Act and Principle Contractualism

HANOCH SHEINMAN
Rice University

Unrejectability is the property shared by things no one could reasonably reject. Following the lead of T. M. Scanlon, modern contractualists hold *Principle Contractualism*: An act is obligatory when conformity to unrejectable principles requires its performance. This article entertains *Act Contractualism*: An act is obligatory when its performance is unrejectable. The article hypothesizes that Principle Contractualism owes its initial plausibility to the assumption that following it somehow realizes unrejectability, if only indirectly. The article then argues that, whereas following Act Contractualism realizes unrejectability, following Principle Contractualism realizes a convoluted, principle-mediated, non-causal conformity relation between acts and unrejectability. But then the notion that this relation is what matters ultimately in action does not seem to enjoy independent plausibility. After interrogating Scanlon's objection that the challenge to the principle-based nature of contractualism is 'misconceived', I conclude that Act Contractualism is the more fitting contractualist theory of obligation.

When we apply some Contractualist formula, such as the Kantian and Scanlonian Formulas, we don't consider particular acts on their own. We ask which are the principles that everyone could rationally choose, or that no one could reasonably reject.

Derek Parfit, *On What Matters*

Note that whereas consequentialists have on the whole embraced the single-level act-consequentialism at the expense of the two-level rule consequentialism, contractualists have been virtually unanimous in their preference for two-level versions of contractualism; indeed, single-level contractualism is virtually unheard of. One consequence of it is that, whereas rule-consequentialists have been forced to carefully formulate and defend rule consequentialism, the two-level character of contractualism has received little in the way of critical scrutiny.

Nicholas Southwood, *Contractualism and the Foundation of Morality*

I. INTRODUCTION

Modern contractualists claim that the moral status of acts is determined by whether they conform to principles that are justifiable to each person or *contractually* justifiable, and that principles are contractually justifiable when no one could reasonably reject them as public standards of behavior – in short, when they are *unrejectable*. Here are some examples.

© Cambridge University Press 2011
Utilitas Vol. 23, No. 3, September 2011
doi:10.1017/S095382081100015X

https://doi.org/10.1017/S095382081100015X Published online by Cambridge University Press

T. M. Scanlon describes judgments of right and wrong as 'judgments about what would be permitted by principles that could not reasonably be rejected'; 'when we address our minds to a question of right and wrong, what we are trying to decide is, first and foremost, whether certain principles are ones that no one, if suitably motivated, could reasonably reject'.[1] According to Stephen Darwall, it 'is a hallmark of contractualist theories that they hold principles of right to have a distinctive *role*'; we are committed to regulate 'our conduct by principles that are acceptable, or not reasonably rejectable, to each as free and rational agents'.[2] Gary Watson holds that practical 'reason expresses itself as a fundamental commitment to act in accordance with principles to which all rational beings could agree'; actions are right or wrong 'if they are allowed or disallowed by principles that it would be unreasonable to reject'.[3] Thomas Nagel has 'adopted the central feature of Scanlon's account of contractualism – the idea that the right principles to govern a practice are those which no one could reasonably reject, given the aim of finding principles which could be the basis of general agreement among persons similarly motivated'; '[a] course of action is prohibited . . . if and only if every universal rule of conduct which would permit it falls within [the] class of rejectable principles'.[4] According to Nicholas Southwood's 'deliberative contractualism, morality's foundations are to be located in facts about what common code we would agree to live by if we were perfectly *deliberatively rational*'; '[b]y "a common code" is meant a relatively comprehensive set of common principles, permitting, forbidding, and requiring certain conduct in certain circumstances. By "principles" is meant anonymous, conclusive prescriptions.'[5] Finally, Derek Parfit has recently adopted both Kantian Contractualism:

> Everyone ought to follow the principles whose universal acceptance everyone could rationally will, or choose.

and Scanlonian Contractualism:

[1] *What We Owe to Each Other* (Cambridge, Mass., 1998), pp. 4, 189.

[2] *The Second-Person Standpoint* (Cambridge, Mass., 2006), pp. 301, 300.

[3] 'Some Considerations in Favor of Contractualism', *Rational Commitment and Social Justice*, ed. J. Coleman and C. Morris (Cambridge, Mass., 1998), pp. 174, 176.

[4] *Equality and Partiality* (Oxford, 1991), pp. 36–37. Again, Nagel approvingly quotes Scanlon's general claim that '[a]n act is wrong if its performance under the circumstances would be disallowed by any system of rules for the regulation of behavior which no one could reasonably reject as a basis for informed, unforced general agreement' (36). Incidentally, Nagel treats Scanlon's general claim (an act is wrong iff some *principle* forbidding it is *un*rejectable) as equivalent to the claim that an act is wrong iff every principle *permitting* it is rejectable. But as Shelly Kagan notes, Scanlon's formula prohibits less and is preferable. 'The Unanimity Standard', *Journal of Social Philosophy* 24 (1993), pp. 129–54, at 148–9.

[5] *Contractualism and the Foundations of Morality* (Oxford, 2010), pp. 88, 102.

https://doi.org/10.1017/S095382081100015X Published online by Cambridge University Press

Everyone ought to follow the principles that no one could reasonably reject.

Parfit writes:

> When we apply some Contractualist formula, such as the Kantian and Scanlonian Formulas, we don't consider particular acts on their own. We ask which are the principles that everyone could rationally choose, or that no one could reasonably reject, if we were choosing the principles that everyone would accept.[6]

Modern contractualists then conceive of the moral status of acts in terms of unrejectable principles; they hold a principle-based conception of right or obligation. Specifically, modern contractualism is distinguished by an asymmetric two-stage structure that gives principles justificatory priority. At the *primary* justificatory stage, we justify some principles by showing that no one could reasonably reject them as public standards of behavior.[7] At the *derivative* justificatory stage, we justify some act by showing that it conforms to these justified principles. The act derives its contractualist justifiability from that of the principles but not the other way around.

To be sure, modern contractualists do talk about contractually justifiable or unrejectable acts. However, such talk is meant as shorthand for talk about acts that conform to contractually justifiable or unrejectable principles. Modern contractualists never ask whether an act is unrejectable in the primary sense of the expression; they always ask whether it is unrejectable in the derivative sense of conformity to unrejectable principles. Unless otherwise indicated, I will use 'unrejectable' in the primary sense.

Modern contractualism and rule-consequentialism: difference

Modern contractualism has developed as a reaction to consequentialism. Most pertinent to our purposes is *rule-consequentialism*, the view that the moral status of acts is determined by whether they conform to rules that are *consequentially* justifiable, and that rules are consequentially justifiable when their general acceptance as public standards of behavior would promote the good – in short, when they are *optimific*. Assuming that 'rules' and 'principles' are co-referential, the difference between modern contractualism and rule-consequentialism is that between contractually and consequentially

[6] *On What Matters*, vol. 1 (Oxford, 2011), pp. 355, 360, 363–4, §§ 52–3. Parfit now endorses 'Triple Theory', in which Kantian and Scanlonian contractualism are combined with rule consequentialism (pp. 412–13, § 64).

[7] It is a peculiar feature of modern contractualism that it typically defines justification as unrejectability rather than as acceptability. Compare Nagel, *Equality and Impartiality*, p. 36 with Kagan, 'The Unanimity Standard', pp. 144–7. The crucial point, however, is that the unrejectability/acceptability in question is of *principles*.

https://doi.org/10.1017/S095382081100015X Published online by Cambridge University Press

justifiable principles. The former are principles no one could reasonably reject as public standards of behavior even if their general acceptance as such standards would fail to promote the overall good; the latter are principles whose general acceptance as public standards of behavior would promote the overall good even if someone could reasonably reject them as such standards.

Modern contractualists charge that in justifying principles by the promotion of the overall good, rule-consequentialism 'attempts to derive principles of right from reasons of the wrong kind'.[8] For their part, consequentialists object that modern contractualism cannot avoid appealing to the overall good.[9]

I will assume that modern contractualists win the debate. I will assume that some moral justification is distinctly contractualist and genuinely non-consequentialist. Without getting into the question of what makes principles unrejectable, I will simply assume that principles can be unrejectable (and so contractually justifiable) without being optimific (and so consequentially justifiable), and vice versa.[10] And I will assume that moral justification is contractualist throughout. The principles to accept as public standards are those that are unrejectable, whether or not they promote the overall good.

We make these assumptions in favor of contractualism to isolate our question: Is modern contractualism the most fitting *contractualist* theory of obligation? If not, what is?

Modern contractualism and rule-consequentialism: similarity

So much for the difference between modern contractualism and rule-consequentialism. More important for present purposes is their similarity. They both have an asymmetric two-stage structure that gives rules justificatory priority. Rule-consequentialism is famous for this two-stage structure. The standard contrast is *act-consequentialism*,

[8] Darwall, *The Second-Person Viewpoint*, p. 312.

[9] Consider the choice between principles P_1 and principles P_2 that cannot be generally accepted together. Assume that person A rejects P_1 because their general acceptance harms A, that person B rejects P_2 because *their* general acceptance harms B, and that others are indifferent as between these principles. Suppose finally that exactly one of these rejections is reasonable. Which one is it? Surely, some comparison between A's and B's harms is called for. Arguably, A's rejection of P_1 is reasonable just when their general acceptance harms A more than the general acceptance of P_2 harms B. Cf. Scanlon, *What We Owe*, p. 195; *The Difficulty of Tolerance* (Cambridge, Mass., 2003), p. 135. But if the notion of unrejectability imports interpersonal comparison, it is only fair to ask how it can fail to render contractualist justification – justifiability to each person – consequentialist. Cf. Alastair Norcross, 'Contractualism and Aggregation', *Social Theory & Practice* 28 (2002), pp. 303–14; Derek Parfit, 'Justifiability to Each Person', *On What We Owe to Each Other*, ed. P. Stratton-Lake (Oxford, 2004).

[10] For simplicity's sake, I will also assume that there is always exactly one set of unrejectable principles.

https://doi.org/10.1017/S095382081100015X Published online by Cambridge University Press

which proceeds in just one justificatory stage. According to act-consequentialism, the moral status of acts is determined by whether they are consequentially justifiable, and acts are consequentially justifiable when they promote the overall good – in short, when they are optimific. Act-consequentialists famously complain that the justificatory primacy rule-consequentialists give rules opens them to the charge of rule worship.[11] Modern contractualists second this charge.[12]

You might expect contractualist theories of right to exhibit a similar division. You might expect some contractualists to be *act contractualists*; others, *principle contractualists*. And you might expect the former to complain that the justificatory primacy the latter give principles opens them to a charge of principle worship. This is not what you find. As we have seen, modern contractualists are uniformly principle contractualists.[13] The assumption seems to be that there is a fundamental disanalogy between principle contractualism and rule-consequentialism: while rule-consequentialism has a rule-free consequentialist alternative (act-consequentialism), principle contractualism has no principle-free contractualist alternative. Given this fundamental disanalogy, the rule worship charge does not carry over to the contractualist arena.

I will argue that this assumption is mistaken. Modern contractualism is every bit as guilty of principle worship as rule-consequentialism is of rule worship. Here too the culprit is the justificatory primacy of principles. Here too the most fitting theory of right proceeds in a single justificatory stage, without reference to principles.

II. PRINCIPLE CONTRACTUALISM

Scanlon's moral theory is a paradigm of modern contractualism. We can summarize it as

Principle Contractualism

A has an obligation to do X (and A's failure to do X is wrong) just when principles no one could reasonably reject as public standards of behavior (*unrejectable*

[11] See, for example, J. J. C. Smart, 'Extreme and Restricted Utilitarianism', *Philosophical Quarterly* 6 (1956), pp. 34–54.

[12] Scanlon, *The Difficulty of Tolerance*, p. 27; Watson, 'Some Considerations in Favor of Contractualism', pp. 172–3.

[13] See Shelly Kagan, *Normative Ethics* (Boulder, 1998), p. 243, and Southwood, *Contractualism and the Foundations of Morality*, p. 102.

https://doi.org/10.1017/S095382081100015X Published online by Cambridge University Press

principles) require A do X (equivalently: just when A's failure to do X violates such principles).[14]

Additionally/alternatively:

A's doing X is right (permissible) just when A's doing X conforms to unrejectable principles.

Here I will assume that Principle Contractualism includes both prongs. We take on board Scanlon's assumption that Principle Contractualism is a distinctly *contractualist* theory, and that the notion of unrejectability has a genuinely non-consequentialist content: principles might be unrejectable but non-optimific (overall bad); rejectable, but optimific (overall good). This is a consequence of Scanlon's *individualist requirement*, the 'insistence that the justifiability of a moral principle depends only on various *individuals'* reasons for objecting to that principle and alternatives to it'.[15]

Principle Contractualism is not just any contractualist theory of obligation, however, but one that gives principles justificatory priority over acts. Acts are justifiable (unrejectable) derivatively, by their conformity to independently justifiable (unrejectable) principles. What we are trying to decide when we face a question of right and wrong 'is, first and foremost, whether certain principles are ones that no one, if suitably motivated, could reasonably reject'. The role of principles on this theory 'is fundamental; they do not enter merely as devices for the promotion of acts that are right according to some other standard'.[16] That is to say, principles are no mere 'rules of thumb', useful *indicators* of acts that are independently right or wrong, which must be set aside as soon as we know all the facts about the acts themselves. Instead, principles determine the moral status of acts: they make it the case that an act has the moral status it has.[17]

Related to this principle-based nature of Principle Contractualism is Scanlon's *generality requirement*: principles can only be justified by *generic reasons*, reasons 'we can see people have in virtue of certain general characteristics'. Hence

an assessment of the rejectability of a principle must take into account the consequence of its acceptance in general, not merely in a particular case . . . Since we cannot know, when we are making this assessment, which particular

[14] *What We Owe*, pp. 4, 153, 195; *Moral Dimensions* (Cambridge, Mass., 2008), p. 99; 'How I Am Not a Kantian', in D. Parfit, *On What Matters*, vol. 2 (Oxford, 2011), pp. 116–39.

[15] *What We Owe*, p. 229. But cf. Scanlon's acknowledgment that reasonable (un)rejectability is comparative. *What We Owe*, p. 195; *The Difficulty of Tolerance*, p. 135.

[16] *What We Owe*, p. 189, *The Difficulty of Tolerance*, p. 142.

[17] But cf. n. 39.

https://doi.org/10.1017/S095382081100015X Published online by Cambridge University Press

individuals will be affected by it in which ways . . . , our assessment cannot be based on the particular aims, preferences, and other characteristics of specific individuals. We must rely instead on commonly available information about what people have reason to want.[18]

Scanlon's individualist and generality requirements are natural corollaries to the twin cornerstones of Principle Contractualism. The individualist requirement is integral to the notion that principles are justifiable to each person or *contractually*; the question is whether the principles are unrejectable, not whether they promote the overall good. The generality requirement on the other hand is integral to the primacy of *principles*: acts are justified by their conformity to unrejectable principles (rather than by their own unrejectability). The generality requirement is not an essential feature of contractualist moral theories but of Principle Contractualism. Indeed, the 'particular aims, preferences, and other characteristics of specific individuals', which might well be relevant to whether one should do as some principles require on some particular occasion, seem irrelevant to whether these principles are justified *as public standards of behavior*.

To think otherwise is to conflate the *practical* justification of principles as public standards of behavior (or the justification of their practical acceptance) with the *theoretical* justification of these principles as independently true claims (or the justification of their theoretical acceptance). Consider first the question of whether we are justified in accepting some principle P as an independently true moral principle. To accept P as true is to believe that, whenever P requires some particular act X, X is obligatory (and its non-performance wrong). One can reasonably reject P as false by focusing on the details of a single case, however peculiar. In so doing, one could refute the principle by counterexample, namely a single case in which the principle gives the wrong verdict, even if the principle gives the right verdict in all other cases.

None of this makes good sense when it comes to the practical justification of principles as public standards of behavior (or the justification of their practical acceptance). Here the question is not whether P is true (has no counterexample), but whether we, as a society, should accept P as a public standard of behavior – whether we should *institute* P, *make* it a social principle, *weave it into* our practice. There are different ways to accept principles as public standards of behavior. We can make them into law. Or we can promulgate, teach, and enforce them in various other ways. But regardless of the particular way in which principles are accepted as public standards, one cannot

[18] *What We Owe*, pp. 204–5.

https://doi.org/10.1017/S095382081100015X Published online by Cambridge University Press

reasonably reject them as such standards just by pointing to one person who can reasonably violate them on one abnormal occasion. Here Scanlon's generality requirement resonates with the Aristotelian wisdom that it is eminently reasonable to accept some principles as public standards of behavior even though any principles would occasionally require us to do unreasonable acts.[19]

III. ACT CONTRACTUALISM

The justificatory primacy of principles under Principle Contractualism raises the possibility of a contractualist theory of obligation that gives principles no justificatory role. Consider

Act Contractualism

A has an obligation to do X (and A's failure to do X is wrong) just when no one could reasonably reject A's doing X as a particular instance of behavior (in short, when it is *unrejectable*).[20]

Alternatively/additionally:

A's doing X is right (permissible) just when A's doing X is unrejectable as a particular instance of behavior.

Here I will assume that Act Contractualism includes both prongs.[21] Like Principle Contractualism, Act Contractualism is a distinctly contractualist, non-consequentialist theory of obligation. The key concept of justifiability to each person on grounds no one could reasonably reject is equally at work here. Nevertheless, this is a very different theory of obligation. Most notably, justification under Act Contractualism proceeds in one stage. The moral status of an act is determined by its own unrejectability, independent of that of principles. Now unlike principles, acts cannot be (un)rejectable *as public standards of behavior*; indeed, acts are not standards. But if we can reject principles as public standards of behavior, I do not see why we cannot also reject acts *as individual instances of behavior*.

Isn't Act Contractualism trivial? 'Of course,' so goes the thought, 'if one has an obligation to do some act, no one could reasonably reject it. And of course, if no one could reasonably reject one's act, one has an obligation to do it!' But this is not obvious at all. First, I suspect that the appearance of triviality conflates unrejectability with obligatoriness

[19] Cf. Aristotle. *Nicomachean Ethics*, trans. T. Irwin (Indianapolis, 1999), i.3.4, ii.2.3.

[20] The possibility of an act-contractualist theory is raised in Kagan, *Normative Ethics*, p. 242.

[21] On this assumption, Act Contractualism implies that an act is right just when it is obligatory. Most act-consequentialists hold this view, but it is equally available to the Act Contractualist.

https://doi.org/10.1017/S095382081100015X Published online by Cambridge University Press

(or rejectability with wrongness). But as Derek Parfit points out, this 'overlooks the fact that, when we apply some Contractualist formula, we cannot appeal to our beliefs about what acts are wrong'.[22] After all, Act Contractualism purports to explain why acts are obligatory/wrong. Rather, in applying the Act Contractualist formula we appeal to 'non-deontic' facts about the evaluated acts, for example, the fact that, compared to the harm imposed on one person by that act, the harm imposed on another person by the only alternative act is much greater. And second, Act Contractualism is incompatible with both act-consequentialism and Principle Contractualism. It is incompatible with act-consequentialism because, as per our assumption, unrejectability is a distinctly contractualist, non-consequentialist property. More to the point, it is incompatible with Principle Contractualism. To see how, notice that Act Contractualism cannot abide by Scanlon's generality requirement. As I have argued, the plausibility of this requirement depends on its confinement to the justifiability or unrejectability of principles. When it comes to the justifiability or unrejectability of particular acts, the insistence that 'our assessment cannot be based on the particular aims, preferences, and other characteristics of specific individuals'[23] loses all plausibility. Suppose that A performs some act that is G. Such acts are nearly always harmless, yet owing to B's eggshell skull, A's act is seriously harmful to B. Surely insofar as we are interested in the unrejectability of A's act, apart from its relation to any principles, the sheer rarity of B's condition is neither here nor there. Of all the people in the world, A's act affects B in particular. B's special vulnerability seems as relevant to the act's justifiability as anything.

So while Principle Contractualists must embrace the generality requirement, Act Contractualists must reject it. The upshot is incompatibility. An act whose non-performance violates unrejectable principles, and is for that reason obligatory under Principle Contractualism, might be rejectable, and for that reason wrong, under Act Contractualism. To deny this is to assume that the acts unrejectable principles require are themselves unrejectable. But it is far from clear that Principle Contractualists accept this assumption. The principles they typically advocate as unrejectable do not seem to make the (un)rejectability of acts a condition of obligation.[24]

[22] *On What Matters*, vol. 1, p. 360 (§ 53). Parfit calls this the 'Deontic Beliefs Restriction'. He makes the point in relation to Principle Contractualism, but the point applies with equal force to Act Contractualism.

[23] *What We Owe*, pp. 204–5.

[24] A notable example is Scanlon's Principle of Fidelity, discussed in section VI.

https://doi.org/10.1017/S095382081100015X Published online by Cambridge University Press

All I have done so far is carve out conceptual space for Act Contractualism. But, as we have seen, modern contractualists do not hold this position; they are Principle Contractualists. So we have discovered a peculiar position in conceptual space: big deal! But wait. The structural difference between Act and Principle Contractualism exactly parallels the one between act- and rule-consequentialism. And we have already noted that rule-consequentialism is commonly thought to be vulnerable to a rule-worship objection. Can we not say something similar about Principle Contractualism?

I think we can. The two-stage justificatory structure of Principle Contractualism generates a problem of worship, which does not arise under the single-stage structure of Act Contractualism. Far from a mere conceptual possibility, Act Contractualism is the more fitting contractualist theory of obligation. My argumentative strategy will be to identify the *foundation* of contractualism – the one idea that makes most sense of the project of constructing a contractualist theory of right action and to which both Act and Principle Contractualism seem to owe their initial plausibility – and argue that only the former is a genuine manifestation of that core idea.

IV. FOUNDATIONAL CONTRACTUALISM

What is that core idea? It trivially includes the view that what matters ultimately in action is to be understood *in terms of* justifiability to each person or unrejectability. On reflection, however, this view is too weak to motivate the project of constructing a contractualist theory of right action. Thus consider the *countercontractualist*, *devil contractualist*, and *devil principle contractualist* for whom what matters ultimately in action is rejectability, unrejectability in a society of devils, and conformity to principles that are unrejectable in a society of devils, respectively. These characters understand the ultimate moral point of action in terms of unrejectability alright (and so are technically contractualist), but their attitudes do not come close to capturing what is appealing about the contractualist project. Let us say that some property or value V is *what matters ultimately in action* (or is *action's ultimate moral point*) when

(1) V has a non-derivative, intrinsic moral significance for how we should act, and
(2) V takes priority over all other values or properties in our action.

To make this definition clearer, we can imagine an agent who is wholeheartedly committed to the notion that V is action's ultimate moral point. Such an agent treats V as of intrinsic moral significance. Presumably, she seeks to realize V in her action for its own sake. But

https://doi.org/10.1017/S095382081100015X Published online by Cambridge University Press

this is not all, for she might treat other properties or values in the same way. To treat V as what ultimately matters in action she must also give V priority in deciding how she acts. When the realization of V and other values are incompatible, she chooses to realize V. Then I hypothesize that the motivating idea behind contractualist theories of right is:

Foundational Contractualism

What matters ultimately in action is unrejectability.[25]

Foundational Contractualism is the simplest and most plausible contractualist answer to *foundational consequentialism* – the view that what matters ultimately in action is goodness. What could the ultimate moral point of action possibly be if not that things go well? Just this: that things be unrejectable! To be sure, any theory that understands action's ultimate moral point in terms of unrejectability is technically contractualist. But Foundational Contractualism is not just any old foundationally contractualist theory; it is an independently plausible view. And it might well be the motivating idea behind contractualist theories of obligation. If the moral point of action is to be put in terms of unrejectability, the simplest and most plausible candidate for that point is unrejectability itself. It seems to me that Principle Contractualism borrows much of its plausibility from the tacit assumption that it somehow reflects Foundational Contractualism, if only indirectly. It seems to enjoy plausibility by association.

Foundational Contractualism hardly entails Principle Contractualism. In particular, it is clearly compatible with Act Contractualism. Discharging our obligations under Act Contractualism is guaranteed to realize unrejectability in this world by producing unrejectable acts. Is this not exactly what Foundational Contractualism commits us to want?

Now, Act Contractualism is a theory of acts of individuals, but it might make sense to extend it to collective acts we perform as a society, including the acceptance/institution of principles as public standards of behavior. So extended, Act Contractualism requires our society to accept/institute principles just when such acceptance/institution is unrejectable. So here too Act Contractualism requires us to realize unrejectability, this time by populating the world with unrejectable collective acts of principle acceptance.

So Act Contractualism directly reflects Foundational Contractualism, an independently plausible theory of action's ultimate moral point.

[25] Cf. e.g. Scanlon's claim that 'taking what can be justified to others – what they have reason to will [or could not reasonably reject] – as the most fundamental moral idea is the essence of contractualism' ('How I am Not a Kantian', in Parfit, *On What Matters*, vol. 2, p. 139).

https://doi.org/10.1017/S095382081100015X Published online by Cambridge University Press

I will now try to show that this is more than can be said for Principle Contractualism.

V. THE FOUNDATION OF PRINCIPLE CONTRACTUALISM

At least before Act Contractualism is put on the table, it is natural to believe that Principle Contractualism reflects Foundational Contractualism, albeit indirectly. Since acts are justified by their conformity to unrejectable principles, unrejectability is the source of justification:

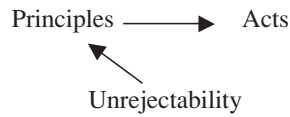


Figure 1. Principle Contractualism (arrows representing justification)

What we have here is a chain of justification with unrejectability as a kind of first mover unmoved: everything traces its justification to unrejectability (which traces its justification to nothing else). Put differently, unrejectability is where justification ends. Therefore, one might naturally conclude, unrejectability is what matters ultimately in action!

But this natural thought is equivocal. It is true that unrejectability under Principle Contratualism is what matters ultimately in action in the sense that acts trace their justification ultimately to unrejectability in the way represented by the diagram: justification ends with unrejectability. But the sheer fact that unrejectability is the *justificatory terminus* does not mean that it is what matters ultimately in action in the sense of Foundational Contractualism. To treat unrejectability as what matters ultimately in action in the relevant sense is to seek to realize this value for its own sake and give it priority in one's action. What value does the Principle Contractualist seek to realize for its own sake and prioritize in his action, exactly?

One thing is clear: unrejectability it is not. Let me explain. For a start, we can say that Principle Contractualism requires the realization of some justification relation R between acts and unrejectable principles. Now, does someone who seeks to realize and prioritize R in her action also treat unrejectability as what matters ultimately in action? Not necessarily. Suppose that R could be realized without realizing the property of being unrejectable, to wit without making it the case that some act, principle, or anything else for that matter is in fact unrejectable – without adding a single unrejectable to this world. Should you be forced to choose between realizing R and realizing unrejectability, Principle Contractualism would instruct you to realize

https://doi.org/10.1017/S095382081100015X Published online by Cambridge University Press

R. In so choosing, you would neither seek to realize unrejectability nor prioritize it in your action.

But the supposition is in fact true. Since R is a conformity relation, it can be realized without realizing unrejectability, however indirectly. What unrejectability could conformity to unrejectable principles possibly realize? Consider the possibilities. Obviously enough, such conformity does not necessarily add (i) unrejectable *acts* to our world; in the case the previous paragraph describes, it requires you to forego the unrejectable act. So perhaps it adds (ii) unrejectable *principles*? But principles as such are propositions, and propositions are just not the kind of thing that acts can bring into or out of existence. (The notion that conformity adds principles is as plausible as the notion that nonconformity subtracts them.) Perhaps the most relevant possibility is that conformity to unrejectable principles adds (iii) *societal acceptance/institution* of principles as public standards of behavior. Alas, to add socially accepted principles is to cause their societal acceptance/institution, and conformity is a non-causal relation. Mere conformity to principles does not tend to make them generally accepted in society.[26]

What matters ultimately in action under Principle Contractualism is not unrejectability of acts, principles, or societal principle acceptance, but *conformity* to unrejectable principles. In this respect, Principle Contractualism is perfectly analogous to rule-consequentialism. Under rule-consequentialism, acts trace their justification ultimately to goodness (via rules). But this hardly entails foundational consequentialism. That view – the core idea to which consequentialist theories of action seem to owe their initial plausibility – is that what matters ultimately in action is goodness, where this is understood as the claim that goodness has intrinsic moral significance and enjoys priority in action. In fact, what matters ultimately in action under rule-consequentialism is conformity to optimific rules.[27]

The notion that Principle Contractualism treats unrejectability as what matters ultimately in action in the relevant sense is just as natural and mistaken as the notion that rule-consequentialism treats goodness in this way. But by acquiescing in the standard talk of

[26] The other side of the coin is that mere nonconformity to principles has no tendency to make them generally unaccepted in society.

[27] Shelly Kagan has offered me the following objection. Suppose what matters ultimately in action is satisfying God's will, and God's will is conflicted: He wants us to conform to the Ten Commandments, but he also wants us to do some acts that happen to violate them. Doesn't the theory that requires conformity to the Ten Commandments reflect our fundamental commitment just as well as the theory that requires performance of the God-wanted acts? Yes, but only because following these theories equally satisfies/fail to satisfy God's will. By contrast, the doing of good/unrejectable acts, conformity to good/unrejectable rules realizes no goodness/unrejectability.

https://doi.org/10.1017/S095382081100015X Published online by Cambridge University Press

unrejectable principles I have in fact understated the case. Such talk, which has been accurate enough for our purposes so far, makes Principle Contractualism and unrejectability look closer than they actually are. It creates the impression that, while what matters under Principle Contractualism is not unrejectability proper, it is at least conformity to principles that are unrejectable proper, in the primary sense of the word. Surely, while Principle Contractualism does not justify acts non-derivatively, at least it justifies principles non-derivatively, by their own unrejectability? Figure 1 (see p. 299) illustrates this standard picture, under which the moral status of acts is only once removed from unrejectability. This impression is false, however. In truth, not only does Principle Contractualism justify acts derivatively; it also justifies principles derivatively. Principle Contractualism justifies principles by their relation to contractually justifiable, unrejectable collective acts (or events). Specifically, principles are contractually justifiable by virtue of the fact that the societal *acceptance*/*institution* of these very principles as public standards of behavior is unrejectable. Standard talk of unrejectable principles is merely shorthand for talk of unrejectable principle acceptance. The moral status of an act is really twice removed from unrejectability proper:

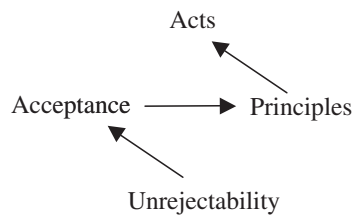


Figure 2. Principle Contractualism (arrows representing justification)

I have said that what really matters ultimately in action under Principle Contractualism is conformity to 'unrejectable principles'. We are now in a position to provide a more accurate description of this putative value. It is neither unrejectability nor unrejectable principles, but some three-place, principle-mediated relation to unrejectability, namely the relation that holds between an act X, principles P, and a collective act/event E just when

(1) The failure to do X violates P.
(2) Were E to occur, P would become socially accepted public standards of behavior.
(3) E is unrejectable.

https://doi.org/10.1017/S095382081100015X Published online by Cambridge University Press

Conditions (2)–(3) gloss the claim that principles P are unrejectable in a way that removes the false impression that they are unrejectable proper.

We sum up the discussion by making the foundational commitment of Principle Contractualism explicit. Rather than indirectly reflecting Principle Contractualism, Principle Contractualism directly reflects

Foundational Principle Contractualism

What matters ultimately in action is conformity to principles whose societal acceptance/institution as public standards of behavior is unrejectable.

But Foundational Principle Contractualism hardly seems to enjoy anything like the independent plausibility of Foundational Contractualism. In fact, as soon as the foundational nature of Principle Contractualism comes to light – as soon as we look at it as a theory about what matters ultimately in action, about action's ultimate moral point – it becomes implausible in a rather obvious way. It may well be natural to think that what matters ultimately in action is unrejectability, that action's ultimate moral point is to realize unrejectability in this world, making things unrejectable. But just how natural is it to think that what matters ultimately in action is conformity – *any* conformity? Why would any resident of this world ever forego unrejectability in the actual world for the sake of conformity to principles whose societal acceptance/institution in some other possible world is unrejectable? Why would anyone who thinks unrejectability is important enough to serve as action's justificatory terminus ever care about conformity, except when it is unrejectable? Why isn't the belief that rejectable conformity can ever trump unrejectable nonconformity irrational? Why is it not rather like the instrumentally irrational thought that a means that fails to achieve its end might nevertheless trump a means that achieves it – so long as it conforms to principles whose societal acceptance/institution achieves that end in some other possible world?

Thus the principle-worship objection to Principle Contractualism is best understood as a dilemma: Either Principle Contractualism is simply a theory of obligation that purports to reflect an independently plausible theory of action's ultimate moral point (Foundational Contractualism), in which case it fails on its own terms; or else it is its own foundational theory (Foundational Principle Contractualism), in which case it lacks independent plausibility. Either way, Principle Contractualism seems guilty of principle worship – or conformism.

https://doi.org/10.1017/S095382081100015X Published online by Cambridge University Press

VI. ACT AND PRINCIPLE CONTRACTUALISM ILLUSTRATED

We can illustrate the conflict between Act and Principle Contractualism by considering Scanlon's theory of promissory obligations. The specific part of this theory is a complicated principle I shall summarize as the

Principle of Fidelity (PoF)

A has an obligation to do what A leads B to expect A to do.[28]

Crucial for our purposes is Scanlon's claim that PoF is unrejectable: a principle no one could reasonably reject as a public standard of behavior.[29] Let us accept this claim, for argument's sake. Now consider the proverbial desert island deathbed promise.

Desert Island: You and Bo, the only survivors of a shipwreck, find yourselves on a desert island. Bo is dying, but you will be rescued after his death. Bo asks you to promise him to donate $5,000 to his rich golf club immediately upon returning to the mainland. You sincerely promise to do so. Upon returning to the mainland, your poor daughter Jo asks you for $5,000 to pay for a language course. Unfortunately, $5,000 is all you have available at the moment. No one but you knows about your deathbed promise to Bo, who is no longer alive.

Insofar as the choice between Act and Principle Contractualism is concerned, Desert Island seems to be as good a test case as any. These theories seem to generate conflicting verdicts in this case, for it seems to involve precisely the kind of considerations that are relevant to the rejectability of acts but not to that of principles, including the fact that the promisee is dead, the third-party beneficiary of the promise (the rich club) does not know about it (and so does not expect to get what's promised), and your promise-breaking will help someone else (your poor daughter) considerably more than your promise-keeping would help the beneficiary (the $5,000 will help your daughter much more than the club).

Clearly, since you led Bo to expect you to donate the money to his club, PoF requires you to keep your promise. And we are assuming that PoF is unrejectable: no one – not even Jo or you – could reasonably reject PoF as a public standard of behavior. Therefore, Principle Contractualism seems to require you to keep your promise. However, if we can assume that PoF is unrejectable as a public standard of behavior, we can also assume that your conformity to PoF is rejectable as a particular instance of behavior. Indeed, it does not seem implausible to maintain

[28] Scanlon, *What We Owe*, p. 304.
[29] *What We Owe*, pp. 304–9. Scanlon writes as if PoF is individually unrejectable. I think Principle Contractualism is best understood in terms of *jointly* unrejectable principles. I will conveniently gloss over the distinction.

https://doi.org/10.1017/S095382081100015X Published online by Cambridge University Press

that Jo and/or you could reasonably reject your promise-keeping act in this case. Jo could point out, for example, that she needs the money more than the club (recall that Jo is poor, the club is rich, and the marginal utility of money is diminishing), that, unlike the club, she has a standing expectation that you help her, and that, unlike the club, she has a meaningful long-term relationship with you. She could also add that, since no one knows about your promise (not even her!) breaking the promise would have no tendency to undermine public trust in promises. Now presumably, rejectability is essentially comparative in the following sense: you always reject something relative to some alternative, which you consider unrejectable (again, relative to the first thing). Principle Contractualists accept this with respect to principles, but then I see no reason not to endorse this assumption with respect to acts.[30] If so, the rejectability of your promise-keeping act would entail the unrejectability of your incompatible promise-breaking act. We can assume, as it seems plausible to do, that your promise-breaking is unrejectable. In particular, we can assume that neither Bo nor the club can reasonably reject that act. The club was never promised anything, and Bo, who was promised, is no longer alive, and is not vulnerable to harm.[31] All in all, it is not implausible to assume that, while the individual rejections of your promise-keeping are reasonable, those of your promise-breaking are not. In other words, it is not unreasonable to assume that Act Contractualism requires you to break your promise.

The point of this story is not to convince you to break your desert island promises, but to illustrate the fact that Act and Principle Contractualism are genuinely distinct theories. Desert Island simply strikes me as the kind of case in which these theories are likely to deliver conflicting verdicts. It illustrates my earlier point that Scanlon's generality requirement applies to the unrejectability of principles only. Take the unusual fact that the promisee in this case is dead. This fact hardly seems relevant to whether anyone could reasonably reject PoF as a public standard of behavior. After all, the overwhelming majority of promisees live to see their promises kept or broken! But of course, in deciding whether Jo's (Bo's) rejection of your promise-keeping (promise-breaking) *act* is reasonable, we had better take account of this crucial feature of the case, however unusual. The fact that the average promisee is alive and well is beside the point. For at issue is the unrejectability of breaking a promise to Bo in particular, not to the

[30] Scanlon, *What We Owe*, p. 195; *The Difficulty of Tolerance*, p. 135. See also n. 9.

[31] A preference-satisfaction theorist can maintain that your promise-breaking posthumously harms Bo, by failing to satisfy his preferences. But this is problematic in more than one way. Apart from depending on a preference-satisfaction theory of well-being, it also seems to conflate Bo and his life.

https://doi.org/10.1017/S095382081100015X Published online by Cambridge University Press

average promisee. Or take the unusual fact that the only beneficiary of the promise (the club) does not know about it. This fact is hardly too particular to be relevant to the unrejectability of your particular action. When it comes to the rejectability of some particular act, we had better examine that act in all its idiosyncratic particularity, including 'the particular aims, preferences, and other characteristics of specific individuals'; we had better not confine ourselves to 'commonly available information'.[32]

Suppose now that you wholeheartedly endorse

Foundational Contractualism

What matters ultimately in action is unrejectability.

What would you do? Remember: you believe that unrejectability is what matters ultimately in action. That is to say, you believe that unrejectability has intrinsic moral significance to how we should act and takes priority over all other values or properties in our action. Presumably, then, you seek to realize unrejectability in your action for its own sake and give it priority in deciding how you act. But in the present case you can do this by, and only by, making things unrejectable. That is why I think you would break your promise to Bo. By breaking that promise, you perform an unrejectable act, thereby realizing unrejectability. By keeping your promise on the other hand you perform a rejectable act, thereby realizing rejectability. As a wholehearted Foundational Contractualist who identifies action's ultimate moral point with *un*rejectability, the choice between breaking and keeping your promise should strike you as a no-brainer.

It may be objected that I am ignoring the possibility that, by keeping your promise, you add at least one unrejectable, namely the unrejectable principle PoF. But if the discussion in the previous section is accurate, you cannot add principles to this world, only their societal acceptance/institution, and you cannot add societal acceptance of principles simply by conforming to them. In no sense does your conformity to the unrejectable PoF add unrejectability to our world.[33] What you do realize by keeping your promise to Bo is not unrejectability but a three-place principle-mediated relation that holds between an act

[32] Cf. Scanlon, *What We Owe*, p. 204.

[33] It is no good replying that, by keeping your promise you indirectly add unrejectable *conformity*. If an act's conformity to unrejectable principles does not make the act unrejectable, however indirectly, it also does not make the act's conformity unrejectable, however indirectly.

https://doi.org/10.1017/S095382081100015X Published online by Cambridge University Press

X, principles P, and a collective act/event E just when

(1) The failure to do X violates P.
(2) Were E to occur, P would become socially accepted public standards of behavior.
(3) E is unrejectable.

To suppose that by realizing this convoluted relation to unrejectability you can somehow indirectly realize unrejectability or make things unrejectable is like supposing that I can somehow realize aesthetic value, albeit only indirectly, by painting in the studio of a great artist.

Now I have taken Desert Island to be a case in which your promise-keeping act is reasonably rejectable. Specifically, I have all but assumed that Jo can reasonably reject your promise-keeping. But why couldn't Bo also reasonably reject your promise-breaking? Why not think that both acts open to you are rejectable? (Or maybe they are both unrejectable.) I have three comments to make in response to this complaint. First, just because someone rejects something does not make it rejectable in the relevant sense; one's rejection must be reasonable. And as I have noted, the reasonableness of rejectability is an essentially comparative affair. In the end, there is no escape from looking at the individual objections to incompatible acts and asking which is stronger. Presumably, when we declare an act rejectable relative to an alternative, we are also declaring the alternative unrejectable. To me it seems that the individual objections to your promise-keeping in Desert Island are stronger than the individual objections to your promise-breaking. So I am inclined to say that your promise-breaking in this case is unrejectable and your promise-keeping is rejectable. Second, notice that it is no good complaining that your promise-breaking in this case seems to you to be wrong. As we have seen, we cannot appeal to such moral intuitions in applying the concept of rejectability.[34] Finally, suppose that my first point is mistaken and both of your incompatible acts are rejectable in this case. Then Act and Principle Contractualism would still generate incompatible verdicts in this case. Since your promise-keeping is rejectable, Act Contractualism still does not imply that you are obligated, or even permitted, to keep your promise. It neither requires nor permits you to keep your promise (but neither does it require or permit you to break it). Similarly if your acts are both unrejectable. And since your promise-keeping is required by an unrejectable principle (PoF), Principle Contractualism still implies that you are permitted and required to keep your promise. (Similarly if your acts are both unrejectable.)

[34] See n. 22.

https://doi.org/10.1017/S095382081100015X Published online by Cambridge University Press

You might now suspect that my story begs the question. In that story, both PoF and your promise-breaking are unrejectable. But does not my assumption that a principle and its violation can both be unrejectable presuppose the falsity of Principle Contractualism? It does not. As we have seen, to say of PoF that it is unrejectable is shorthand for saying that no one could reasonably reject it as a public standard of behavior, i.e. to reject its societal acceptance/institution. By contrast, to say of your promise-breaking that it is unrejectable is shorthand for nothing; it is to say, well, that no one could reasonably reject your promise-breaking (as a particular instance of behavior). If Principle Contractualism were true, your promise-breaking would be wrong on account of violating an unrejectable principle. But that would not make it rejectable, except in the derivative sense that it violates an unrejectable principle. But in *that* sense, describing an unrejectable principle's violation as unrejectable is tautologous. It has no tendency to show that your promise-breaking is rejectable proper, in the primary sense of Foundational Contractualism. In sum, the assumption that breaking your promise to Bo violates the unrejectable PoF is compatible with the assumption that it is unrejectable. And that would be so even if Principle Contractualism were true.

A contractualist who thinks you have an obligation to keep your promise to Bo should bite the bullet – namely accept that Foundational Contractualism implicitly commits you to Act Contractualism – and simply reject Foundational Contractualism in favor of Foundational Principle Contractualism. We have already seen the weakness of this response, however. Foundational Contractualism is not just any old foundational contractualist theory; it is the simplest and most natural such theory, a theory one might find appealing even apart from reflection on theories of obligation. If we were to understand action's ultimate moral point in terms of some attractive value such as goodness, justice, or unrejectability at all, what could we plausibly understand that point to be if not that value itself – goodness, justice, or unrejectability? 'Conformity!' replies the champion of Foundational Principle Contractualism, 'Conformity to principles whose societal acceptance would be good, just, or unrejectable'. But since such conformity can be bad, unjust, or rejectable, its identification with action's ultimate moral point seems unmotivated. Surely the sheer fact that the determinative conformity is a relation to an attractive value is not enough; remember the countercontractualist. But then what is so great about a conformity relation to unrejectability?

The next two sections entertain objections. Section VII examines the possibility that the choice between Act and Principle Contractualism is illusory, with reference to Scanlon's suggestive remark that the choice

https://doi.org/10.1017/S095382081100015X Published online by Cambridge University Press

is misconceived. Section VIII asks whether my argument is guilty of double standard.

VII. ACT CONTRACTUALISM AND COLLAPSE

The case for Act and against Principle Contractualism seems analogous to the case for act and against rule-consequentialism. In each case, one theory of obligation seems open to a worship/conformism objection to which the other seems immune. In each case, the objection seems to stem from the two-stage justificatory structure of the offending theory. But as I have noted, modern contractualists remain unmoved by this analogy. The tacit assumption is that the problem of worship is distinctly consequentialist and does not carry over to the contractualist arena. The distinction between act- and rule-consequentialism has no contractualist analogue. If so, my distinction between Act and Principle Contractualism is a distinction without a difference.

The argument implicitly at work in the contractualist literature is reminiscent of the one-time popular argument that rule-consequentialism collapses into act-consequentialism, only here the argument runs in the other direction. Presumably, the idea is that Act Contractualism is not a genuine alternative to Principle Contractualism because these theories are extensionally equivalent, namely give the same results in all cases.

Scanlon explicitly considers the analogy with rule-consequentialism. Given the familiar controversy between act- and rule-consequentialism, he writes, 'it would be natural to ask why justification of our actions to others should proceed by way of principles at all. Why not consider individual acts instead? Put in this way, the question is misconceived.' The question Scanlon rejects as misconceived is the question I have been pressing on behalf of Act Contractualism. Scanlon thinks the question is misconceived because

> [t]o justify an action to others is to offer reasons supporting it and to claim that they are sufficient to defeat any objections that others may have. To do this, however, is also to defend a principle, namely one claiming that such reasons are sufficient grounds for so acting under the prevailing conditions.[35]

The second sentence threatens tautology; it appears to be saying something like this:

> (1) To offer undefeated reasons for doing some particular act X is also to defend the act-specific principle that there are undefeated reasons for doing X.

This truism says nothing about undefeated reasons for accepting the act-specific principle *as a public standard of behavior* – about

[35] Scanlon, *What We Owe*, p. 197.

https://doi.org/10.1017/S095382081100015X Published online by Cambridge University Press

its unrejectability in the practical sense of Principle Contractualism. Hence it does not show that Act Contractualism collapses into Principle Contractualism.

According to a second interpretation, Scanlon's claim amounts to this:

(2) To offer undefeated reasons for doing some particular act X is also to defend the act-general principle that there are undefeated reasons for doing every act that is G, where G is some non-trivial property of X.

Unlike (1), (2) is not tautologous. A *particularist* can reject (2) by claiming that the reasons that defeat objections to X defeat objections to no other act; the only principle we defend when we offer undefeated reasons for doing X is the trivial one that there are undefeated reasons for doing every act that is identical with X.[36]

Let us assume that (2) is true, and ask whether this proves the choice between Act and Principle Contractualism misconceived. The answer is that it simply fails to engage that choice, for both theories entail (2). If Principle Contractualism is true, whenever you offer undefeated reasons for doing some particular act X you defend the act-general principle that there are undefeated reasons for doing every act the non-performance of which violates unrejectable principles. If Act Contractualism is true, whenever you offer undefeated reasons for doing some particular act X, you defend the act-general principle that there are undefeated reasons for doing every unrejectable act. As we have seen, these are two different theories; their deliverances might diverge (as they seem to do in Desert Island). Hence claim (2) does not show that Act Contractualism collapses into Principle Contractualism.

Scanlon effectively makes claim (2) when he writes that judging some particular act to be wrong is equivalent to judging that 'it is wrong for some reason, or in virtue of some general characteristic'.[37] But again, this does not affect Act Contractualism, which is committed to the notion that when acts are wrong they are wrong for a reason – their rejectability.

It may now be objected that unrejectability is not a basic property or reason; it is rather a 'second-order' property/reason that is grounded in 'first-order' properties/reasons. For example, your promise-breaking act is unrejectable because your failure to pay the club can no longer affect the promisee, your daughter needs the money, the club is not aware of your promise, and so on. Therefore, so goes the objection, while your act is wrong for all the reasons that make it rejectable, it is not really wrong 'in virtue of some general characteristic'. But whether or not

[36] Cf. Jonathan Dancy, *Ethics without Principles* (New York, 2004).
[37] Scanlon, *What We Owe*, pp. 197–8.

https://doi.org/10.1017/S095382081100015X Published online by Cambridge University Press

this objection is available to Scanlon, it seems false. For a second-order property/reason is still a genuine property/reason. Consider goodness or the property of being optimific. Presumably, it too is a second-order property, which is grounded in first-order properties. Presumably, no two acts are non-optimific for exactly the same reasons. But if act-consequentialism is correct, they are all wrong for a single reason – being non-optimific.[38]

The only interpretation of Scanlon's misconception charge that engages the contrast between Act and Principle Contractualism is this:

(3) To offer undefeated reasons for doing some particular act X is also to defend the societal acceptance/institution as public standards of behavior of some principles P that require the doing of X.

This, however, threatens to beg the question. If Act Contractualism is true, one can offer undefeated reasons for doing some particular act X without defending the societal acceptance/institution of any principles as public standards of behavior. In addition, if the discussion so far is basically right, then (3) does not seem very plausible. Recall Desert Island. Given the fact that, compared to not giving the money to the club, not giving it to your daughter is considerably more detrimental, you might have an undefeated reason to break your promise to give the money to the club. But this hardly tends to show that we, as a society, have an undefeated reason to accept some principle that requires promise-breaking in similar situations. In fact, we may well have an undefeated reason to accept PoF, which requires you to do the opposite.

Let me mention one more possible rationale for Scanlon's dismissal of the Act Contractualist challenge. If unrejectable principles were super-simple, they could require us to perform acts that lead to disaster. Scanlon takes pains to exclude this possibility by stressing that unrejectable principles include many qualifications:

We are not morally required to keep a promise no matter what. The clearest part of the principle is this: the fact that keeping a promise would be inconvenient or disadvantageous is not normally a sufficient reason for breaking it, but 'normally' here covers many qualifications.[39]

This is a corrective to the generality requirement. While unrejectable principles focus on the typical case, they also take account of typical deviations from that case. Presumably, PoF does not require one

[38] Compare Parfit's comment that 'Scanlon shows that . . . we have and can usefully appeal to intuitive beliefs about what are reasonable grounds for rejecting moral principles. That is Scanlon's greatest contribution to our moral thinking' (*On What Matters*, vol. 1, p. 370, § 54). I would just add that, if Scanlon is right about this, then we can also usefully appeal to intuitive beliefs about what are reasonable grounds for rejecting acts.

[39] *What We Owe*, p. 199. See also, *Moral Dimensions*, pp. 21–2.

https://doi.org/10.1017/S095382081100015X Published online by Cambridge University Press

to keep one's promise in a typical emergency case. Thus Principle Contractualism is immune to the objection that conformity to unrejectable principles can lead to disaster.[40]

But the Act Contractualist complaint is not that conformity to Principle Contractualism can lead to disaster; it is that it can lead to *rejectability*, disastrous or not. After all, Scanlon's claim that unrejectable principles can have many qualifications is not meant to retract the generality requirement. But you cannot have the positive side of the generality coin (conformity to unrejectable principles typically requires unrejectable acts) without the negative (conformity to unrejectable principles atypically requires rejectable acts). Desert Island illustrates such atypical cases. But Principle Contractualism applies to every action, however atypical.

There is, however, another natural way to take the claim that principles are only 'normally' decisive and have many exceptions.[41] We could take this qualification to suggest that principles are mere 'rules of thumb', useful indicators of acts that are independently right or wrong, indicators that must be set aside when we know all the facts about the (un)rejectability of the evaluated acts themselves. But of course, this interpretation would contradict the notion that the role of principles 'is fundamental; they do not enter merely as devices for the promotion of acts that are right according to some other standard'.[42] It would, in other words, effectively repudiate Principle Contractualism in favor of Act Contractualism.

VIII. ACT CONTRACTUALISM AND ACT WORSHIP

The final objections I consider accuse me of double standard.[43] According to the first, Act and Principle Contractualism have the same two-stage asymmetric justificatory structure: each gives justificatory priority to something. While Principle Contractualism gives justificatory priority to principles, justifying principles directly by their own unrejectability and acts only indirectly by their conformity to unrejectable principles, Act Contractualism gives justificatory priority

[40] For a similar defense of rule-consequentialism, see Brad Hooker, *Ideal Code, Real World* (New York, 2000).

[41] Scanlon, *What We Owe*, p. 199 and *Moral Dimensions*, pp. 21–4.

[42] Scanlon, *The Difficulty of Tolerance*, p. 142.

[43] Compare the discussion of consequentialism in Shelly Kagan, 'Evaluative Focal Points', and Phillip Pettit and Michael Smith, 'Global Consequentialism', both in *Morality, Rules and Consequences*, ed. B. Hooker, E. Mason, and D. Miller (Edinburgh, 2000), pp. 121–33 and 134–55 respectively. The objection in the text is not attributable to these authors, however.

https://doi.org/10.1017/S095382081100015X Published online by Cambridge University Press

to acts, justifying acts directly by their own unrejectability and principles only indirectly by their *proscription* of rejectable acts. While Act Contractualism is a direct contractualist theory of acts, it is an indirect contractualist theory of principles. But if there is something wrong with justifying acts indirectly by their relation to principles, there must also be something wrong with justifying principles indirectly by their relation to acts. Therefore, if Principle Contractualism is guilty of principle worship, Act Contractualism is guilty of act worship.

The objection mistakenly assumes that Act Contractualism justifies principles indirectly in terms of acts. Act Contractualism is a moral theory of action. It answers the question:

When, and in virtue of what, does agent A have an obligation to do X?

As a moral theory of action, Act Contractualism does not purport to justify anything other than acts.[44] To be sure, Act Contractualism can have moral implications for principles, at least when extended to 'collective acts'. So extended, Act Contractualism enjoins us, as a society, collectively to accept – promulgate, teach, enforce, or otherwise institute – principles whose institution no one could reasonably reject. But then the justifiability of our collective principle acceptance/institution consists of its own non-derivative unrejectability; it matters not whether the acts proscribed by the principles are unrejectable.

The crucial asymmetry between Act and Principle Contractualism comes to this. While both are moral theories of acts, only Act Contractualism justifies acts non-derivatively, by their own unrejectability; Principle Contractualism justifies acts only derivatively, by their conformity to principles whose societal acceptance/institution is unrejectable. Principle Contractualism is vulnerable to the worship charge because it justifies acts derivatively, regardless of whether they are rejectable. Nothing of this sort can be said of Act Contractualism. Whatever Act Contractualism justifies it justifies by its own unrejectability. And Act Contractualism justifies acts only. In particular, it does not justify principles. Now again, it may or may not make sense to extend Act Contractualism to collective acts. If it does, then Act Contractualism can be said to justify the societal acceptance/institution of principles through their promulgation, teaching, and enforcement. But then such acceptance is just another 'act', to be justified non-derivatively by its own unrejectability. Justification under Act Contractualism

[44] It is not an instance of what Kagan would call 'indirect contractualism' or Pettit and Smith would call 'local contractualism'.

https://doi.org/10.1017/S095382081100015X Published online by Cambridge University Press

proceeds in exactly one stage; the only justification is direct or non-derivative.

We can imagine a rejoinder. As I define Act Contractualism, goes the rejoinder, it does not really make a distinctive claim about the justification of acts; it is really just a special case of

Direct Contractualism

Everything is justified by its own unrejectability.[45]

This entails that acts are justified by their unrejectability, but also that principles are justified by *their* unrejectability. What is more, it entails that chairs, stones, and atoms are justified by their unrejectability. Act Contractualism, so goes the thought, is simply one consequence of Direct Contractualism among infinitely many. If so, Act Contractualism does not tell us anything special about acts.

I have no qualm with Direct Contractualism. I tend to agree that, to the extent that it makes sense to talk about the contractualist justification of some object, that object can only be justified non-derivatively, by its own unrejectability. But I am wary of the implication that we justify principles – or chairs or stones or atoms – in the same sense in which we justify acts. The concept of justification at work in moral theories of action is essentially practical. And the only thing that seems to me to admit of practical justification is acts (individual and perhaps also collective). Indeed Act Contractualism does not simply justify acts; it justifies acts in the practical sense of showing them to be obligatory or permissible – or more generally in the sense of being supported by practical reasons (reasons for action). By contrast, stones, chairs, and atoms are neither obligatory nor permissible nor supported by practical reasons. In that respect, principles are like stones rather than like acts. Acts are special not because other things are justified derivatively by their relation to justified acts.[46] Acts are special because only they can be justified in an essentially practical way. Your action can be justified in the sense that you have a reason for that action; a chair or principle cannot be justified in the sense that you have a reason for that chair or principle. The insistence that acts are to be justified by their own unrejectability therefore means that acts play

[45] Cf. the discussion of 'direct' and 'global' consequentialism in Kagan, 'Evaluative Focal Points', pp. 149–151 and Pettit and Smith, 'Global Consequentialism'.

[46] Cf. Kagan's remark that 'the distinctive claim of the act consequentialist [is this:] it is only acts that are to be evaluated directly in terms of the good, and that rules are to be evaluated only indirectly, in terms of the best acts' ('Evaluative Focal Points', p. 149).

https://doi.org/10.1017/S095382081100015X Published online by Cambridge University Press

a unique role in the only practical justification there is, their own justification.[47]

The final objection I consider is another double standard charge. I have claimed that what matters ultimately in action under Principle Contractualism is not unrejectability but some *relation to* unrejectability. It may now be objected that the same is true of Act Contractualism. What matters ultimately in action to the Act Contractualist is not unrejectability, but the relation that holds just when acts realize unrejectability. Call this the *realization* relation. If Principle Contractualism fails to reflect Foundational Contractualism because it prizes some conformity relation, Act Contractualism fails to reflect Foundational Contractualism because it prizes the realization relation. Or so goes the objection.

The truth in the objection is that every contractualist moral theory of action makes *some* relation to unrejectability morally determinative. But all relations are not created equal, and the objection ignores the uniqueness of the realization relation in the present context, the context of moral theories of action. Of all the different relations to some putative value V – goodness, justice, unrejectability – only the realization relation is such that its realization in action realizes V itself. Consequently, only a theory that requires the realization of the realization relation to V treats V as what matters ultimately in action. There is simply no difference between seeking and prioritizing goodness, justice, or unrejectability in one's action and seeking and prioritizing its *realization* in one's action. To realize a value in one's action is to realize its realization.

Not so with non-realization relations to V. Seeking and prioritizing V in one's action – however indirectly – is one thing; seeking and prioritizing some non-realization relation to V in one's action is another thing altogether. Consider three agents, corresponding to three non-realization relations to goodness. The counterconsequentialist, angel consequentialist, and angel rule-consequentialist seeks to realize badness, goodness in a society of angels, and conformity to rules that would realize goodness in a society of angels, respectively. Each of these characters prizes some relation to goodness, but none can be said to seek, prioritize, or realize goodness itself – however indirectly. The same point applies to the simple rule-consequentialist. And an exactly analogous point applies to Principle Contractualism.

The realization relation, then, has unique practical significance. The good, just, or unrejectable agent is a goodness, justice, or unrejectability

[47] What if we redefine Direct Contractualism in terms of the broader concept of evaluation? We can certainly evaluate stones, even if we cannot justify them. But then Act Contractualism is not a consequence of Direct Contractualism.

https://doi.org/10.1017/S095382081100015X Published online by Cambridge University Press

realizer. As the only contractualist theory of obligation that requires us to realize unrejectability, Act Contractualism is the only contractualist theory that reflects Foundational Contractualism.

IX. CONCLUSION

Scanlon writes that, while modern contractualism shares with rule-consequentialism the important feature that the defense of individual actions must proceed via a defense of principles, the

> role of principles in contractualism is fundamental; they do not enter merely as devices for the promotion of acts that are right according to some other standard. Since it does not establish two potentially conflicting forms of moral reasoning, contractualism avoids the instability which often plagues rule [consequentialism].[48]

But this only proves the troubling analogy. For the role of rules in rule-consequentialism *is* in fact fundamental – as fundamental as is the role of principles in Principle Contractualism. And *that* is precisely the source of its instability: its criterion of right (conformity to good principles) is incompatible with the independently plausible consequentialist theory of action's ultimate moral point from which it borrows its initial plausibility (foundational consequentialism). And that is just as true of Principle Contractualism. In both cases, there is a better alternative, namely a theory of obligation that justifies acts without reference to principles, and proceeds in a single justificatory stage. That theory determines the morality of acts by their own primary, consequentialist, or contractualist justifiability. In the contractualist case, that more fitting theory of obligation takes the form of Act Contractualism.[49]

sheinman@rice.edu

[48] *The Difficulty of Tolerance*, pp. 241–2.

[49] I am greatly indebted to Hagit Benbaji, Stephen Darwall, Brad Hooker, Shelly Kagan, Greg Klass, Matt Kramer, Alastair Norcross, Philip Pettit, Henry Richardson, Tim Scanlon, Michael Smith, and an anonymous reader for *Utilitas* for their useful conversations or comments. Earlier drafts were presented at the Rice University Humanities Research Center, Cambridge Forum for Legal and Political Philosophy, Bar-Ilan University Philosophy Colloquium, Ben Gurion University Philosophy Colloquium, Georgetown Contract and Promise Workshop, and Georgetown Law and Philosophy Workshop. I would like to thank the audiences in these forums for the valuable discussions.

https://doi.org/10.1017/S095382081100015X Published online by Cambridge University Press