

*Non-tuism*¹

DONALD C. HUBIN
Ohio State University
Columbus, OH 43210-1365
USA

I

Contractarians view justice (or, more ambitiously, all of morality) as being defined by a contract made by rational individuals. No one supposes that this contract is actual, and the fact that it is merely hypothetical raises a number of questions both about the assumptions under which it would be actual and about the force of hypothetical agreement that is contingent on these assumptions.

Particular contractarian theories must specify the circumstances of the agreement and the endowments, beliefs, desires, and degree and type of rationality of the agents. How these issues are settled determines the force of the hypothetical agreement. The fact that ignorant people who desired only universal suffering would, under duress, agree to a certain principle gives us no reason to believe the principle is a correct moral principle or to think it rational to accept or act on it: some counterfactual assumptions undermine entirely the moral force of hypothetical agree-

1 In the course of thinking through the issues discussed here, I have incurred significant debts to several people. It was a paper by Jan Narveson, cited below, that first made me aware of the confusions surrounding the notion of non-tuism. My thinking was clarified by talks with Dan Farrell, Mike Morris, Diana Raffman, Peter King, Calvin Normore, Peter Vallentyne, Chris Morris, and David Gauthier. Finally, referees' comments from Howard Sobel and Richmond Campbell were uncommonly helpful.

ment. On the other hand, to take people just as they are, with their current beliefs, desires, endowments, and all, is to endorse their ignorance and mistakes as well as any previous injustice that affects their bargaining power.²

One of the more controversial assumptions made by contractarians concerns the desires of the parties to the contract. On some conceptions of rationality, this issue would not arise. However, contemporary contractarians assume an instrumental (means/ends) conception of rationality that ties the rationality of choice to the desires of the chooser.³ On *this* conception, it seems essential to make restricting assumptions about the desires of the contracting parties if determinate results are to be obtained. Of course, determinate results may be renounced by the contractarian; she may hold that even at the most fundamental level, the substantive nature of justice (or morality) is contingent on the desires of the parties to the contract.⁴ I shall focus, though, on contractarians who harbor hopes for determinate results. Rawls has such hopes; in the attempt to fulfil them he assumes not only that the contractors take no interest in the interests of others, but that they are motivated only by desires for the 'primary social goods' of powers and opportunities, income and wealth, rights and liberties, and the social bases of self-esteem. Even if one seeks to avoid such strong substantive assumptions, there is, as we will see, a powerful motivation for placing some limits on the desires of the contractors. It is a motivation that even David Gauthier, who eschews many of Rawls's idealizations about the contracting situation, does not resist.

In Rawls's theory of justice, the hypothesis under which agreement would take place is far removed from the actual situation — stripping individuals not only of knowledge of their particular endowments and characteristics, but also of their desires, values and aims in life. Theories that employ such a rarefied conception of the individual raise what might be called 'the rational compliance problem' — they fail to show that we have any reason to comply with the terms of the agreement that

2 David Gauthier offers a parable illustrating this in *Morals by Agreement* (Oxford: Clarendon Press 1986), 190-1. Unless otherwise indicated, page numbers in the text refer to this work.

3 I use 'desire' here as a surrogate for a variety of subjective conative states, some of which may be more properly called 'ends,' 'aims,' or 'goals.'

4 This sort of view is adopted by James Buchanan, *The Limits of Liberty* (Chicago: University of Chicago Press 1975) and *Freedom in Constitutional Contract* (College Station, TX: Texas A & M Press 1977); and Gilbert Harman, *The Nature of Morality* (New York: Oxford University Press 1977).

would be made by such individuals.⁵ Gauthier's contractarian theory, as defended in *Morals By Agreement*, seeks to avoid this problem by taking individuals 'as they are' — *more or less*. Like Rawls, though, he makes assumptions about the nature of the contractors' interests in one another. At times he assumes mutual unconcern — that individuals take 'no interest in one another's interests' (100). Elsewhere, he makes the weaker assumption of non-tuism, which requires only that individuals not take 'an interest in the interests of those with whom they exchange' (87). The distinction between these two is important, and will sometimes be of concern here. When it is, I will indicate this; when it is not, I will generally use 'non-tuism' to refer indifferently to either requirement.

The assumption of non-tuism seems designed to avoid two sorts of difficulties, one associated with what might be called 'negative' concern with others and the other with 'positive' concern. If the contractarian allows agents to be motivated by desires that other agents fare poorly, prospects for guaranteeing a mutually advantageous agreement seem bleak — an illustration of the extreme case being the following. Imagine just two people, one who has quite ordinary desires and another who desires only that the first person's desires be thwarted as much as possible. In this case, the outcome of an idealized bargaining situation that places no restrictions on the agents' desires will be the *ex ante status quo*, or whatever the preferred bargaining theory defines as the default ('no-agreement') outcome. This does not constitute an interesting moral constraint — especially for Gauthier, for whom the no-agreement point is amoralism.⁶

5 Gauthier makes this point, no doubt with Rawls in mind, when he says, 'Those who claim that moral principles are objects of rational choice in special circumstances fail to establish the rationality of actual compliance with these principles' (17). The rational compliance problem — the problem of showing that compliance is rational — is not, as even a casual observer of human nature knows, the same thing as the problem of ensuring compliance.

6 Gauthier appears to hope for more from his moral theory: '...moral constraints must apply in the absence of other-directed interests...indeed they must apply whatever preferences individuals happen to have' (100). This claim can be interpreted in various ways. Taken *sans phrase* it might appear to mean that regardless of the preferences of individuals and the situation in which they find themselves, moral restrictions exist and are applicable to them. It will become apparent, I think, that Gauthier cannot simultaneously satisfy *this* requirement, carry out his contractarian project, and succeed in showing that moral restrictions are rationally binding. Probably, though, we should take Gauthier's claim to mean only that when people's preferences and external circumstances are such that a mutually advantageous bargain would be struck in the initial situation, then the terms of the bargain apply regardless of the preferences people actually have. This requirement appears

On the other hand, agents motivated by a desire for the well-being of others may arrive too easily at agreements that coincide with our pre-theoretical moral convictions. The contractarian seeks to show that agents not directly motivated by any altruistic desire would have reason to agree to constraints on the pursuit of their own ends — constraints that serve, quite intentionally, to allow others to pursue their own ends. Some contractarians seek to show, in addition, that agents have reason to *comply* with these constraints. The compliance thesis becomes more interesting when we assume the agents in question are not motivated by direct positive concern for others.

This last point suggests that the rejection of positive concern with others is an optional piece of panache — that the contractarian's argument succeeds *a fortiori* without it. But this is not so. Unless the positive concern with others is equally shared, it appears that those with more fellow-feeling fare worse in the social bargain — a result that contractarians are loath to embrace. Of course, a presumed equality of positive concern could be achieved by assuming increased concern on the part of the mean-spirited, but this would rather obviously undermine the interest of the social contract — certainly to *these* people. The rejection of positive concern seems an attractive way to avoid apparent skewing of the results of the bargain; it seems to avoid 'double-counting' of those who are the object of the positive concern.⁷ As we shall see, though, it is simply not true that if the contractarian argument for the rationality of compliance works for agents without concern for others, it will work *a fortiori* for those with positive concern for others.

Screening off positive concern for others is no less controversial than screening off negative concern. The goal of rationalizing morality by employing the contractarian approach, I will argue, undermines the rationale for excluding tuistic concerns, whether negative or positive.

Gauthier is our target.⁸ He is not, of course, the only contractarian to make assumptions about mutual concern, but in addition to offering rich

consistent with the contractarian project and the goal of rationalizing moral restrictions. However, I argue below that this latter goal puts severe limits on what restrictive assumptions we are warranted in making about the motivations of agents in the initial situation. As a result, the cases in which a mutually advantageous bargain would be struck may be far fewer than Gauthier believes.

7 I have more to say about this rationale below, 462 ff.

8 The specific assumptions Gauthier makes in *Morals by Agreement* regarding mutual concern have been clarified and criticized by Christopher Morris in 'The Relation Between Self-Interest and Justice in Contractarian Ethics,' in Ellen Frankel Paul, ed., *Gauthier's New Social Contract (Social Philosophy & Policy 5 [1988])*, 119-53; and by

insights concerning the nature and justification of such assumptions, he embraces the project of establishing that it is rational for an agent to comply with the outcome of the social bargain, properly conceived.

While Gauthier's discussion of non-tuism and mutual unconcern is rich in insights, it is also misleading in places. To clarify matters, I draw attention to a distinction crosscutting that between non-tuism and mutual unconcern. This distinction provides at least two plausible senses of each term. Both of these senses are more tolerant than some have taken 'non-tuism' (/mutual unconcern) to be. And I suggest other senses of the requirement of non-tuism (/mutual unconcern) that seem to exclude what some have taken to be ruled out by this assumption.

After distinguishing various interpretations of what Gauthier might mean by 'non-tuism' and examining the sometimes surprising implications of these, I offer a general contractarian rationale for making restrictive assumptions about the preferences of the contractors. This rationale does not warrant the assumption of non-tuism. I then discuss several other rationales for the assumption of non-tuism suggested by Gauthier's discussion and argue that none of these is compelling. Finally, I argue that the assumption of non-tuism cannot be reconciled with Gauthier's fundamental moral project, which is to provide a rational grounding for morality.

Peter Vallentyne in 'Contractarianism and the Assumption of Mutual Unconcern,' *Philosophical Studies* 56 (1989) 187-92. Gauthier has responded to some of these criticisms in 'Morality, Rational Choice, and Semantic Representation: A Reply to My Critics,' *Social Philosophy & Policy* 5 (1988) 213-17, suggesting that the assumption of non-tuism can be given a less problematic role. The retrenchment is tentative, though. He says:

I should like the revision in the role and status of nontuism I have...sketched to be considered, not as a fixed alteration in my theory, but as a provisionalsuggestion. At the present time, it seems to me that this revision is needed to accommodate Morris's objections. But I should of course be pleased to find that a less radical change would suffice. ('Morality, Rational Choice, and Semantic Representation: A Reply to My Critics,' 217)

I maintain that, contrary to Gauthier's wish, in fact a *more* radical change is needed. Nothing less than dropping the assumption of non-tuism will be consistent with the primary goal of Gauthier's project. If he is serious about his assumptions and the criteria he explicitly endorses for the adequacy of a moral theory, he should be pleased to be rid of the assumption of non-tuism.

II Understanding the Assumption of Non-tuism

Interest and Interests

Begin by distinguishing an agent's *interest* from his *interests*. There is some (though not complete) support from ordinary discourse for using these labels to mark the distinction I have in mind; but it is the distinction that matters, not the labels. An agent may take an interest in things that are not in her interest (genuinely altruistic desires are but one example); she may fail to take an interest in her own interest (true selflessness is an extreme example of this); or she may take an interest in things that are directly contrary to her interest (that is, she may have directly self-destructive interests). I take what an agent desires, cares about, values, etc., to mark her *interests*—these things are what she is interested in. Her *interest* is her well-being.⁹

Gauthier defines both 'non-tuism' and 'mutual unconcern' in terms of 'taking an interest in the *interests*' of others. (See above, 443.) We cannot determine from these bare definitions whether Gauthier is referring to the desires and values (i.e., *interests*) of others or to their well-being (i.e., *interest*).¹⁰ The first reading, though, has more extensive and explicit textual support and coheres better with Gauthier's general approach.¹¹ On this account, the proposed assumptions require that

9 The well-being of an agent *might* include the satisfaction of her desires, but plausible accounts will not restrict it to this. See my 'Prudential Reasons,' *Canadian Journal of Philosophy* 10 (1979) 63-81.

10 The definition of 'mutual unconcern' is semantically ambiguous because we don't know whether 'interests' is used collectively or distributively. The definition of 'non-tuism' is univocal given the labelling I have endorsed for the above distinction, but we can hardly stick Gauthier with this labelling.

11 There are passages in *Morals by Agreement* that seem based on a confusion of interest and interests in the senses suggested in the text here. For example, Gauthier says: 'co-operation is possible only among contemporaries who actually interact. Although an individual can do much to benefit or harm his descendants, only those whose lives overlap with his can benefit or harm him in return' (298). But this is not true if 'benefit' and 'harm' are interpreted, as Gauthier explicitly requires, in terms of the utility level of agents. For as our considered desires can be about anything whatsoever, and our utility is determined by the degree to which our desires are satisfied, we can surely be benefited or harmed in the sense of having our utility level raised or reduced by actions of distant future people (and indeed by people with whom we have no causal relations at all). For a related discussion, see my and Mark Lambeth's 'Providing for Rights,' *Dialogue* 27 (1988), 497; reprinted in Peter Vallentyne, ed., *Contractarianism and Rational Choice* (Cambridge: Cambridge University Press 1990).

agents' desires be (in some sense) independent of the desires of others. Let us begin with this understanding of the requirements.

Independence of Utility Functions

Reading the requirements of non-tuism and mutual unconcern as excluding desires directed at the *interests* (values, preferences, desires, etc.) of others finds direct support in the text. When Gauthier clarifies these requirements, he says that they ensure that people's utility functions are 'strictly independent; [that] no person gains or loses simply from the *utilities* of others' (86; my emphasis).¹²

The notion of two utility functions being strictly independent could mean several things, but in this case it presumably means that they are *logically* independent. A contractarian will certainly not require that people's desires be such that the satisfaction of each individual's desires is *actually* not dependent on the degree to which the desires of others are satisfied.¹³ Problems of justice arise precisely because there exists this sort of conflict between people's desires (preferences, utility functions, etc.).¹⁴

There are two ways in which two people's utility functions could fail to be logically independent. First, one or both might desire that the other's desires be satisfied or frustrated. Here, the failure of independence is due to the content of the individuals' desires *considered severally*. Call this 'intrinsic logical dependence,' since the logical de-

12 Peter King has pointed out to me that the requirement that people's utility functions be independent and that the people take no interest in the interests (understood in terms of preferences) of others are not equivalent. Parents frequently have desires about the content of the desires of their children. This is not equivalent to having desires about the degree of satisfaction of the desires of the children. Thus, two people could have independent utility functions (in the sense that their levels of utility are logically independent) and yet 'take an interest in the interests of others.' Most of us have derivative desires about the content of the interests of others. If we enjoy live jazz, we hope that enough of our neighbors do so as well, to support local jazz clubs. Those who are preferentially conformist or non-conformist also take an interest in the interests of others without necessarily gaining or losing from the utilities of others.

13 As Gauthier points out, '[w]ere the scarcity faced by each person not aggravated by the presence of her fellows, then however self-biased she might be, her activities would bear little relation to those of others, and neither conflict nor co-operation would result' (114).

14 The most elegant statement of this is still Hume's, I believe. See §III, Part I of *An Enquiry Concerning the Principles of Morals* (La Salle, IL: Open Court 1960).

pendence is intrinsic to the individuals' desires (preferences, etc.). Second, while neither has desires about the desire satisfaction of others, they may desire states of affairs that are logically related in some way. For example, one might desire that a state of affairs obtain and another that it not obtain. Here, the failure of independence is due to the content of the desires of both individuals *considered jointly*. Call this 'extrinsic logical dependence.'

For a number of reasons, it is objectionable to take the requirements of non-tuism and mutual unconcern to rule out sets of preferences that are merely *extrinsically* logically dependent. In the first place, the offending desires offend only in conjunction with the desires of others. There is no individual desire of which it is true that *that* desire must be removed if logical independence is to be re-established. If I desire that a tract of land be developed and you desire that it not be developed, our desires are logically unsatisfiable as a set. But which is the offending desire?¹⁵ There is no more reason to dismiss one than the other in attempting to achieve independence of our utility functions.¹⁶

However we solve the problem of which desires to filter out, it will remain that extrinsic logical independence of utility functions is a far stronger requirement than one would initially take non-tuism or mutual unconcern to be. For example, one would not ordinarily think either the desire that an area be developed or the desire that it not be developed violates the requirements of non-tuism or mutual unconcern; nor would one take either to constitute taking an interest in the interests of others.

15 The problem of identifying the offending desire in a set of extrinsically logically dependent desires arises because extrinsic dependence is, obviously, not an intrinsic property of a desire. We might, then, choose to reject all desires that could be logically related to the desires of others. The *possibility* of extrinsic logical dependence is an intrinsic property. Unfortunately for this strategy, it is an intrinsic property of *all* desires; no desire is such that someone *cannot* have a desire that is extrinsically logically related to it.

16 When two people have desires that conflict, it is, no doubt, good strategy for the resolution of conflict to try to find more basic desires that ground these directly conflicting desires but are not *themselves* directly in conflict. Perhaps I desire that the tract of land be developed only because I see this as a means to housing the homeless, and you desire that it not be only because you see development as a threat to wildlife. Shifting discussion to the more fundamental desires of housing the homeless and preserving wildlife habitat — desires that are not in logical conflict — raises the possibility of solutions that are mutually satisfactory. But it is not always possible to 'retreat' to desires that are not in logical conflict (or otherwise logically related). In the text I assume that the conflicting desires are basic (non-derivative) desires.

Suppose, then, that we take the requirements to address only desires that generate *intrinsically* logically dependent utility functions — that is, we rule out by these requirements only those (second-order) desires that directly regard the desire satisfaction of others. There are, I believe, good reasons for thinking that even these requirements, so construed, are unjustifiably strong.¹⁷ But they are also weaker than one might suppose.

Such a requirement will rule out desires that the desires of others be thwarted or that they be satisfied. However, it will not rule out certain other desires that one might initially think to be tuistic or other-directed. For example, my desire that you suffer eternal damnation does not violate this requirement. There is nothing intrinsic to this desire that connects it logically to your desires. (I assume here that Hell is not *analyzed* in terms of frustrated preferences. If I have begged the question against your conception of Hell, substitute ‘eternal physical pain’ for ‘eternal damnation.’) You are free to desire eternal damnation for yourself, to desire that you not be eternally damned, or to have no desires about the matter. In the first two cases, our utility functions will turn out to be logically dependent, but only extrinsically so. In the last case, they will be logically independent. Desires for the suffering of others (or for their pleasure) are not intrinsically logically connected to the desires of others or to their utility functions.

A desire for a positional good, which many take the assumptions of non-tuism and mutual unconcern to rule out,¹⁸ is consistent with these assumptions if they are interpreted in terms of the absence of intrinsically dependent utility functions. This is true at least of most desires for positional goods. My desire to make more money than anyone else in my neighborhood is not connected with the desires of others in such a manner as to make my utility function intrinsically dependent on those of my neighbors. They may or may not have desires about our relative incomes. If they do not, our utilities are independent. There *is* a case in which desire for positional goods might violate the requirements as we are now interpreting them; it is when one has desires concerning her degree of preference satisfaction relative to others. For example, if I were to desire that my preferences be better satisfied than those of others with whom I am interacting, this might violate the requirement that my utility function not be intrinsically dependent on those of such persons (thus violating

17 See Vallentyne, ‘Contractarianism and the Assumption of Mutual Unconcern,’ and Morris, ‘The Relation Between Self-Interest and Justice in Contractarian Ethics.’

18 Jan Narveson clearly holds this view in ‘McDonald and McDougal, Pride and Gain, and Justice: Comment on a Criticism of Gauthier,’ *Dialogue* 27 (1988) 503-6.

non-tuism and, *a fortiori*, mutual unconcern). This, though, is not the garden variety positional good.

My desire that you be my slave (or that I be yours), that others suffer (or that they flourish), that goods be distributed equally (or unequally) — none of these is tuistic in the sense that it necessarily involves ‘taking an interest in the *interests* of others’ or that it entails that utility functions are intrinsically (or even extrinsically) logically dependent.

Non-tuism, understood in the way we have been suggesting, precludes second- (and higher-) order preferences directed toward the preference satisfaction of the person with whom we are interacting. (Mutual unconcern precludes such preferences about the preference satisfaction of any others.) Non-tuism does not exclude preferences that are other-directed — only those that are directed toward the *preference satisfaction* of others. As a result, it will exclude less than it might appear to. Nevertheless, for reasons to be discussed below (465-7), the non-tuism requirement excludes too much; it undermines Gauthier’s major moral claim, *viz.*, that it is rational to comply with the agreement that would be made in the appropriate initial bargaining position.

Indifference to the Well-Being of Others

So far we have been looking at the requirement of non-tuism as excluding desires about the *interests* (utilities, desires, preferences, etc.) of others. Alternatively, we might take non-tuism to rule out desires directed at the *interest* (well-being) of others. So understood, it would exclude, for example, altruistic and sadistic desires, as well as desires that everyone fare equally well. While this interpretation coheres poorly with the thrust of Gauthier’s work, some might think it is suggested by certain of his remarks. For example, in defending the assumption of non-tuism, he says, ‘it is neither unrealistic nor pessimistic to suppose that beyond the ties of blood and friendship, which are necessarily limited in their scope, human beings exhibit little positive fellow-feeling’ (101). ‘Fellow-feeling,’ I propose, is ordinarily taken to suggest something more than the mere desire that others’ desires be satisfied. Another motivation for trying to interpret ‘non-tuism’ as ruling out concern for the *interest* (well-being) of others is that it is a more natural interpretation of these terms, which are commonly understood, insofar as there is such an understanding, to rule out the sorts of desires mentioned above.¹⁹

¹⁹ This is especially true with respect to ‘mutual unconcern.’ ‘Non-tuism’ is more a quasi-technical term, of which, perhaps, there is no *common* understanding (at least

If a theory requires us to assume that individuals take no interest in the well-being of others, it ultimately owes us a substantive account of 'individual well-being' — one that distinguishes it from individual utility level. Gauthier does not offer such an account in *Morals by Agreement*.²⁰ If we assume that avoiding experiences like pain, depression, and anxiety, and having experiences like pleasure, joy, and contentment are *part* of a person's well-being, then this conception of 'non-tuism' will better coincide with our ordinary conception of 'selfishness,' excluding desires that are usually thought of as tuistic or 'other-directed' — ordinary altruism and sadism, for example.

Understanding 'non-tuism' in this way will not, however, exclude everything that might reasonably be thought of as tuistic. For example, many think that positional and competitive desires are tuistic. I might desire to outdo you in some competition, or to make more money than you, or to be better at cracking safes than you are. None of these things is necessarily connected with your well-being. Your life may go just as well for you (or even better) if you lose the competition, make less money than I, and are less skilled than I at safe-cracking; and I may care not a whit that your life goes just as well for you, for what I aim at has to do not with your well-being but with something else. I seek not your ill, but some state that is defined by comparison with you — to be better, worse, smarter, dumber, stronger, weaker, taller, shorter, more beautiful, uglier, or whatever, than you are.

To exclude these desires, if we decide they need to be excluded, we could reject positional desires directly. We could assume that people have no desires for states that are defined by comparison to others. This is significantly different from the idea that utility functions be inde-

among common people). Even with respect to 'non-tuism,' though, there is a history of broader usage than Gauthier's. (See P.H. Wicksteed's *The Common Sense of Political Economy* [London: MacMillan 1910], 163-82.)

20 Indeed, Gauthier presents and defends a purely subjective theory of value — more precisely, one that makes value dependent only on individual preferences. This may seem to preclude offering the sort of account mentioned in the text. But perhaps not. One might, consistent with such a subjective theory of value, hold that sense can be made of a purely naturalistic account of individual well-being (or flourishing) that is not wholly dependent on individual preferences. While such a concept has, this sort of subjectivist might insist, no direct moral significance — certainly not that attached to the concept of *moral value* by a consequentialist — it is a well-defined concept, and one that can be used in understanding 'non-tuism.' The fact, though, that Gauthier offers no account of 'individual well-being' distinct from utility maximization is further indication that this sense of 'non-tuism' is not what he has in mind.

pendent or that people not be concerned with the well-being of others. I do not think it a defensible assumption,²¹ and I shall argue later that it, too, would be destructive of Gauthier's project.

Be that as it may, even this won't exclude all the sorts of desires one might expect non-tuism to exclude. For example, I desire that you not have a certain object — say, a Rolls Royce Silver Shadow. This is not a positional desire. If we want the initial bargaining position and the bargaining itself not to be influenced by such desires directed at others, we might exclude desires that directly concern the commodity bundles enjoyed by others.

Not all concern with the lives of others can be understood in terms of concern for their utility level, for their well-being, for their position vis-à-vis ours or for the commodity bundles they consume. One might, for instance, desire that others be physically revolting, stupid, clever, understanding, wise, or compassionate. None of these is necessarily

21 Narveson seems to assume that desires for positional goods must be excluded in order for the bargaining situation to be a non-zero-sum game — for there to be a mutually advantageous bargain possible (505). This, of course, is not true. In the first place, we may have other desires that make the situation non-zero-sum. Second, the various positional desires may not be opposed to one another. So, for example, I may desire to make more money than you and you to have better developed pectoral muscles than I (or even, less plausibly, to make less money than I). Whether a situation is zero-sum or not depends on the collective preferences of the agents. Garden variety positional desires are not intrinsically related to the desires (preferences) of others; hence, they do not necessarily render a situation zero-sum. Finally, even if each positional desire I have is directly opposed by a positional desire of yours we may still be in a non-zero-sum game. Suppose that I want to have both a larger income and more advanced academic degrees than you, and you have similar desires with respect to me. It may still be that I care more about 'winning' in the category of income and you care more about 'winning' in the degrees category. We will, perhaps, find ourselves in a non-zero-sum game. Even if Narveson were correct in believing that desires for positional goods must be excluded if a non-zero-sum game is to exist, this would constitute a tenuous defense of the exclusion. Presumably the reasoning is that if the game is zero-sum there will be no mutually advantageous bargaining solution and if there is no such bargaining solution, then contractarian theories of morality will be silent about such cases. This will constitute a reason for excluding desires for positional goods only if one assumes (a) the correctness of contractarian theories of morality, and (b) that morality is not silent on these cases. The first of these is frequently what is at issue in such discussions. The second is by no means one to which a contractarian must commit herself. Indeed, where the situation is one of strict competition, where there is no possibility for mutual advantage, it seems more consonant with the contractarian project to hold that there are no moral restrictions than to hold that the moral restrictions are those it would be reasonable to institute were we to have different preferences — ones that did not put us in strict competition.

connected with their preferences, well-being, status, or with the commodity bundles they consume. (Though, presumably, if they consume too little, they will be none of these things.) Yet these desires will influence the initial bargaining position and the bargain itself as much as would other desires that some seek to exclude by appeal to assumptions of non-tuism and mutual unconcern. If there is some inappropriate skewing of the bargaining situation that is produced by desires regarding the utility levels or well-being of others, it is unclear why it would not also be produced by desires of the sort just mentioned.

Unconcern with Others

We considered, first, an interpretation of 'non-tuism' that ties it to an agent's concerns with the *interests* (desires, values, utilities) of others. On this interpretation, the non-tuism requirement rules out far less than one might expect; it does not rule out desires for the suffering of others, for example. Next, we considered an interpretation that related non-tuism to an agent's desires concerning the *interest* (well-being) of others. This interpretation, too, made the requirement more tolerant than anticipated, allowing, for instance, desires for positional goods. We mentioned, in passing, the possibility of adding to the definition of 'non-tuism' clauses to exclude desires for positional goods or desires concerning the commodity bundles enjoyed by others. We turn now to more radical interpretation of 'non-tuism' that would make the requirement, in a fell swoop, exclude all of the desires discussed above. This interpretation would have it that any desire containing an essential direct reference to any other individual or group of individuals is tuistic. At times, Gauthier seems to take 'mutual unconcern' in this way,²² but it does not seem to be his considered interpretation.

22 'Problems of inheritance may seem to be dissolved by the assumption of mutual unconcern. A person who takes literally no interest in *her fellows* and their interests must be entirely indifferent to what befalls her possessions and privileges when she is no longer able to derive benefit from them' (300; my emphasis). Here, mutual unconcern is taken to include the assumption that agents take no interest in other agents, not just that they take no interest in the preference satisfaction of other agents.

It is not clear, by the way, why Gauthier thinks that a person who takes 'no interest in her fellows and their interests' must be unconcerned with the posthumous disposition of her property. It does not, of course, follow from the fact that one takes no interest in one's fellows that one takes an interest only in oneself. Perhaps, rather than (or in addition to) non-tuism, Gauthier means to assume that a person's utility is entirely a function of the commodities and services he consumes

This conception of non-tuism leads to a highly restrictive assumption that is very difficult to justify. Why should these desires be excluded? Why may the desires I have about my car and the survival of the snail darter have their effect on the definition of the baseline for bargaining and on the interests that motivate the social bargain but my desire that a friend survive her bout with cancer not? In addition, such a maneuver clearly requires some account of who is to be considered an 'other individual.' (This apparently innocent phrase conceals a philosophical briar patch.) Perhaps grown children are clearly 'other individuals' in the appropriate sense regardless of how much love and concern we have for them. What, though, of minor children — in particular, very small children? What of future generations? What of the severely retarded or mentally deranged? What of animals and fetuses? If desires for the well-being of animals are precluded, the contractarian loses one of the obvious ways to include within his theory restrictions on cruelty to animals.²³ Gauthier, 'trusting theory rather than intuition,' may be unconcerned with whether his theory accords with 'the supposed "plain duties" of conventional morality' (269). But even supposing we dismiss prohibitions on cruelty to animals as *merely* supposed duties of *merely* conventional morality, Gauthier must be concerned with how, 'if morality is to fit within the domain of rational choice,' we are to justify excluding from the motivation of the bargainers their real life concerns for the well-being of 'animals, the unborn, the congenitally handicapped and defective [who] fall beyond the pale of a morality tied to mutuality' (268). It would be especially puzzling how morality is supposed to fit within the domain of rational choice since the only reason such concerns would not be included in this derivative way in a morality tied to mutuality is that the particular moral theory in question gives a privileged position to a person's desire that *he* not suffer — a position it denies to his desire that his dog not suffer.

or of the experiences he has. At one point he says, 'each person's utility is strictly determined by the goods he consumes and the factor services he provides' (86). Whether this endorses the position described here depends, of course, on how we understand 'goods.' In any event, Gauthier never defends the *extremely* strong assumption that a person's preferences not extend to states of affairs beyond his experience. Such an assumption seems both indefensible and unnecessary for the defense of Gauthier's theory.

23 Given that 'to make covenants with brute beasts, is impossible' (Hobbes, *Leviathan* [London: Collier-MacMillan 1969], 109), contractarians typically include the interests of animals in the theory by pointing to human concern for them. (See, for example, D.A.J. Richards, *A Theory of Reasons for Acting* [Oxford: Clarendon 1971], 182.)

Gauthier has elsewhere made it clear that concern for third parties is not a violation of the requirement of non-tuism,²⁴ which precludes concern only when it is directed toward an individual with whom one is interacting. His comments suggest that we should take him everywhere to be intending 'non-tuism,' even when he says, speaking loosely, 'mutual unconcern.' Since animals, the severely retarded, etc., are not parties to the social contract, I take this to settle the question of Gauthier's position on the inclusion of their interests in the social bargain.²⁵ Such interests are included (derivatively) provided those party to the bargain care about them.

Perhaps, in light of the above considerations, other contractarians would agree with Gauthier that my concern for the well-being of my dog, like my concern for the condition of the 1962 Studebaker warehoused in my garage, does not violate non-tuism. But, then, why is it that my concern for my Studebaker and my dog do not violate non-tuism while my concern for my grown children and my friends does? These people may be third parties to many particular interactions in which I am involved, but they are, presumably, parties to the social contract. In that context they are individuals with whom I am interacting, and if non-tuism requires that I have no concern for those with whom I am interacting, then it will screen off my concern for these individuals.

III Justifying the Assumption of Non-tuism

Contractarian Presuppositions

It is clear from the preceding that there are several plausible, but quite different, conceptions of the assumption of non-tuism. I think the interpretation of 'non-tuism' as requiring only that individuals' utility functions be free of intrinsic logical dependence coheres best with Gauthier's project and is most consistent with his explicit assertions. As I have indicated, however, it is not consistent with everything he says, and some, in the mistaken belief that they are following Gauthier, have taken the assumption to rule out far more than this. These misunderstandings motivated the task we have just completed: teasing out the ambiguities

24 'Morality, Rational Choice, and Semantic Representation,' 214-15

25 It is not clear, however, whether it solves the problem raised by Chris Morris that Gauthier invokes it to solve. But I set this issue aside here.

in 'non-tuism.' We now turn to the task of justifying the assumption of non-tuism.

It is helpful to begin at the beginning of the contractarian's story. What assumptions, if any, is the contractarian, *qua* contractarian, committed to making about people's preferences? We might think that the requisite assumptions are just those that must be made for there to be a mutually advantageous bargaining situation. Perhaps we can gain some insight into *this* issue by considering everyday cases. When I purchase a car, I am concerned to minimize my expenditure, and I have no intrinsic concern for the profit of the dealer. I am certain, too, that he is out to maximize his profit and has no intrinsic concern for my well-being. One might think, then, that a model for economic relations and bargaining is that of egoists rationally pursuing their self-interest — that this is what distinguishes economic relations from co-operation based on shared purposes.

The model of egoists cooperating for self-interested reasons is important because it reminds us that cooperation is possible even in the absence of shared ends; that is, it points out clearly that, as P.H. Wicksteed puts it, 'co-operation extends beyond the limits of common purpose' (171). This fact is crucial to the contractarian's (and I think, more generally, the liberal's) conception of justice and morality. But the egoistic model of economic relations is misleading, for, as Wicksteed emphasizes, 'the economic relation...is necessary alike for carrying on the life of the peasant and the prince, of the saint and the sinner, of the apostle and the shepherd, of the most altruistic and the most egoistic of men' (171). Complete selfishness is not destructive of economic relations, but neither is it a precondition of them.

According to Wicksteed what is necessary for the economic relation is non-tuism: 'the economic relation does not exclude from my mind everyone but me, it potentially includes everyone but you' (171). Taken literally, this suggests a form of non-tuism that excludes any interest *at all* in those with whom we interact, but it is not entirely clear whether Wicksteed meant the exclusion to be so sweeping. Perhaps he meant to exclude only concern with the well-being or with the utility of the person with whom we are interacting. Perhaps he did not distinguish clearly between these.

Wicksteed takes a step in the right direction in shifting the focus from altruism and egoism in general to some relation we have to the specific person with whom we are interacting. He errs, though, in thinking that the required assumption has to do with fellow-feeling toward that person — that it requires a sort of localized absence of altruism or, more generally, concern directed at others. We make bargains with loved ones, after all: parents with children, wives with husbands, brothers with sisters. It is plainly wrong to suppose that the economic relation is

absent where there is concern for the person with whom we are interacting, her well-being or her utility. To assume that the content of the bargain is uninfluenced by this concern is equally mistaken. Parents may negotiate with their children precisely to promote the well-being, and even the desire satisfaction, of the children.

What in fact is needed for economic relations and a potentially mutually advantageous bargaining situation is this: that the outcome of everyone following her utility-maximizing strategy is sub-optimal, in the sense that there is some alternative outcome preferred by everyone. In order for this to obtain, the parties to the bargain (exchange) must have all-things-considered preferences (desires) that define a non-zero-sum game.

Does this mean that contractarian theory assumes that preferences are such that these conditions will be met? I think not. Suppose that preferences are such that the game is one of pure conflict. Imagine two people whose preference rankings over the various outcomes are inverse images of one another. Among perfectly rational and informed agents with these preferences, there can be no bargaining and no mutually beneficial outcome.²⁶ But the contractarian need not assume that people's preferences are not in this sort of conflict. He must simply deny that considerations of morality and justice apply in such a case. Such a denial may cause him no discomfort.

Alternatively, suppose two people have exactly the same preference ranking over outcomes. Such people might cooperate based on what Wicksteed calls 'common purpose,' and they may need to coordinate their actions, but they would have no need of bargaining in the normal sense.²⁷ They might make agreements of a sort — agreements designed to coordinate action — but these would not be agreements to restrict the pursuit of their own aims. We could view these agreements as generating benign moral rules (benign, that is, from a rational egoist's point of view, because they don't restrict people from doing what they want); or we could, following Gauthier, decline to call such rules 'moral rules.' In any case, if the contractarian insight is plausible in the familiar cases in which preferences are in conflict but are not diametrically opposed, it

26 A bargaining theory can still offer an outcome, of course. But any plausible theory will offer only the default, 'no-agreement' outcome. Thus, the theory can be universal in the sense of always offering an outcome for any coherent set of preferences, but some of these outcomes will not plausibly be considered the outcomes of bargains, even if they are the outcomes of a bargaining theory.

27 Again, a bargaining theory may offer a solution in these cases, but there seems to be no natural sense in which the parties have struck a bargain.

will also give plausible answers to the case in which preferences are perfectly coincident.

What we are doing here, of course, is distinguishing between assumptions required for contractarianism to yield correct results and assumptions required for contractarianism to yield moral restrictions. I am suggesting that the assumption that preferences are such that there is a 'mixed game' (rather than one of pure conflict or pure coordination) is, from the contractarian perspective, the second sort of assumption.²⁸ If the assumption is not correct in a given circumstance, the contractarian should hold that her theory offers the correct answer that there are no moral restrictions in such circumstances.

Of course, on the contractarian view, it is the 'mixed game' case that is interesting. For simplicity, the contractarian might restrict her attention to such cases. One way to do so would be to assume *directly*, merely as a heuristic, that in real life the contractors face a mixed game. But there is another, historically more influential, way to focus attention on such cases: make assumptions about the world that are *sufficient* to guarantee that the contractors will find themselves in such a game. This is Hobbes's way. On the standard interpretation of Hobbes, he assumes that goods are scarce, that human desire for these goods is insatiable, that humans are sufficiently comparable in their abilities and powers as to be mutually vulnerable and that they are motivated only by concern for their own well-being (a much stronger assumption than non-tuism or even mutual unconcern). These assumptions seem sufficient to ensure a mixed game. But they are certainly not necessary for such a result. Psychological egoism, a strong assumption about the content of the agents' preferences, seems not to be a part of the contractarian's central insight about justice as being a mutually attractive bargain. The same should be said about non-tuism and mutual unconcern.

While I believe that the relatively weak assumptions that preferences form a mixed game and that the result of individual utility maximization is sub-optimal are all the contractarian must assume in order to generate moral restrictions (and that even this is not a presupposition of the

28 In 'Morality, Rational Choice, and Semantic Representation: A Reply to My Critics,' Gauthier suggests that he must assume that people are likely to confront prisoners' dilemma situations. But this assumption is necessary, he makes quite clear, in order to show 'that it is advantageous for each person to comply with constraints that it would be rational for all to agree to, provided that others may be expected to be generally similarly compliant' (215). That is, he sees it as necessary for the derivation of (rationally binding) moral rules, not for the correctness of the contractarian position.

correctness of the theory), there might be grounds for adopting stronger assumptions. What the contractarian takes those grounds to be depends on what criteria she imposes for the adequacy of a contractarian theory. We shall look at Gauthier's views on this soon, but first we must distinguish two roles played by non-tuism in his contractarian theory as presented in *Morals by Agreement*, and we must develop one of these roles — one about which we have said little so far.

Non-tuism and the Baseline for Bargaining

One obvious role played by motivational assumptions like non-tuism in a contractarian theory is in determining the rational agreement point. This is assumed to be determined by, among other things, the agents' preference functions. The assumption of non-tuism will (typically) lead us to consider different utility functions than those determined by the full set of an agent's desires.

Gauthier employs non-tuism in this role. But it plays another role in his theory as well. It is used to define the baseline for bargaining. As was noted at the outset, a rational bargain from a morally indefensible initial situation hardly seems to carry the moral significance a contractarian desires. But we must have an initial situation; bargaining can only take place against a backdrop of endowments. These initial endowments must be justified, but, on pain of regress, they cannot be justified by the contractarian methodology.

Different contractarians may have different standards in mind in justifying the initial bargaining position. Gauthier is primarily concerned to demonstrate that the situation is such that it would be rational to comply with bargains that would rationally be made in it. Others, more reliant on pre-theoretical intuitions, might be concerned to show that the situation is such that the outcome of a rational bargain would be in accordance with our pre-theoretical convictions about justice and morality. Either, or both, of these considerations may offer grounds for making restrictive assumptions about the preferences of the contractors in defining the baseline for bargaining. We have restricted our discussion to those forms of contractarianism that aim to show that compliance with the social bargain is rationally required. These theories should accept Gauthier's criteria for justifying assumptions about the initial bargaining position: we should make those and only those assumptions necessary to ensure that it is rational to dispose oneself to comply with the terms of a bargain that would be struck.

Gauthier believes that people bring to the bargaining table those resources they would control were everyone to have complied with Gauthier's version of the Lockean proviso. This proviso forbids bettering

oneself through interaction with another that worsens the position of the other. 'Bettering' and 'worsening' are understood in terms of subjective utility, and so the question arises whether this utility is a measure of the satisfaction of all of the agent's preferences or some restricted set of them.

I have elsewhere criticized Gauthier's use of the Lockean proviso to determine the initial situation (Hubin & Lambeth). It is, I think, indefensible either by appeal to pre-theoretical moral convictions or by appeal to the need to solve the problem of rational compliance. Some of the criticisms I have offered of this use of the proviso depend on cases that violate the assumption of non-tuism interpreted either as a restriction on our concern for the utility of others or on our concern for the well-being of others or in one of the other ways suggested above (e.g., as ruling out any preferences regarding others at all, or as ruling out desires for positional goods). For example, the proviso has the unhappy consequence of ruling out using your natural talents to the fullest if doing so would make you better off than you would be in the absence of another and makes her worse off than she would be in your absence even if the only reason for this is that it satisfies your desire to be more accomplished than she in some activity and violates her desire that you not be. It would seem that rational agents would not agree to constrain the use of their natural talents because of this effect on their (positional) desires.

To illustrate: Imagine two independent farmers, McDonald and McDougal, who work opposite ends of a valley. McDonald is more successful than McDougal, but not by exploiting him. However, since the two are quite competitive, each has a strong desire to perform better than the other. But each would prefer being the only farmer in the area to being outdone by the other. Given these competitive preferences, it seems that McDonald is prohibited by the proviso from employing his abilities to their fullest. This is because he prefers the outcome he then gets to that he would get in McDougal's absence and McDougal prefers the outcome he would get in McDonald's absence to that which he actually gets. And this means that McDonald's use of his powers and abilities better his own position and worsens the position of McDougal.²⁹

In *Morals by Agreement*, Gauthier seems to assume that the preferences that define the baseline for bargaining are non-tuistic — sometimes that they satisfy the stronger requirement of mutual unconcern (205). It might be thought that adherence to the assumption of non-tuism (in one form or another) would avoid counterexamples like the above, and so

²⁹ This case is presented, in more detail and with consideration of some responses, in Hubin & Lambeth, 498-9.

rescue the proviso as a tool in defining the initial bargaining position.³⁰ That is, it might be thought that the problematic counterexamples can be avoided by defining the 'bettering' and 'worsening' in relation to the utility function agents would have in the absence of tuistic desires. Such a result would require a broad interpretation of 'tuism' — for the above case, one that included positional desires. But, even were such an interpretation otherwise defensible, the cure is worse than the disease.

Consider the following case. Henry enslaves Beasley. Alas (for Henry), the slavery is not very profitable on most counts. Given custodial costs and diminished incentives for Beasley to labor to his utmost, Henry's net profit *in narrowly economic terms* is negative. What makes it worthwhile to Henry to continue the enslavement is his desire to thwart Beasley's desire to be free (or to make Beasley's life go poorly for him, or to be superior to Beasley in some feature like wealth, liberty, or whatever).

Allowing all desires to be considered in determining the utility levels of the two agents, and assuming that Beasley would be better off free, we can say that Henry violates the proviso; Henry makes himself better off than he would be in Beasley's absence by interacting with Beasley in a way that makes *him* worse off than he would be in Henry's absence. Well and good. But if we determine bettering and worsening of positions relative not to an agent's full set of preferences but only to his non-tuistic preferences, then trouble arises. Considering Henry's non-tuistic preferences only, he does not make himself better off than he would be in Beasley's absence — indeed, he makes himself worse off. But, then, on this interpretation, he doesn't violate the proviso, and he may legitimately take his possession of Beasley to the bargaining table for the social contract.

Gauthier wants the proviso to rule out what he calls 'freeloading' and 'parasitism.' The former consists in obtaining benefits without paying their costs, the latter in obtaining a benefit by displacing the cost onto another. Because costs and benefits are defined in terms of individual utility levels, what counts as parasitism and freeloading relative to all of the desires of the agents involved may no longer qualify as such when considered in relation only to a restricted set of their desires. If we ignore

30 Narveson clearly holds this position (505-6). His comments indicate that he takes the 'non-tuism' assumption to rule out desires for positional goods. (To avoid all of the counterexamples raised in 'Providing for Rights,' the non-tuism assumption would have to be interpreted even more broadly.) It is doubtful that Gauthier ever intended any interpretation except the first one discussed in the text — unconcern with the utility levels of others. For reasons discussed above, this will clearly not justify Narveson's maneuver.

tuistic desires in order to allow an agent to use his *own* talents to the fullest when doing so satisfies his tuistic desires, how are we to avoid allowing him to use the talents of others (against their will) when doing so satisfies his tuistic desires? Many answers could be given, but not by one who seeks, as Gauthier does, to found the right to one's own body and talents on the proviso. Given *this* project, it is unclear how to avoid the dilemma posed by these two cases.

Double-Counting

If we bargain and I am motivated only by concern for my own advantage while you are motivated by a positive concern for both of us, you are liable to get the worse end of the deal, it seems. This is because, in some intuitive sense, my advantage has been double-counted: once because of my own preferences and once because of your concern for me. Double-counting, it might be thought, is to be avoided.

Gauthier is tempted by this sort of rationale for the assumption of non-tuism.³¹ But it is a line that Christopher Morris has shown is not open to Gauthier.³² Morris's objection is that Gauthier's rejection of double-counting depends on an antecedent principle of fairness according to which each is to count for one and none for more than one — a principle to which Gauthier is not entitled if he means to provide us with a fundamental theory of justice.

There is another reason for denying Gauthier appeal to the argument of double-counting. It is that, on the only grounds Gauthier can admit, there is no objectionable double-counting. What does Gauthier take to be counted twice? Presumably, it is the utility of the person who is the object of the (positive) tuistic preferences. In one sense, of course, it is patently false that any person's utility is counted twice. Each person has her own utility function and the bargaining mechanism weighs these equally. However, in another sense, double-counting can clearly take place. If the object of the tuistic concern, Julia, values states of affairs in proportion to her wealth (or to her well-being) in them, and the tuist, Rachel, desires Julia's desire satisfaction, then increases in Julia's wealth (/well-being) get counted twice: once directly, based on her desires; and once indirectly, based on Rachel's desires. But this is because such increases satisfy the desires of two people, one directly and one indirectly. It is difficult to see how, given Gauthier's theory of value, this is

31 See, for example, the introductory comments to *Morals by Agreement* (11).

32 Morris, 138. See also Vallentyne's discussion.

more objectionable than any sort of case in which a state of affairs is 'double-counted' because it satisfies two person's desires. Furthermore, even if Julia and Rachel did present a case of double-counting that Gauthier had grounds for objecting to, this would not constitute a reason to exclude all tuistic preferences. It is certainly possible for a person to fail to take an interest in her own wealth (/well-being). In those cases, the tuistic interest another takes in it will not produce double-counting, and should not be ruled out under this proposed rationale.

The foregoing suggests that allowing tuistic desires does not produce a form of double-counting that Gauthier can recognize and reject. More important, as we shall see, is the apparent fact that even if it were to result in double-counting, one could not show that compliance with the social bargain is rational without allowing tuistic desires to influence the contract.

Non-tuism and Moral Foundationalism

As Morris notes in attacking the double-counting rationale for the assumption of non-tuism, Gauthier takes contractarianism to offer us a fundamental moral theory. Perhaps the foundational character of this project, which undermines the double-counting argument, simultaneously provides an independent argument for the assumption. Perhaps tuistic desires are incompatible with a contractarian theory being morally basic.

Consider the classic case of seemingly objectionable double-counting offered by Raymond Smullyan and discussed by Christopher Morris:

Once upon a time two boys found a cake. One of them said, "Splendid! I will eat the cake." The other one said, "No, that is not fair! We found the cake together, and we should share and share alike, half for you and half for me." The first boy said, "No, I should have the whole cake!" Along came an adult who said, "Gentlemen, you shouldn't fight about this: you should *compromise*. Give him three quarters of the cake."³³

Morris points out that there is a circularity problem here if the adult takes herself to be giving a basic account of just or fair division. After all, the division proposed by the second boy is already proposed as the fair solution. Whatever problems of double-counting the adult's solution might raise, his boast to be offering a foundational theory of justice

³³ *This Book Needs No Title* (Englewood Cliffs, NJ: Prentice Hall 1980), 56. Cited by Morris, 138.

or fairness raises a circularity problem so long as it is founded on a pre-existing moral evaluation.

Perhaps, then, the rationale for assuming non-tuism is to allow the social contract to do its assigned task of generating moral requirements from non-moral assumptions. If the theory needed to *assume* that people are morally motivated, then it at most offers a (possibly important) addendum to moral theory. It cannot be the foundation of morality. Indeed, one might argue, even to *allow* parties to the social contract to be morally motivated endangers contractarianism as a foundational moral theory. A foundational conception of contractarianism can tolerate the contractors having the motivation to 'comply with the terms of the contract, whatever they may be.' It cannot, it seems, tolerate the contractors having some more definite moral motivation. For the *contract* is to tell us what morality requires; the content of moral requirements cannot be known prior to the making of the contract. The idea is that the specific content of the moral motivation people feel in real life is to be explained by the contractarian methodology. For this to be possible, such motivation can't influence the contract itself.

But this approach to justifying the non-tuism assumption does not give us any rationale for rejecting tuistic desires that are not plausibly understood to be based on moral concerns. Negative tuistic desires present the most obvious example, but many others will do as well. Indeed, I think that much positive concern with others, even where it is morally desirable motivation, is not moral motivation.

In the above case, we supposed that the second boy's proposed division of the cake was based on a consideration of the desires of the first boy *independently* of the effects of satisfying those desires on the satisfaction of the second boy's desires. But suppose the case were different. Suppose that the second boy did not favor the even division as a fair or just resolution to their conflict. Suppose the second boy had a non-moral desire that the cake be shared equally, while the first boy has the non-moral desire that he have all the cake.³⁴ If we take the appropriate standard of value to be the subjective one Gauthier assumes, there seems to be no justification for dividing the cake equally.

34 We can suppose that the first boy's utility is linear with the amount of cake he receives and the second boy's is linear with the 'degree of equality' in the division of the cake, where 'degree of equality' could be defined as: $\text{MIN}(c^1, c^2) / \text{MAX}(c^1, c^2)$, where ' c^i ' is the fraction of the cake received by the i^{th} boy. We are to suppose that the latter boy's utility function is the result of a non-moral tuistic concern with the welfare of the other boy.

This is to force the first boy to make all of the concessions. Why should he rationally agree to this?

This suggests that, if the rationale for non-tuism is the necessity of a non-moral foundation for the social contract, the filter that Gauthier should use on preferences is not non-tuism (or mutual unconcern) but amorality. This proposal seems promising since, after all, morality is to be the *product* of the contract. Developing and defending this restriction, though, would require making clear the distinction between moral and non-moral (or amoral) desires. Furthermore, if the contractarian desires, as Gauthier does, to solve the rational compliance problem, she must also show that the exclusion of such moral motivation does not undermine the project of establishing that compliance with the terms of the contract is rational.

Non-tuism and the Rational Compliance Problem

Contractarians have different methodologies and different criteria of adequacy for their theories. While Rawls is primarily concerned that the results of his theory accord with our pre-theoretical convictions in reflective equilibrium, Gauthier's overriding concern is to solve the problem of rational compliance. He seeks to show that it is rational to comply with moral restrictions on action. This is not a new project, of course. But Gauthier is unique in that he accepts, *ab initio*, an individual-utility-maximizing conception of rationality, and argues that agents disposed to maximize their own utility would choose, on grounds of individual utility-maximization, to dispose themselves otherwise (given certain assumptions he claims are plausible). Morality is individual utility-maximization reining itself in for reasons of individual utility-maximization. It is, I think, fair to say that for Gauthier *the* criterion of adequacy of a contractarian theory is the establishment of the rationality of compliance with the outcome of the hypothetical bargain.³⁵

How does the assumption of non-tuism fare on this criterion? Poorly, I think. Regardless of whether we interpret non-tuism to preclude all concern for others, concern for our position vis-à-vis others, concern for their well-being, or concern for their level of desire satisfaction, the assumption of non-tuism undermines the goal of establishing the rationality of compliance with the social bargain.³⁶

³⁵ See, for example, his comments, cited above (454), from *Morals by Agreement* (269).

³⁶ Whether the same can be said for an assumption of non-moral motivation depends crucially on how one conceives such motivation. If all positive concern directed at

On the view that Gauthier accepts, it turns out that, for independent choice situations, actions are rational just in case they maximize the agent's expected utility, where utility is a measure of the agent's considered preferences.³⁷ He puts no restriction on the *content* of these preferences. By a clever and complex argument, he attempts to show that it would be rational *on these grounds* to dispose oneself to restrict the pursuit of one's utility maximization where outcomes are dependent in part on the choices of others. This is the cornerstone of his theory. It will not do for him to show that it *would* be rational for people to restrict their pursuit of individual utility maximization were they to have a set of desires different from what they have — for example, were they to have none of their tuistic desires.

Now imagine that people would agree to a certain set of institutions, rules or outcomes were they to set aside their tuistic concerns and bargain from a hypothetical initial situation that is defined by reference to the proviso (understanding 'bettering' and 'worsening' in relation only to non-tuistic desires). That is to say, imagine that these institutions, rules, or outcomes constitute the solution to the bargaining situation when it is defined in terms of the agent's non-tuistic preferences and the initial endowment is corrected, as Gauthier suggests, to adjust for the effects of freeloading and predation where these are understood relative to the agents' non-tuistic desires. Would agents who are rational in Gauthier's sense choose to dispose themselves to comply with these institutions? There seems to be no reason to suppose that they would generally.³⁸ If, on reflection, my overriding concern is that some foe's desires be thwarted, or that a friend flourish, or that I be better off in

others is conceived of as moral, then the assumption of non-moral motivation will pose the same difficulties for the rational compliance project as does the assumption of non-tuism. But there may be ways of construing the distinction between moral and non-moral motivation that will avoid this problem.

37 As Gauthier presents his views in *Morals by Agreement*, it is the expected utility of dispositions that determines the rationality of the actions they are dispositions to perform; an action is rational if it is in accordance with a disposition that would maximize expected utility for an ideal agent. But while much of his work is dedicated to showing that an ideal agent would have a reason to adopt a disposition not to maximize utility in *interdependent* situations, he believes that no such reason exists with respect to independent action. Thus, while even in independent choice situations actions are judged rational in terms of their conformance with rational dispositions, the rational disposition is to maximize utility in such situations.

38 Of course, it might turn out that a given institution enjoying such a sanction would be one with which we should rationally comply but this would not be simply because it represented the solution to the bargaining problem defined as in the text.

some respect than a competitor, why should I care about the terms of a bargain that I would have made had I lacked such desires?

Gauthier offers an answer in his recent reply to criticisms. He says:

As I now see it, social institutions and practices should be justified by an appeal to a hypothetical agreement based largely on the nontuistic preferences of the parties concerned, because each person expects *ex ante* to benefit if she forgoes the inclusion of her tuistic preferences in determining social arrangements provided others do the same. ('Morality, Rational Choice, and Semantic Representation,' 216)

But in what sense does each expect to benefit? Were Gauthier to allow a non-subjective sense of 'benefit' — in particular, one distinct from overall preference satisfaction — one might be able to argue that excluding one's tuistic preferences could be expected to be to one's benefit. But Gauthier has no such sense of 'benefit.' It is an easy trick, I would suppose, to show that *relative to one's non-tuistic preferences* one can expect to benefit from excluding tuistic desires from the bargain. But it is not a trick worth doing. What Gauthier must show is that *relative to one's full set of considered preferences*, one can expect to do better by excluding one's tuistic preferences from the social bargain.

Is this true? Consider an extreme case. Imagine a person whose only concerns were tuistic.³⁹ How could such a person possibly expect, *ex ante*, to secure a greater degree of preference satisfaction by bargaining with her tuistic preferences excluded? Such extreme examples are useful but unnecessary. Suppose that someone simply cared very deeply about the preference satisfaction of others in addition to caring about other things.⁴⁰ What reason does such a person have, *ex ante*, to expect that her preferences will be better satisfied by an agreement that would be made if she excluded her tuistic desires provided others did so as well? None, I think.

IV Conclusion

Many attacks have been launched against Gauthier's rational-compliance project. I am sympathetic with the critics. Gauthier has not argued successfully that it is always rational to act in accordance with a disposition that it is rational to adopt. Neither has he shown that it is rational

³⁹ Calvin Normore suggested this case to me.

⁴⁰ The same point can be made using the other senses of 'non-tuism.'

to dispose oneself to comply with the terms of a bargain one would make under the idealized assumptions of the bargaining situation for the social contract.

What I have tried to establish here is that the effect of assumptions of non-tuism is an exacerbation of this latter problem. To the degree that Gauthier 'assumes away' the values, concerns, and preferences that human beings actually have (or assumes values, concerns, and preferences they do not have), he undermines his project of establishing that it is rational to dispose oneself to comply with the social contract. And this is as true for positive concern for others as it is for negative concern. What reason have I to dispose myself to comply with a contract that would be made by someone who doesn't value what I value, hate whom I hate, love whom I love?

Received: July, 1990

Revised: March, 1991

Publication title: Canadian Journal of Philosophy

Volume: 21

Issue: 4

First page: 441

Number of pages: 28

Publication year: 1991

Publication date: Dec 1, 1991

Publisher: Canadian Association for Publishing in Philosophy.

Place of publication: Edmonton, Alta.

Country of publication: Canada

Publication subject: Philosophy

Publication frequency: Quarterly

ISSN: 0045-5091

Publication title history: Canadian Journal of Philosophy

Source type: Scholarly Journals

Language of publication: English

Document type: Article

ProQuest document ID: 1297846105

Document URL: <http://proxy.lib.ohio-state.edu/login?url=http://search.proquest.com/docview/1297846105?accountid=9783>

Last updated: 2013-02-23

Database: Periodicals Archive Online

Contact ProQuest

Copyright © 2014 ProQuest LLC. All rights reserved. - [Terms and Conditions](#)