality. Gibbard's argument thus threatens my claim that instrumentalism, rational egoism, and the ethical interpretation aren't competing theories.

In his argument, Gibbard says that this belief is coherent:

(B): I am rationally required to give weight to my future happiness, *even if* doing so wouldn't achieve my goals (because I am indifferent to the future).[42]

And so, he concludes, 'rational' can't *mean* 'instrumentally effective'. And therefore, instrumentalism is a substantive theory.

This argument, however, is fallacious. At most it shows that 'rational' can't have *just* an instrumental meaning, since it has no such meaning in (B).[43] And I agree that 'rational' doesn't *just* have an instrumental meaning, since there are at least two other semantic interpretations of 'rational action' (namely, the ethical and prudential interpretations). In (B), 'rational' (or rather, 'rationally required') is best interpreted prudentially (though it may also be interpreted ethically). Even if I don't care about my future happiness, not giving it weight is imprudent (and violates the ethical principle of giving weight to everyone's future interests). So, Gibbard cannot conclude that 'rational'

[41] I thank Jonathan Bennett for suggesting this genus/species analysis. And I thank William Fitzpatrick for making a similar suggestion.

[42] Gibbard, *Wise Choices*, p. 12.

[43] A different reply is also possible. An instrumentalist might say that (B) is coherent, since even if giving weight to my future happiness doesn't promote my *present* goals, it probably promotes my future goals, since I am likely to care about my future happiness in the future.

https://doi.org/10.1017/S0953820809990252 Published online by Cambridge University Press

has no instrumentalist meaning. And thus he hasn't shown that instrumentalism is a substantive theory.

IV. CHANGING OUR USE OF 'RATIONAL BEHAVIOR'

Using 'rational' in three different ways – ethical, prudential, and instrumentalist – encourages the false belief that there are three rival theories of rational behavior. Also, making 'rational' do so much work can obscure its meaning on particular occasions, since many acts are rational in more than one way. For example, exercising might be rational in terms of self-interest, ethics, and the agent's own desires, so calling exercise 'rational' is less informative than one might think. On balance, we should change our use of 'rational', despite what the different uses have in common.

Should we continue calling self-interested behavior 'rational'? I don't see why; neither clarity nor economy would be lost in just talking about 'self-interested' or 'prudent' behavior. Moreover, calling only self-interested behavior 'rational' suggests that ethical egoism is true: that self-interested reasons are, ultimately, our only practical reasons.[44] The insinuation is dangerous. Studies suggest that economists are especially egoistic, perhaps because they habitually associate 'rational' with self-interest.[45]

I also think we shouldn't use 'rational' ethically, though I see why people do. Ethical behavior *is* the behavior best supported by reasons. And the normative force of 'rational' *is* appropriately aimed at right action. However, it will usually be clearer to talk about what is ethical or what is best (given what the agent knows) rather than what is rational. If one wants to emphasize the relationship between ethics and reasons, it might be better to talk about 'reasonable behavior' than 'rational behavior'. 'Reasonable' is the less corrupted term.

We shouldn't try to eliminate the instrumentalist use of 'rational', for several reasons. First, using 'rational' to refer to the instrumentalist idea is economical because that idea is so complicated. Second, the instrumentalist use is deeply embedded in the decision theory literature; it could not easily be abandoned. Third, it is natural to connect the rationality of behavior to that of belief and desire. The instrumentalist tradition lets us explore these connections, if we find them worthwhile. Fourth and finally, the normative connotations of

[44] On a related note, Scanlon says: 'Given that there are reasons for action other than those provided by an agent's own interests, I see no justification for giving this one class of interests such special status' (*What We Owe to Each Other*, p. 31).

[45] See Robert H. Frank, Thomas Gilovich and Dennis T. Regan, 'Does Studying Economics Inhibit Cooperation?', *Journal of Economic Perspectives* 7.2 (Spring 1993), pp. 159–71.

https://doi.org/10.1017/S0953820809990252 Published online by Cambridge University Press

'rational' are not wholly out of place in the literature of decision theory. For instrumentalists, rationality involves norms such as consistency of choice, coherence of desire, adherence to evidence, and intelligibility of action.

So, the instrumentalist use should stay, and the others should go. It's not that rational action is 'really' what the instrumentalists say it is; it's just that their use of 'rational' is more useful and entrenched than the others.

srachels@bama.ua.edu

https://doi.org/10.1017/S0953820809990252 Published online by Cambridge University Press