Thrasymachus, Reasons and Rationality

Peter G. Woolcock

Thrasymachus, in Plato’s *Republic*, claims that justice is what is in the interest of the stronger party. Cross and Woozley interpret this as the claim that the weak have a duty to serve the interests of the strong. I argue that this interpretation is mistaken and that Thrasymachus agrees with Socrates (and Aristotle) that justice is giving people their fair share. Thrasymachus’ point is that people who act justly in effect are serving the interests of the stronger and, thereby, are acting irrationally. For Thrasymachus, the rational thing to do is to pursue your own self-interest even if it is at the expense of others. Thrasymachus seems to be adopting an instrumentalist account of rationality here. I contrast this with the kind of objectivist account of rationality offered by philosophers like Scanlon, Parfit and Nagel which maintains that there are substantive reasons why one should sometimes pursue the interests of others for its own sake. I argue that this apparent objectivity is an artifact of how talk about reasons works in public, as opposed to private, reasoning. I conclude that Thrasymachus is correct that egoism is rational, but he is mistaken to think that it is the only rational position. Acting justly is also rational. I conclude that, while a just person is rational in sincerely advocating justice, Thrasymachus is irrational in sincerely advocating egoism.

In *The Republic*, the Sophist philosopher Thrasymachus sets out to convince Socrates and others that justice or right is “what is in the interest of the stronger party” (Plato, 1955:65). There has been much debate over what precisely Thrasymachus meant by such a claim. In their book, *Plato’s Republic: A Philosophical Commentary*, Cross and Woozley consider four plausible interpretations. The first, which they call “naturalistic definition”, contends that Thrasymachus is offering us a definition of justice. The second, “the nihilist view”, has Thrasymachus suggesting that there is no such thing as justice, that it is “an illusion from which Socrates and others suffer, and from which he proposes to liberate them”. A third interpretation, which they call “incidental comment”, assumes the fact of “the existence of justice, or at least a belief in its existence, and contents itself with making a remark about justice, intended to bring it into disrepute”. They reject all three of these views and settle, instead, on the view that Thrasymachus really does think that justice is what is in the interest of the stronger party. Thrasymachus is maintaining, they hold, “that it is the duty of the
weaker to help and serve the stronger”. They call this “the essential analysis” view (Cross & Woozley, 1964:32–38).

One consequence of the essential analysis interpretation, they point out, is that it makes Thrasymachus guilty of an incoherence. In defending his view, Thrasymachus not only says that justice is what is in the interest of the stronger but also that something is just only if it is in another’s good, not merely one’s own (Cross & Woozley, 1964:41). Thrasymachus tells us that the terms “just” and “unjust” apply as much to the strong as they do to the weak but, if this is so, then the strong man who is just in that he serves his own interest by exploiting others is, thereby, unjust by not acting for the good of others. One suspects that any remotely competent thinker is unlikely to make such an obvious error. This suggests that the essential analysis interpretation is not really what Thrasymachus had in mind. Moreover, if he had made such an obvious muddle, Socrates would surely have been quick to point it out.

What did Thrasymachus have in mind then? It seems to me that a version of “the incidental comment” interpretation accommodates most comprehensively the things that Thrasymachus says about justice, and does so without finding him guilty of a naïve confusion.

Before I spell out my alternative interpretation, it is important to note a point that Cross and Woozley make about how Thrasymachus’ remarks on justice have been translated. In the rendering given above, H. D. P. Lee uses the word “define” to translate the Greek word dikaiosyne. This would seem to commit us to the “naturalistic definition” interpretation of what Thrasymachus has to say. However, Cross and Woozley point out, this is a misleading translation. Dikaiosyne, they say, is “a very general word meaning ‘answer’, with sometimes a legal flavor about it which may or may not be present here, meaning the defendant’s answer to a charge brought against him”. They go on to note that, while the suggestion “that Thrasymachus is offering an answer to Socrates’ question about justice is quite compatible with its being intended as a definition of ‘justice’ it does not entail that it is one” (Cross & Woozley, 1964:25).

Independently of the issue of the correct translation of dikaiosyne, it is unlikely that Thrasymachus is telling us that what people mean when they use the words “justice” or “right” is “what is in the interest of the stronger”. If the subjects of the ruler, say, were to think that “justice” was identical in meaning with “what is in the interest of the stronger” (in this case, the ruler), why would they think that they thereby had a reason to do it, other than fear of the superior power of the ruler? Thrasymachus, however, sees the just man as someone who believes that the justice or rightness of an act is a reason for doing it in its own right, independently of any intimidation from the ruler. When asked by Socrates, “Will one just man compete with another and want more than his fair share of an act of justice?” Thrasymachus replies, “Certainly not; otherwise, he would not be the simple, agreeable man we supposed him to be”. By contrast, when asked whether the unjust man will “compete with the just and want more than his share in an act of justice?” Thrasymachus says, “Of course he will; he wants more than his share in everything” (Plato, 1955:79).
Implicit here, then, is the notion that justice, of its nature, has something to do with fair shares. As Aristotle was to put it later, “In the popular mind the description ‘unjust’ is held to apply both to the man who takes more than is his due and to the man who breaks the law. It follows that the man who does not seek to break the law and the man who does not take more than he is entitled to will be ‘just’. ‘Just’ therefore means (a) lawful and (b) what is ‘equal’, that is, ‘fair’” (Aristotle, 1953:140). For Thrasyvmaschus, the “fairness” element in this account is primary. The unjust ruler, after all, may well follow the laws that he has established, but he has established them in order to get more than his fair share. He does not, thereby, become just.

If justice means something like the fair allocation of shares, then the strong can come in two types — either strong just people who allocate shares fairly or strong unjust people who take more than their fair share because they have the power to do so. Likewise, weak people can also come in two types — weak just people who expect, and respect, the fair allocation of shares; and weak unjust people who take more than their fair share whenever they can get away with it. As a rule, I will follow Thrasyvmachus in taking the ruler as a typical case of the stronger, and the subjects of the ruler as typical cases of the weaker. This gives us the four categories of the just ruler, the unjust ruler, the just subject and the unjust subject. This formulation matches Thrasyvmachus’ account and seems to give us a perfectly coherent typology.

This is, however, only a preliminary to an exposition of what Thrasyvmachus is trying to say about justice. His central point is that justice, which is giving people their fair share, amounts in practice to nothing but serving the interests of the stronger. It does no good to the person who acts justly but merely results in their being ripped off and exploited. As he puts it:

The just man always comes off worse than the unjust. For instance, in any business relations between them, you won’t find the just man better off at the end of the deal than the unjust. Again, in their relations with the state, when there are taxes to be paid the just [sic — should be ‘unjust’] man will pay less on the same income, and when there’s anything to be got he’ll get it all. Thus if it’s a question of office, if the just man loses nothing else he will suffer from neglecting his private affairs; his honesty will prevent him appropriating public funds, and his relations and friends will detest him because his principles will not allow him to push their interests. But quite the reverse is true of the unjust man. (Plato, 1955:73)

I will say more about it later, but it looks as if Thrasyvmachus’ thesis, put briefly, is that it is irrational to be just, and rational to be unjust. I will call this the “rationality” interpretation of Thrasyvmachus. It falls into Cross and Woozley’s third class of interpretations, the “incidental comment” class in that Thrasyvmachus is making an incidental comment on giving people their fair shares (i.e. justice). His comment is that behaving this way is irrational.

The following dialogue, beginning with a question by Socrates, demonstrates the point:
'Justice is a good quality, I suppose, and injustice a bad?'
'Yet that is just what I do.' (Plato, 1955:78)

The ruler, as this makes clear, can be unjust. He is so when he does not give his subjects their fair share. But the subject can also be unjust. He is so when he does not give his ruler or his fellow citizens their fair share. The subject is just when he does give his ruler and his fellow citizens their fair share, and that unlikely character in Thrasymachus’ scenario, the just ruler, gives himself and his citizens their fair share.

Thrasymachus, then, has given us a factual thesis, namely, that people who act justly are worse off in life, especially with respect to their own happiness, than unjust people. It does not pay to give people their fair share. If you want a happy life, then get as big a share as possible, no matter how unfairly you treat others. In fact, for Thrasymachus, the most rational thing of all is to be as unjust as possible by becoming a tyrant if you can manage it. As he puts it:

Tyranny is not a matter of minor theft and violence, but a wholesale plunder, sacred or profane, private or public. If you are caught committing such crimes in detail you are punished and disgraced: sacrilege, kidnapping, burglary, fraud, theft are the names we give to such petty forms of wrongdoing. But when a man succeeds in robbing the whole body of citizens and reducing them to slavery, they forget these ugly names and call him happy and fortunate, as do all others who hear of his unmitigated wrongdoing. (Plato, 1955:73)

But is Thrasymachus correct to say that rationality requires us to maximise our self-interest? Even if it is true that becoming a tyrant enables you to maximise your self-interest, does it follow that this is the most rational way to behave?

In holding that this is the most rational thing to do, Thrasymachus seems to be endorsing what is called an “instrumentalist” account of rationality. On this account, rationality is not a matter of what your goals are but, rather, is a matter of what means you choose in order to achieve those goals. You are rational when you choose those means that you believe are most likely to achieve those goals. Further, you act rightly when you choose those means that, in fact, will achieve your goals. It is possible,
therefore, to act rationally without, thereby, acting rightly, although the rational agent will always aim at acting rightly, that is, in a way that will actually achieve his or her goals overall. In order to determine whether or not you have acted rationally, on this account, we need to know facts about you, about what you desire or care about or what matters to you most. Your rationality is a function of your attitudinal states. You are irrational, on this account then, when you do not try to avoid what you would regard as self-defeat, that is, when you act in a way you believe would foil, rather than further, what matters to you (perhaps even brings about what you most wanted not to happen).

There is, however, an account of rationality that holds that certain goals, of their nature, are things that it is irrational to pursue, even if they are pursued by the most effective means available. One of the goals that this account regards as irrational is the goal of maximising one’s self-interest irrespective of the interests of others. It maintains that there are acts that are *substantively* rational, not just *procedurally* rational, for example, acting justly. On this account, we do not need to know facts about what you desire or care about or what matters most to you. Rather we assess whether or not you’re pursuing certain desires, say, is the rational thing for you to do in terms of whether or not you have good reasons for doing so. T. M. Scanlon calls this kind of account “reasons fundamentalism” (Scanlon, 2014:2) but I will use Parfit’s term “objectivism” (Parfit, 2011:3) and will refer to Thrasymachus’ view as “subjectivism”.

It is a view common among objectivists that one sometimes has decisive reason to care for the interests of others, not as a means to one’s own ends, but for its own sake. Scanlon, for example, holds that “For a person in control of a fast moving automobile, the fact that the car will injure and perhaps kill a pedestrian if the wheel is not turned is a reason to turn the wheel” (Scanlon, 2014:2), and this is true regardless of any attitudinal state of the driver. It would be true even if the driver wanted to kill the pedestrian out of revenge, say. Parfit tells us that, “If I am judging who deserves some prize, that would give me a reason to ignore the fact that one of the contestants is my best friend” (Parfit, 2011:33). It is an implication of his example that I would have this reason even if I cared more about pleasing my friend than I did about acting in an honourable way. So, if the objectivists are correct, then Thrasymachus must be mistaken in his view that it is always rational to care only about one’s own self-interest.

Objectivism is primarily an account of what it is to have a reason to act, and defines what it is to be rational in terms of what one has most, or even decisive, reason to do. Derek Parfit, for example, puts it as follows: “In most cases, I believe, some possible act of ours would be *rational* if we have beliefs about the relevant facts whose truth would give us sufficient reasons to act in this way”. It would be “what we *ought rationally* to do if these reasons would be decisive” and it would be “*less than fully rational* if we have beliefs whose truth would give us clear and decisive reasons *not* to act in this way”. It would be “*irrational* if these reasons would be strongly decisive” (Parfit, 2011:34).

In order, then, to determine whether or not a person has acted rationally we would need to know whether or not he had good reason for doing so. The notion of a “reason”
appears to be logically prior to the notion of “rationality”. It is important, therefore, if we are to be in a position to evaluate this account of rationality, to understand what it means by “a reason”. According to Parfit, however, “the concept of a reason is indefinable in the sense that it cannot be helpfully explained merely by using words. We must explain such concepts in a different way, by getting people to think thoughts that use these concepts”. Nonetheless, we can get some grip on the notion, he suggests, if we think of the fact that gives us a reason for having some attitude, or for our acting in some way, as “counting in favour” of the attitude or act (Parfit, 2011:31).

What, then, is thought to be the problem with subjectivism? I will look at an objection raised by T. M. Scanlon. He is concerned that the subjectivist seems to think that the mere fact that “one has some desire, intention, or other attitude” settles the question of whether or not one has a reason to act (Scanlon, 2014:13). Thrasymachus, for example, would maintain that the tyrant who desires to maximise his chances of happiness thereby has a reason to further his self-interest by ruthlessly exploiting the citizens he rules. However, Scanlon points out, “one can always ask oneself why one should have these attitudes — whether they can be justified in the relevant way” (p. 13). So, in the case of the tyrant, the question is always open, both to himself and to others, whether or not he ought to have such a desire? And this is an appeal that is not answered merely by reiterating that he, in fact, has this desire. It is a challenge to him to offer a reason, to himself as much as to anyone else, as to why he should not replace this desire, say, with a concern that everyone have their share of happiness, i.e. with a concern for justice. Why should he have his happiness at the expense of everyone else?

Scanlon notes that the philosopher, Christine Korsgaard (1996:38), has tried to answer this objection. He says that she “recognises, indeed emphasises, the possibility of this kind of reflective ‘stepping back’ when one is thinking about what reasons one has. In such a situation, she says, a person must keep on asking ‘why’ until she comes to a point at which it is ‘impossible, unnecessary or incoherent to ask why again’. This is what she calls the search for the unconditioned” (Scanlon, 2014:13).

Scanlon does not find Korsgaard’s response satisfactory. “But when”, he says, “is it unnecessary to ask any further? I would say that this depends on the substantive merits of the answer one has reached — on whether this answer is clearly correct, or whether there is any reason to doubt it” (Scanlon, 2014:14).

Contrary to Scanlon, it seems to me that Korsgaard is on the right track here. However, in order to show why this is so, it will be helpful to make a distinction between cases of what I will call “public reasoning” as opposed to cases of “private reasoning”.\(^1\)

In cases of public reasoning, the parties to an argument may need to hide from their

\(^1\) Scanlon (2014:12–13) refers to a related distinction made by Gilbert Harman, namely, the distinction between reasoning when two people are arguing about what reason for action one of them has (“external reasoning”) and reasoning about what reasons one has oneself (“internal reasoning”) (Harman, 2007:3). My distinction, however, focuses on what it would be self-defeating for people in public discourse to offer as reasons.
audience the reasons they have for proposing a particular course of action. This may need to be done in order to prevent the audience using that knowledge to forbid that course of action and, thereby, thwart the goal that the course of action was meant to achieve. They have to be careful that what they put on the table is not self-defeating for them. By contrast, an individual (a woman, say), who engages in private reasoning, can be entirely honest with herself about what her reasons are without, thereby, putting the achievement of her goals in jeopardy.

The nature of public reasoning, I believe, is such that it mistakenly gives the impression that reasons have the objectivity that the objectivists claim for them. Consider the case of a man whose only goal, like the man whom Thrasymachus thinks is rational, is the furtherance of his own self-interest, who is out to get as big a share of wealth, power and, thereby, happiness, as possible, regardless of the cost to others. In a situation where he cannot force others to obey him, his only choice is to convince them by argument. However, unless he is lucky enough to have an audience whose only concern is the furtherance of his self-interest, it will be self-defeating of him to declare to them that his proposed courses of action have as their purpose his happiness at the expense of theirs. He may, then, privately run through the options that are most likely to get their agreement, possibly by deceit or manipulation, but it would be self-defeating to reveal this reasoning to them too, even though it contains no incoherence or inconsistencies as a piece of private reasoning.

The upshot of this is that, when the parties have no power over each other, it will usually be self-defeating to declare that the reason why other people should agree to a course of action is merely that this would achieve some desire you have, or that it is required by some attitude or like or dislike you have, or that you care that it be done, or that it matters to you that it be done, or that it is in your self-interest that it be done. None of these are likely to be seen by your audience as a reason why they should comply. Thomas Nagel argues that even talk of desiring to desire does not capture what it is for something to be a reason. Such second-order desires, in his view, are not sources of motivation but “simply the manifestation in our motives of the recognition of certain rational requirements” (Nagel, 1997:108). As we have seen, however, we do not need Nagel’s explanation because, unless you are very lucky, people don’t care enough about these facts about yourself to act to bring them about. In fact, they are likely to regard it as a bad precedent to allow other people to get away with offering things such as these as reasons. It is not in their interests, individually or collectively to do so.

This brings out an important feature in the logic of the word “reason”. As we have seen, objectivists plausibly treat a reason for an action as a “consideration in favour” of that action. An argument that could be offered for objectivism\(^2\) is the reluctance we have to say that a person has a reason for an action unless we endorse his or her doing that action. To illustrate: suppose you tell me that my nephew wants more than

\(^2\) This point was prompted by discussions with Garrett Cullity.
anything else to be rich but the only plausible way that this can happen is if he inherits my fortune. Does this mean that my nephew has a reason to murder me, namely, that this is the only way he will become rich and get what he wants more than anything else? Thrasymachus, it seems, would say “Yes”. But what about me, the uncle? What would I say? Do I really think that it is a consideration in favour of my nephew murdering me that this would get him the riches he wants more than anything else? I may agree that my nephew thinks he has a reason. I would probably go so far as to say that, given the facts, it is the rational thing for him to do (the “self-defeat-avoiding” thing). But I won’t go so far as to say that he really has a reason. To say that would appear to endorse his murdering me, and I definitely don’t do that.

And what about you? Would you say that he really has a reason to kill me, that his becoming rich at the cost of my death really is a consideration in favour of his committing a murder? Now, suppose I asked you this question in a public reasoning situation where, in this particular instance, lots of your friends, neighbours and colleagues are listening in. Even if you privately think that it is a perfectly good reason for killing someone that it gets you what matters to you most, provided you can get away with it, it would be self-defeating of you to let your audience know this fact about the kind of person you are. The rational thing for you to do in the circumstances is to condemn my nephew’s design, to make it clear to everyone that you do not think it is a consideration in favour of killing someone that it gets you what you most want. To do otherwise is to make them suspicious and fearful of you and, thereby, make them less likely to cooperate with you on your proposals because they suspect that you have no scruples about harming them if it suits you and you can get away with it.

We could make a distinction, as many philosophers do, between a motivating reason and a justifying reason, that is, a distinction between what one thinks one has a reason to do (a “motivating reason”) and what one really has a reason to do (a “justifying reason”). We could say, then, that my nephew has a motivating reason to kill me, but he does not have a justifying reason (Smith, 2004:174). What we do not need to do, however, is to claim that there is anything more to the notion of a justifying reason other than an instrumental account. My refusing to acknowledge that my nephew has a justifying reason to kill me is a matter of my not seeing his inheriting my wealth as a consideration in favour of his committing murder, of my refusing to endorse his behaviour. Likewise for you if you refuse to see it as a reason. There does not need to be some non-instrumental, substantive entity around that we are refusing to attach to his action. We do not need to see these internal states as the “manifestation in our motives of the recognition of certain rational requirements” as Nagel advises us to do.

We see, then, that much of the apparent objectivity of reason-talk is just an artifact of what reasons it would be self-defeating for people to allow in public discourse in situations where argument rather than force is the means by which agreement is reached. Certain things are, in effect, ruled out as reasons in public discourse, not because they lack some characteristic in themselves that other reasons have but, rather,
because it would be self-defeating to allow them to count as reasons in contexts where no-one has a predominance of power.

There is an exception to this, namely, the situation of a philosophical discussion. In a philosophical discussion, to borrow a phrase from phenomenology, one “brackets out” what one’s personal commitments in ordinary life are. A philosopher (say a man) who wants to know whether there is reason to be moral need not be someone who has any actual doubts about the matter. His is a theoretical exercise. He wants to understand what kind of a thing moral talk is, what (if anything) makes a moral claim true or false, and so on. Elizabeth Anscombe, it seems to me, is a famous example of misunderstanding this when she says “But if someone really thinks, in advance, that it is open to question whether such an action as procuring the judicial execution of the innocent should be quite excluded from consideration — I do not want to argue with him; he shows a corrupt mind” (Anscombe, 1981:40). This is might well be an appropriate response in a non-philosophical context where the participants are revealing to each other the values they actually endorse — their real moral commitments — but this is not what is going on in a philosophical discussion.

This, I think, explains a phenomenon that Christine Korsgaard (1996:33) has noted. As Scanlon puts her point, “She imagines two people disagreeing about whether something is a reason for a certain action, and she observes that it is mere reiterative stone-kicking for one party to say, in the face of the other’s denial, ‘But it just is a reason!’” (Scanlon, 2014:13). Bernard Williams describes such a move in the face of someone who denies the irrationality of their action as “bluff” (Williams, 1981:111). I suggest that this kind of case mostly occurs in philosophical discussions, not in non-philosophical, everyday ones. If it is a philosophical context, then the stone-kicker (say, a woman) still has her mind in everyday mode. She thinks that her opponent (say, a man) is revealing what he values are whereas, in all likelihood, he is not challenging its status as a reason but, rather, is wondering what it is about it (or anything else) that makes it a reason.

To return to the claim that one must repeat asking “Why?” until one comes to a point at which it is impossible, unnecessary or incoherent to ask why again. This can take both a public reasoning and a private reasoning form. In the public reasoning version, someone keeps asking you why until you come to the point where you have no further reason to offer. It is, of course, possible that the explanation for why you have no further reason to offer is that the position you have adopted lacks a justification. Korsgaard, however, thinks that your position would be justified if you reached it in a way that made it impossible, or unnecessary or incoherent to ask why again. How can this be so?

Consider, for example, the case where someone asks you what reason you have for thinking that people ought not to engage in acts of causing other people gratuitous pain, such as torturing them merely for your own pleasure. In a non-philosophical situation you are likely to be shocked by such a question. The mere asking of it seems to suggest that your questioner (say, a man) sees nothing wrong in torturing people
for fun. Your likely response here is some variation on Anscombe's "I do not want to argue with him; he shows a corrupt mind". Suppose, however, that you choose to continue the discussion (probably while edging a little further away from your questioner). You are likely to reply along the lines of, "Well, how would you like it if someone hurt you just for fun?" You are, in effect, asking him whether or not he himself already knows that there is something about pain that gives all intelligent sentient beings a reason not to want it to happen to themselves unless there is some non-gratuitous reason for it. The very nature of pain, the kind of thing it is in itself, is a reason-giving fact to such beings.

Suppose, however, that he pushes you further and says, "Fair enough, the very nature of pain gives me a reason to avoid it being gratuitously inflicted on me, but what reason is there for me not to gratuitously inflict it on others?" Assuming again that, somewhat against your better judgement, you continue the conversation, what can you say in reply? Aren't you left with something like, "Are you telling me that you don't see the very nature of pain as a reason not to gratuitously inflict it on people? If it is really true that you don't, then I cannot offer you any further reason. What could a further reason be?"

Likewise, if someone tells you that he doesn't value justice, what else can you do to convince him otherwise but try to draw his attention to features of justice that he might have missed, features that you think reveal or confirm its value? However, given that he refuses to acknowledge the very nature of these things as a reason to act morally, and you are one of those others whom he might treat unjustly, you are likely to draw this conversation to a close before he decides to harm you. Moreover, you are likely to warn others to avoid him like the plague.

That, it seems to me, is more or less the way the discussion would go in a case of public reasoning in a non-philosophical context. It would be self-defeating, outside of a philosophical context, to admit to other people over whom you do not have a more or less invincible power, that you do not see the very nature of pain as a reason in itself to avoid gratuitously inflicting it on other people. It would, of course, be even more irrational, even more self-defeating, to let them know that you actually thought you had a reason to hurt them for your own pleasure (suitably qualified to exclude consenting masochists). In this way, it becomes irrational to ask why again, which is a kind of incoherence, even if it is not exactly either unnecessary or impossible.

In a case of private reasoning, however, as in an interior monologue, it need not be self-defeating to admit to oneself that one does not see the very nature of pain as a reason not to inflict it on others, even if you think its nature gives you a good reason to avoid having it gratuitously inflicted on you.3

3 Something like this approach is advocated by John McDowell (1995).

4 For a thorough account of why there is no logical impediment, such as contravening universalizability, to holding that one can take one's own pain as a reason for action yet ignore the pain of others, see D. H. Monro, Chapter 16.
How is the subjectivist to explain why some people see the very nature of pain as a reason not to inflict it gratuitously on others, whereas other people, such as Thrasy-machus, do not? The explanation lies in the kind of people they are, the kind of states that are internal to them, such as desires and attitudes. However, the moral person (a woman, say) does not see her internal states as the justification for her rejection of gratuitously inflicting pain on others. As we have seen, she sees such a policy as justified by the very nature of pain itself. If she is asked what it is about the nature of pain that justifies the policy of not inflicting it gratuitously on others she will not say that it is justified by her being in some internal state rather than another. It is not her desire that people not be hurt that she sees as underpinning the wrongness of hurting people for fun but the very nature of pain itself. That you desire something, then, will often fail to count as a reason for action in the public domain, even on a subjectivist account. Instead, the publicly allowable reason will lie in some fact about the action itself or its effect.\(^5\)

A similar constraint does not exist with internal reasoning. Suppose you decide that what you care about most, what matters to you more than anything else, is living a life devoted to pleasure, regardless of its effects on others (except when this might reduce your pleasure) — in other words, the kind of life that Thrasy-machus thinks it is rational for any of us to pursue. As Scanlon pointed out above, you can always ask yourself whether this is the life you ought to be leading. Is it a life which has reason on its side? You don’t have to convince anyone else, so you don’t have to restrict what you propose as a reason to only those items acceptable to others.

In what sense, then, could it be a life that you ought not to be leading? Well, it could be a life that is ultimately self-defeating for you. Just because you think that it is what you care about most, this may not be true. You may come to find that it is empty and leaves you friendless, and that these are conditions you wish to avoid even more than you desire pleasure. Of course, at the time when you falsely conclude that the pursuit of pleasure is what you care about most, it would be irrational of you not to pursue pleasure. A rational person (say, a woman) is one whose intention is always to act in ways consistent with what she \textit{thinks} is what she cares about most. Nonetheless, it may well turn out that it was the mistaken thing to do from the point of view of what really mattered to you most, that is, it may turn out that there was a reason not to do it. It was present to you as a motivating reason but failed to be a justifying reason. You thought you were avoiding self-defeat while actually bringing it about.

On this understanding of a “justifying reason”, what is a justifying reason in private deliberation may not be acceptable as a justifying reason in public deliberation. In private deliberation, you have a justifying reason to perform an act if the act will achieve what really matters to you most, even if you are not aware of what that is at the time. It is an act such that, were you not to do it, you would defeat yourself. In public deliberation, however, there are certain considerations that it would be self-defeating

\(^5\) This meets some of the objections to subjectivism raised by Talbot Brewer (2002).
for people to allow other people to count as being in favour of an act, as we saw earlier in the discussion of gratuitous pain. Even if it is true that it would be self-defeating for you not to live a certain kind of life, it may well be self-defeating for other people to treat your self-defeat as a consideration in favour — a reason — for their allowing you to live that kind of life, especially if it involves your using them as mere means to your own ends. It is unlikely to be among the range of reasons allowed as justifying reasons in public discourse.

It has not been shown, then, that Thrasymachus is mistaken to think that it can be rational to perform acts aimed at furthering one's own self-interest regardless of the harm it does to others. It would be rational to do so if one's own self-interest was what one thought one cared about most. This leaves one open to two kinds of error — one might be mistaken that this act really does further one's own self-interest or one might be mistaken in the belief that one's own self-interest was what one cared about most. Committing either of these errors, however, need not impugn the rationality of what one has done although, as a rational agent, one would prefer to make neither of them.

Unless one thought one was making one or both of these errors in one's private reasoning, what sense could be given to the claim that maybe one ought to do something other than what one most cares about? It makes sense as a challenge in public reasoning but, once you are sure that this is the thing that really matters to you most, it would seem incoherent, at the least, to ask why again in a private reasoning case. Of course, it makes sense for the man who thinks that he values his own pleasure above everything else to wonder whether he might turn out to be mistaken. Perhaps he will come to realise that he really values justice but, unless there is evidence present to him now that he will do so, he would not be acting as a rational agent to do anything other than to seek to maximise his own pleasure. However, the fact that his pursuing his own pleasure at their expense was the rational thing for him to do would not be the kind of thing that other people would usually allow as a reason in public discourse. The mere fact that it was the rational thing from his point of view would not count as a consideration in its favour from their point of view.

But Thrasymachus is not merely saying that it can be rational to pursue one's own self-interest at the expense of others. He is also saying that this is the only way to be rational. In his view, anyone who acts justly, who does not seek more than his fair share of the good things in life, is a fool. There is nothing in the analysis of rationality offered so far that warrants this conclusion. Thrasymachus has not shown us that it must be self-defeating to give others their fair share. Whether it is or not depends on what we most care about, and there seems no logical impediment to our caring most that we, and other people, behave justly. In fact, there is strong empirical evidence that many people think this way (Bloom, Chapter 16). There is something about the nature of justice, of fairness, of moral virtues such as honesty and generosity that many people find worthwhile in their own right, as the kind of thing that is reason-giving in itself. If these people acted in a way inimical to the realisation of these values, they would regard themselves as bringing about self-defeat.
While doing what you believe will avoid self-defeat is always the rational thing to do, whether it is acceptable to admit this in a case of public reasoning depends very much on what an individual counts as self-defeat. It would not be rational of Thrasymachus, for example, to declare that self-defeat for him would be failing to maximise his self-interest regardless of its effects on others. But it would be perfectly rational for a just man to admit that self-defeat for him would be his doing the unjust thing.

In this respect, people who value justice and fairness have an advantage over Thrasymachus. They can manifest the virtue of integrity, being able to match what they believe privately with what they say and do publicly. This option is not open to Thrasymachus except, perhaps, in a philosophical context such as the one where he feels free to say to Socrates what he really thinks. Even there, Socrates treats him rather gently, taking what he says as a theoretical position rather than a set of attitudes that Thrasymachus actually endorses, even though Thrasymachus rather obviously endorses them. After all, if Socrates really thought that Thrasymachus was the kind of person whose ideal life was “robbing the whole body of citizens and reducing them to slavery” (Plato, 1955:73), we would expect him to display even greater contempt for Thrasymachus than he does.

Bibliography

Anscombe, 1981

Aristotle, 1953

Bloom, 2013

Brewer, 2002

Cross and Woozley, 1964
Harman, 2007

Korsgaard, 1996
Christine Korsgaard, Sources of Normativity. Cambridge: Cambridge University Press.

McDowell, 1995

Monro, 1967

Nagel, 1997

Parfit, 2011

Plato, 1955

Scanlon, 2014

Smith, 2004

Williams, 1981
Bernard Williams, “Internal and External Reasons”. In Moral Luck. Cambridge: Cambridge University Press.