

DE GRUYTER

REASON, JUSTIFICATION, AND CONTRACTUALISM

THEMES FROM SCANLON

Edited by Markus Stepanians and Michael Frauchiger

LAUENER LIBRARY OF
ANALYTICAL PHILOSOPHY

EBSCO Publishing : eBook Collection (EBSCOhost) - printed on

2/12/2023 10:26 AM via

AN: 1333763 ; Markus Stepanians, Michael Frauchiger.; Reason,
Justification, and Contractualism : Themes From Scanlon

Accessions335141

Markus Stepanians and Michael Frauchiger (Eds.)
Reason, Justification, and Contractualism

Lauener Library of Analytical Philosophy



Edited by
Wilhelm K. Essler, Dagfinn Føllesdal,
and Michael Frauchiger

Volume 7

Reason, Justification, and Contractualism



Themes from Scanlon

Edited by
Markus Stepanians and Michael Frauchiger

DE GRUYTER

ISBN 978-3-11-073843-8

e-ISBN (PDF) 978-3-11-073367-9

e-ISBN (EPUB) 978-3-11-073375-4

ISSN 2198-2155

Library of Congress Control Number: 2021939131

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie;
detailed bibliographic data are available on the Internet at <http://dnb.dnb.de>.

© 2021 Walter de Gruyter GmbH, Berlin/Boston

Printing and binding: CPI books GmbH, Leck

www.degruyter.com

Brief Book Overview

This book collects major original essays developed from lectures given at the award of the Lauener Prize 2016 to T. M. Scanlon for his Outstanding Oeuvre in Analytical Philosophy. In 'Contractualism and Justification', Scanlon identifies some difficulties in his theory and explores possible ways to deal with them. In 'Improving Scanlon's Contractualism', D. Parfit recommends revisions and extensions of Scanlon's theory, while R. Forst suggests in 'Justification Fundamentalism' that Scanlon may want to replace reason with justification as his foundational concept. T. Nagel raises fundamental questions concerning 'Moral Reality and Moral Progress', and S. Mantel offers in 'On How to Explain Rational Motivation' a critical discussion of Scanlon's cognitivist theory of motivation. Z. Stemplowska does the same for Scanlon's conception of responsibility in 'Substantive Responsibility and the Causal Thesis', and S. Olsaretti suggests in 'Equality of Opportunity and Justified Inequalities' an alternative to Scanlon's arguments against economic inequalities. All contributors receive extensive replies by Scanlon. For anyone interested in T. M. Scanlon's seminal work in moral and political philosophy, the present volume is utterly indispensable.

<https://doi.org/10.1515/9783110733679-001>

Acknowledgments

We would like to express our special thanks to Tim Scanlon both for having been very actively involved in the Lauener Symposium held in Bern in 2016 on themes from him and for his generous and productive way of collaborating with us on the preparation of this volume. And, beyond that, we owe gratitude to him for having made it possible for Derek Parfit's animating symposium paper—presumably Parfit's last paper—to be available for the present volume. What is more, we would like to express our deep appreciation and thanks to Dagfinn Føllesdal, the president of the Lauener Foundation, for his tireless, spirited involvement in the attempt to take the occasion of the symposium to bring together again three of the most influential contemporary moral philosophers: Scanlon, Nagel, and Parfit. Derek Parfit's unexpected death, only four months after his participation in the Symposium alongside Scanlon and Thomas Nagel, increases the significance of the Lauener event in honor of Scanlon, since it made it possible for these three interlocutors and truly outstanding philosophers to spend one more time together, enjoying their conversations, especially at lunch and dinner, on a wide range of subjects.

Moreover we would like to sincerely thank all other contributors to this volume for the commitment they have shown in thoroughly reworking their symposium papers for publication in this collection.

Many thanks also go to our fellow members of the Lauener Foundation Council (Dagfinn Føllesdal, Alex Burri, Claus Beisbart, and Dieter Jordi) for their friendly encouragement as well as to Wilhelm K. Essler and Dagfinn Føllesdal qua coeditors of the book series “Lauener Library of Analytical Philosophy”, edited on behalf of the “Lauener-Stiftung”. Beyond that we would like to thank Christoph Schirmer and Mara Weber at De Gruyter as well as Matthias Wand for their much-appreciated cooperation in this book project.

Acknowledgment notes which are specific to certain individual contributions can be found at the beginning of the respective articles.

Table of Contents

Brief Book Overview — V

Acknowledgments — VII

Markus Stepanians

Instead of an Introduction: Scanlon's Project — 1

Rainer Forst

Deontological Communitarianism. Laudation for Thomas M. Scanlon — 11

T. M. Scanlon

Contractualism and Justification — 17

Rainer Forst

Justification Fundamentalism: A Discourse-Theoretical Interpretation of Scanlon's Contractualism — 45

Susanne Mantel

On How to Explain Rational Motivation: Where Internalism and Externalism Meet — 59

Thomas Nagel

Moral Reality and Moral Progress — 83

Serena Olsaretti

Equality of Opportunity and Justified Inequalities: How the Family Can Be on Equality's Side — 91

Derek Parfit

Improving Scanlon's Contractualism — 109

Zofia Stemplowska

Substantive Responsibility and the Causal Thesis — 119

T. M. Scanlon

Responses to Forst, Mantel, Nagel, Olsaretti, Parfit, and Stemplowska — 131

X — Table of Contents

Susanne Mantel

Words of Thanks — 155

Contributors — 157

Index — 159

Markus Stepanians

Instead of an Introduction: Scanlon's Project

In search of a non-utilitarian theory of morality

Scanlon first conceived of contractualism as a general theory of morality in the spring of 1979 while on leave from Princeton at All Souls College, Oxford. The main challenge, as he saw it at the time, was to find a viable alternative to the then still dominant utilitarian moralities in their various forms. Scanlon shared the view of many moral theorists that, despite more than 100 years of serious attention, refinement and discussion, utilitarianism had failed to provide a plausible reconstruction of our deepest moral commitments. Scanlon considered act utilitarianism to be “wildly at variance” with some of our most cherished moral beliefs,¹ and he regarded the retreat to a form of rule utilitarianism as “an unstable compromise” (Scanlon 1982, p. 102).

Discontent with utilitarianism (or more broadly: consequentialism) was not new, of course. It had already led Rawls in the 1950s in search of a non-utilitarian foundation for a theory of justice. Rawls found it in “the concept of fairness which relates to right dealing between persons who are cooperating with or competing against one another, as when one speaks of fair games, fair competition, and fair bargains” (Rawls 1958, p. 178). Rawls claimed that it is this idea of fairness “which utilitarianism, in its classical form, is unable to account, but which is expressed, even if misleadingly, by the idea of the social contract” (Rawls 1958, p. 164). Famously, Rawls went on to develop this ‘contractualist’ notion of fairness as “right dealing between persons” into a comprehensive theory published in 1971 as *A Theory of Justice*.

Scanlon agreed with Rawls's view about the untenability of utilitarianism, shared Rawls's diagnosis of utilitarianism's inherent lack of fairness and joined him in crediting the concept of fairness to “the social contract tradition going back to Rousseau” (Scanlon 1998, p. 5).² However, Scanlon felt that, by limiting

1 Scanlon illustrates this point in *What We Owe to Each Other* with his Transmitter Room example: Scanlon 1998, p. 235.

2 A second source of influence on Scanlon's thinking at the time was Thomas Nagel. Here is Nagel's 'contractualist' critique of utilitarianism: “[Utilitarian justifications] are really justifications to the world at large which the victim, as a reasonable man, would be expected to appreciate. However, there seems to me something wrong with this view, for it ignores the possibility

himself to the development of a non-utilitarian theory of justice, Rawls had undersold the explanatory potential of the contractualist key notion of fairness: “Contractualism has been proposed as the alternative to utilitarianism before, notably by John Rawls in *A Theory of Justice*. Despite the wide discussion which this book has received, however, I think that the appeal of contractualism as a foundational view has been underrated” (Scanlon 1982, p. 103). Scanlon was convinced that the contractualist idea of fairness, reformulated by him as fair-mindedness in the sense of “a shared willingness to modify our private demands in order to find a basis of justification that others also have reason to accept” (Scanlon 1998, p. 5) provided not only an auspicious starting-point for a theory of justice but for morality in general.³ Moreover, and perhaps as a result of Rawls’s narrow focus on justice, *A Theory of Justice* did not shake utilitarianism’s status as the dominant theory of morality.⁴ Scanlon felt that it was still the view towards which most philosophers found “themselves pressed when they try to give a theoretical account of their moral beliefs” (Scanlon 1982, p. 102). Given the untapped explanatory potential of contractualism for general morality and Scanlon’s (widely shared) belief that utilitarian reconstructions of morality tended to be “wildly at variance” with some of our most cherished moral commitments, the need to extend Rawls’s contractualist framework beyond justice appeared obvious.

that to treat someone else horribly puts you in a special relation to him, which may have to be defended in terms of other features of your relation to him. The suggestion needs much more development; but it may help us to understand how there may be requirements which are absolute in the sense that there can be no justification for violating them. If the justification for what one did to another person had to be such that it could be offered to him specifically, rather than just to the world at large, that would be a significant source of restraint” (Nagel 1972, p. 137).

3 However, by the time Scanlon wrote *What We Owe to Each Other* he had concluded that morality does not constitute a unified normative domain and that contractualism accounts only for interpersonal duties – hence the title of the book. For reasons of simplicity I will take ‘morality’ and ‘moral’ to refer only to this interpersonal realm in what follows.

4 It also did not help that Rawls’s decidedly political conception of justice as “the first virtue of social institutions” (*A Theory of Justice*, p. 3) is far narrower than the classical notion of justice as the primary virtue of persons in its various (commutative, distributive, rectifying ...) varieties. See Aristotle’s *Nicomachean Ethics*, book five.

The attractions of utilitarianism – three foundational questions

But what is the source of utilitarianism's persistent appeal despite its inability to account for many of our most firmly held moral beliefs? Whence utilitarianism's enduring attractiveness? According to Scanlon, the "wide appeal of utilitarianism is due ... to philosophical considerations of a more or less sophisticated kind which pull us in a quite different direction than our first-order moral beliefs" (Scanlon 1982, p. 102). These philosophical considerations of a more or less sophisticated kind are the utilitarian's persuasive answers to certain foundational challenges any general theory of morality has to confront. For Scanlon, it is primarily the soundness of utilitarianism's metaphysical foundations that accounts for the widespread feeling that, despite its counterintuitive moral implications downstream, utilitarianism must be the right moral theory after all. This diagnosis sets the bar for any serious non-utilitarian rival: "What a successful alternative to utilitarianism must do, first and foremost, is to sap this source of strength by providing a clear account of the foundations of nonutilitarian moral reasoning" (Scanlon 1982, p. 102).

What are the most important foundational challenges a viable moral theory has to meet? The first question concerns the subject matter of morality. What are moral judgements about and how is their truth or falsity determined? Empirical judgements are made true by facts about sensual objects, mathematical judgements by facts about numbers and geometrical figures, psychological judgements by facts about objects of consciousness. But what are the facts of morality and their objects? The utilitarian's answer to the question of subject matter is that moral judgements are about individual well-being, and that it is the impact of moral judgements on well-being that determines their truth-value. This seems plausible. That well-being plays a pivotal role in morality appears evident and gives the utilitarian axiology a very commonsensical ring. According to Scanlon, it is primarily the persuasiveness of the focus on well-being that explains the persistent appeal of utilitarianism.

Any non-skeptical answer to the question of subject matter and truth-conditions immediately gives rise to a second, epistemic question: How are the objects of morality given to us? And how do we go about deciding the truth or falsity of moral judgements? The utilitarian's answer to this second, epistemic question is as commonsensical as his answer to the first. Since moral judgements are about individual well-being, the epistemic problem concerns our access to facts about well-being, and there seems to be no mystery about how we achieve this. Finding out about how someone is doing is a bread-and-butter routine of our daily lives,

and we are pretty good at it. Admittedly, it has proven quite difficult to translate the intuitive straightforwardness of this daily practice into a theory for the general and systematic assessment of well-being. But these seemingly technical problems have done little to shake the confidence of utilitarianism's adherents in the general feasibility of such a theory.

Third and finally, there is a foundational question regarding the action-guiding significance of moral judgements. In his early 'Contractualism and Utilitarianism' Scanlon introduces the issue by observing that sometimes concepts developed for one subject matter have surprising applications in others (Scanlon 1982, p. 104). A notoriously puzzling example of such a conceptual cross-over is the effectiveness of mathematics in the natural sciences.⁵ Perhaps less perplexing, but still in need of explanation is the effectiveness of moral judgements for determining the will to act: "In the case of morality the main connection is, or is generally supposed to be, with the will. Given any candidate for the role of subject matter of morality we must explain why anyone should care about it" (Scanlon 1982, p. 104).

It is worth noting, however, that for Scanlon the question of action-guiding significance is neither about convincing the morally tone-deaf nor about the psychological preconditions of practical efficacy. Rather, it is about the moral force and authority of moral judgements that strikes most people as inescapable and is felt even by those who resist it. In order to account for this action-guiding significance we have to look beyond the mental acts at their propositional contents and the reasons we grasp and come to understand in them. A viable moral theory has to make the force of these reasons intelligible: "A satisfactory moral philosophy [...] must make it understandable why moral reasons are ones that people can take seriously, and why they strike those who are moved by them as reasons of a special kind of stringency and inescapability" (Scanlon 1982, p. 104f.).

How does the utilitarian account for the action guiding significance of moral judgements? Given his response to the first and second foundational question, the utilitarian's answer is obvious enough. If judgements of moral right and wrong are about well-being, the normative force of moral judgements is easily explained by their tendency to affect well-being. Acting on our moral judgements can and often does make people better or worse off. It is therefore the inherent potential to affect people's well-being that explains the action-guiding signifi-

⁵ A classical reflection on this phenomenon for mathematics is Wigner 1960. The mathematical case is particularly puzzling because it appears to straddle the gap between the abstract and the concrete.

cance of moral judgements. As Scanlon notes, this obvious and commonsensical solution of the problem is another strength of utilitarianism: "It seems evident to people that there is such a thing as individuals' being made better or worse off. Such facts have an obvious motivational force; it is quite understandable that people should be moved by them in much the way that they are supposed to be moved by moral considerations" (Scanlon 1982, p. 104).

Scanlon's mature reformulations of the three foundational questions

How does the contractualist answer the three foundationalist questions? Scanlon states his mature view in *What We Owe to Each Other* from 1998. The book opens with a reaffirmation of the importance of the question of subject matter, epistemic access, and action-guiding significance for any viable theory of morality.⁶ Scanlon now reformulates the problem of subject matter explicitly as concerning the ontology of moral truth. It is about the "ontology of morals – for example, about the metaphysical status of moral facts" (Scanlon 1998, p. 2). The epistemic question is recast in *What We Owe to Each Other* as the problem of adequately characterizing the distinctive method of moral reasoning for moral conclusions: "... we arrive at the judgment that a certain kind of action would be wrong simply by thinking about the question in the right way, sometimes through a process of careful assessment that it is natural to call a kind of reasoning. But what kind of reasoning is it?" (Scanlon 1998, p. 1). Finally, the question of the action-guiding significance of moral judgements reappears in *What We Owe to Each Other* as the problem of accounting for their striking reason-giving force: "... the fact that a certain action would be morally wrong seems to provide a powerful reason not to do it, one that is, at least normally, decisive against any competing considerations. But it is not clear what this reason is. Why should we give considerations of right and wrong, whatever they are, this kind of priority over our other concerns and over other values?" (Scanlon 1998, p. 1).

However, Scanlon suggests in *What We Owe to Each Other* a methodological reversal of priorities relative to the approach taken in 'Contractualism and Utilitarianism.' In his early sketch of contractualism from 1982, Scanlon ranked the

⁶ In Scanlon's 'Contractualism and Justification' in the present volume, the foundational questions receive yet another restatement. In this essay, the issues concerning subject matter and epistemic decision procedure are compressed into "the question of content", and the problem of action-guiding significance reappears as "the question of acceptance."

metaphysical question of subject matter as “the first philosophical question about ... morality” (Scanlon 1982, p. 103). In *What We Owe to Each Other* the importance of the problem of subject matter is demoted and played down: “... this metaphysical question is not, for me at least, the primary issue” (Scanlon 1998, p. 2). Moreover, Scanlon now thinks that to tackle the problem of the ontology of moral truth first is to put the cart before the horse. Doing so fails to appreciate that moral judgements are not meant to describe anything. Their point is practical and normative, not theoretical and descriptive. First and foremost, moral judgements are answers to the question: ‘What is the right thing (for me) to do?’ We answer it by engaging in a peculiar type of deliberation towards a conclusion with action-guiding significance. Given the thoroughly practical point of moral judgements, the correct methodological strategy is to reverse the original ranking of the three foundational questions. We should start our inquiry with an explanation of the action-guiding significance of moral judgements, followed by an account of the kind of reasoning that results in moral conclusions: “Thus, of the three questions about right and wrong with which I began—the questions of subject matter, method of reasoning, and reason-giving force—it is the second and especially the third which I take to be of primary concern. Accordingly, I take the reason-giving force of judgments of right and wrong as the starting point of my inquiry” (Scanlon 1998, p. 3). Furthermore, Scanlon now argues that satisfactory explanations of the second and the third question would give us, in effect, all we need to know about the first. All questions concerning the ontology of moral truth are already answered, if only implicitly, in the contractualist’s response to the second and third question: “If we could characterize the method of reasoning through which we arrive at judgments of right and wrong, and could explain why there is good reason to give judgments arrived at in this way the kind of importance that moral judgments are normally thought to have, then we would, I believe, have given a sufficient answer to the question of the subject matter of right and wrong as well. No interesting question would remain about the ontology of morals—for example, about the metaphysical status of moral facts” (Scanlon 1998, p. 2).

The contractualist’s answers to the three foundational questions.

Scanlon has argued that any successful alternative to utilitarianism must, first and foremost, sap utilitarianism’s source of strength by offering a clear account of the foundations of non-utilitarian moral reasoning. For the contractualist to

make good on his claim to provide such an alternative, his answers to the three foundational questions have to be as compelling as those of his utilitarian rival; and in order to successfully displace utilitarianism as the dominant moral theory, they'd better be better.

So how does the contractualist answer the methodologically primary question of the action-guiding significance? He acknowledges, as he must, the great significance of individual well-being. The contractualist denies, however, that the beneficial consequences of our actions are the only things that matter. Scanlon argues that, more often than not, there is another highly relevant element in play. The ultimate source of this additional element is the inherent capability of human beings to grasp and assess reasons for wanting certain opportunities and for objecting against certain treatments. The fact that all of us have and act upon reasons makes the question 'Why – for what reason(s) – did you (not) do this?' (whether we address it to ourselves or to others) always pertinent, proper and even inescapable. For the contractualist, fair-mindedness begins with the recognition of this inherent reason-responsiveness. It calls for the cultivation of a steady and constant willingness to justify one's actions on grounds that those who share our inherent reason-responsiveness can accept.

Scanlon illustrates the force of this additional constraint besides considerations of well-being with a vivid description of his own emotional reaction to Peter Singer's ethical reflections on the Bangladesh famine of 1974: "Certainly sympathy of the usual kind is one of the many motives that can sometimes impel one to do the right thing. It may be the dominant motive, for example, when I run to the aid of a suffering child. But when I feel convinced by Peter Singer's article on famine [Singer 1972], and find myself crushed by the recognition of what seems a clear moral requirement, there is something else at work. In addition to the thought of how much good I could do for people in drought-stricken lands, I am overwhelmed by the further, seemingly distinct thought that it would be wrong for me to fail to aid them when I could do so at so little cost to myself" (Scanlon 1982, p. 115). For the contractualist, one's indifference to the plight of others does not only show a lack of benevolence (as manifest in the failure to exert one's causal powers to increase well-being). What makes the failure to respond to urgent needs morally wrong in the first place is that there is nothing one can say for oneself in response to the pertinent question 'Why – for what reason(s) – did you (not) do this?' Scanlon's generalization and elaboration of this demand for justification leads him to answer the question of action-guiding significance as follows: "When I ask myself what reason the fact that an action would be wrong provides me with not to do it, my answer is that such an action would be one that I could not justify to others on grounds I could expect them to accept" (Scanlon 1998, p. 4).

What is the contractualist's response to the epistemic question concerning the distinctive method of reasoning we apply in practical deliberations for moral conclusions? The contractualist criterion of wrongness sets the terms for a reason-based justificatory procedure. Given a proposed action, we start by formulating a general principle that would allow it. We then inquire what possible objections (based exclusively on how its general acceptance would impact the objector's life and prospects) could be raised by any fair-minded person against that principle and in favor of its negation. Finally, we ask ourselves whether the objections are reasonable and strong enough to override the reasonable grounds other reasonable persons might have for adopting it. If the objections against the principle were to trump those in favor of it, all actions covered by this general rule are thereby shown to be wrong. If the principle under discussion survives reasonable criticism and remains undefeated, all actions sanctioned by it are acceptable and therefore morally permissible (cp. Scanlon 2014, p. 96).

Where do the contractualist's answers to the questions of action-guiding significance and epistemically right reasoning leave us with the problem of the ontology of moral truth or subject matter? We have seen that, for Scanlon, the importance of the metaphysical question of subject matter is derivative and secondary in relation the two other questions about the epistemology and action-guiding significance. Scanlon argues that, by answering the latter questions, we "have given a sufficient answer to the question of the subject matter of right and wrong as well. No interesting question ... remain[s] about the ontology of morals – for example, about the metaphysical status of moral facts" (Scanlon 1998, p. 2). We must therefore ask: What can we learn from the contractualist's answers to the first two questions about the third question concerning the ontology of moral truth? We have seen that moral questions are decided by a reason-based decision procedure concerning the rejectability of principles; and that the action-guiding significance and authority of the decisions thus reached is to be explained in terms of reasons derived from principles compatible with the ideal of universal agreement about the principled regulation of behavior. The contractualist's answer to the question of subject matter is therefore that the objects of moral judgements are ultimately facts about reasons not to be treated in certain ways.

In *What We Owe to Each Other* Scanlon anticipates the concern that the concept of a reason is too obscure to be basic: "It might be objected that this is to explain right and wrong in terms of something else—the idea of a reason—that is equally in need of philosophical explanation" (Scanlon 1998, p. 1). Scanlon admits that the concept of a reason needs explanation; but he strenuously denies that this explanation must take the form of a conceptual analysis or breakdown into simpler notions. He argues that it cannot take this form since the concept of

reason is primitive and fundamental. This is where we reach metaphysical bed-rock, and the spade is turned. According to Scanlon, the concept of a reason neither allows nor needs further conceptual reduction; but it does allow and needs philosophical elucidation: “I do not believe that we should regard the idea of a reason as mysterious, or as one that needs, or can be given, a philosophical explanation in terms of some other, more basic notion. In particular, the idea of a reason should not be thought to present metaphysical or epistemological difficulties that render it suspect” (Scanlon 1998, p. 3). It took Scanlon another 16 years to undergird and substantiate his belief in the non-reducibility of the concept of a reason by developing a highly sophisticated theory of practical reasons in *Being Realistic About Reasons* from 2014. The book elucidates the concept of a practical reason by describing its role, function and place in our normative conceptual scheme. This account of the key concept of a reason and the book's elaborate defense of the doctrine of ‘Reasons Fundamentalism’ puts the capstone on Scanlon's impressive contractualist construction.

Where does all this leave us with the problem of finding a metaphysically sound alternative to utilitarian (or consequentialist) moral theory? Scanlon's contractualism proposes that we replace the utilitarian account of action-guiding significance in terms of well-being with the ideal of fair-minded justifiability among inherently reason-responsive actors. As an answer to the epistemic question, Scanlon suggests that we supplant the utilitarian assessment of rightness and wrongness in terms of impact on well-being with a decision procedure based on reasonable justifiability. Finally, the contractualist asks us to move away from the utilitarian emphasis on well-being as the sole value towards a more general concern with practical reasons and the principled objections they support.

Has Scanlon succeeded in his original goal from 1979 to unleash the full explanatory potential of contractualism and of overthrowing utilitarianism as the dominant moral theory? In view of the glacial speed of progress in philosophy I'm inclined to quote Chinese premier Zhou Enlai's answer from 1972 in response to a journalist's question whether the French Revolution was successful: “It is too early to tell.”

References

- Nagel, Thomas 1972, “War and Massacre”, *Philosophy & Public Affairs*, Vol. 1, No. 2, pp. 123–144.
- Rawls, John 1958, “Justice as Fairness”, *The Philosophical Review*, Vol. 67, No. 2, pp. 164–194.

- Rawls, John 1971, *A Theory of Justice*. Cambridge, MA: Harvard University Press.
- Scanlon, Thomas M. 1982, "Contractualism and Utilitarianism" in Amartya Sen and Bernard Williams, eds., *Utilitarianism and Beyond*, Cambridge University Press, 1982, pp. 103–128.
- Scanlon, Thomas M. 1998, *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.
- Scanlon, Thomas M. 2014, *Being Realistic About Reasons*. Oxford: Oxford University Press.
- Singer, Peter 1972, 'Famine, Affluence and Morality,' *Philosophy and Public Affairs* 1 (1972), pp. 229–43.
- Wigner, Eugene, "The Unreasonable Effectiveness of Mathematics in the Natural Sciences", *Communications in Pure and Applied Mathematics*, Vol. 13, No. 1, 1960.

Rainer Forst

Deontological Communitarianism. Laudation for Thomas M. Scanlon

Dear President Professor Føllesdal,
Dear Dr. Frauchiger,
Dear Representatives of the Lauener Foundation,
and, especially,
Dear Professor Scanlon and Dear Lucy Scanlon,
Dear Dr. Mantel,
Dear Ladies and Gentlemen,

It is a great honor and pleasure to address you this afternoon in the form of a laudation for the winner of the Lauener Prize 2016 for an Outstanding Oeuvre in Analytic Philosophy, Professor Thomas Michael Scanlon from Harvard University. The award recognizes the achievements of one of the truly great moral and political philosophers of our time whose work has had and will continue to have a lasting positive impact on our profession. Scanlon's contractualist approach to explaining the core of morality is a powerful theory that achieves the optimum that one can hope for in moral philosophy – to take us decisive steps beyond the debate between consequentialist and Kantian theories in its traditional form.

Before going on to explain in a bit more detail how Scanlon achieves this, I would like to add a more personal observation as to why this opportunity to speak to you today gives me such great pleasure. It was exactly 25 years ago that I went to Harvard as a visiting doctoral fellow to study with John Rawls and Tim Scanlon. Both of them took excellent care of me and devoted much of their precious time to discussing my at the time pretty vague ideas about how to steer a course in the debate between liberal and communitarian theories of justice, a generosity for which I am eternally grateful. Whenever I returned to Harvard ever since then, we would meet and talk, and as everyone who knows Tim Scanlon will attest, a conversation with him is always a truly enlightening experience. In dialogue as well as in his writings he is a model not only of conceptual clarity and argumentative rigor but, more than that, he is also a truly Socratic thinker. His thinking unfolds in a constant debate with alternative ways of seeing things and has the great virtue of hermeneutic philosophy – if I may speak in such terms in the context of an award in analytic philosophy – namely, to make the opponent's point as strong as possible. Scanlon never takes the easy way out in a debate, and he is also an excellent critical reader of his own work, creating a systematic philosophy by way of elaborating and oc-

<https://doi.org/10.1515/9783110733679-004>

casionally changing his views, while nevertheless holding onto certain core ideas (about which I will try to say something in a moment).

But first let me say – if I may – a few things about the career of this remarkable philosopher. For, as with so many who are drawn into this field and profession, it was more a fortuitous than a planned occurrence. Tim Scanlon grew up in Indianapolis, his father being the son of an Irish immigrant family who had become a successful lawyer. Tim shared his father's interest in constitutional legal questions, and when he went off to college in Princeton, his intention was eventually to return to Indianapolis to practice law. His main academic interests at the time, however, were in mathematics and philosophy, and thus he focused on logic and the philosophy of mathematics, the subject of his senior thesis with Paul Benacerraf. After finishing college, he went to Oxford for a year to continue to study philosophy rather than law and also became more interested in moral and political philosophy. So when he went to graduate school at Harvard, he wrote a thesis in logic with Burton Dreben, but also developed a keen interest in the issues being pursued by a professor there called John Rawls.

In 1966, he was hired as an assistant professor by Princeton University, where he remained as Professor until 1984, before leaving for Harvard at the request of Rawls to join the faculty there – a clear case of an offer that he could not reasonably reject. At Harvard, he eventually became Alford Professor of Natural Religion, Moral Philosophy and Civil Polity, a chair from which he retired in a beautiful ceremony in spring 2016. On that occasion, Scanlon reminded his audience of the great teachers and colleagues from whom he benefited – some of whom, like Thomas Nagel and Derek Parfit, are present in this room – as well as of the many great students he trained who have themselves become central figures in moral and political philosophy, like Sam Scheffler and Elizabeth Anderson, to name just two among a large and very impressive group. And he reminded us that he developed his ideas at a time that was marked by “great confidence in the possibility of progress through thinking clearly and carefully about philosophical problems.” Scanlon's work has contributed in an exemplary way to foster this confidence, a fact attested to by the many honors he has received, among them the invitations to deliver the prestigious Tanner and Locke Lectures, his appointment as Fellow of the American Academy of Arts and Sciences, and the prize being conferred upon him today.

In his philosophical development, Scanlon gradually moved away from logic and the philosophy of mathematics to political and legal philosophy and in 1971 founded, together with Tom Nagel and Marshall Cohen, *Philosophy and Public Affairs*, which would become one of the leading journals in political philosophy. His first publications in that field immediately made a mark, such as his 1972 essay “A Theory of Freedom of Expression,” a topic in which he had been inter-

ested since his youth and to which he would return regularly in later writings, always carefully revising his theory. The approach he took to rights there is broadly consequentialist in arguing that rights protect especially valuable human interests by avoiding “bad consequences” – and at the same time it has a deontological aspect in according a special place to the value of personal autonomy among the interests to be protected, thus constraining other consequentialist considerations. Here, and this is true of other contexts as well, we can detect the influence of a great liberal thinker, John Stuart Mill, on the thought of another, Thomas Scanlon. I think this is worth keeping in mind if one wants to understand both Scanlon’s strong liberal beliefs and his attempt to overcome the dualism between consequentialism and deontology – though the further development of his thought exhibits an increasing leaning towards deontology, a trajectory nicely documented by the volume of collected papers entitled *The Difficulty of Tolerance*. Like Rawls and many others, Scanlon thinks that considerations of moral rightness should guide the way we think normatively about basic concepts of law and politics. Thus the major move in Scanlon’s thought, I believe, was from the assumption that moral rightness is mainly about inquiring into valuable states of affairs based on judgments about personal well-being to the idea that there is a core morality of interpersonal respect which obliges us to justify our actions to every other person as, in my terms, a justificatory equal, on the basis of general principles which are justifiable to all such persons.

The breakthrough to the approach that Scanlon would call “contractualist” came in a paper from 1982 entitled “Contractualism and Utilitarianism,” one of the most influential articles in recent moral philosophy. The term contractualism can be a bit misleading, because Scanlon neither argues that morality is a matter of rational choice in a situation of limited information nor that morality is fixed by some kind of hypothetical contract. Rather, the approach explains a certain structure of morality and the justification of moral truth by way of the famous formula of moral wrongness:

“An act is wrong if its performance under the circumstances would be disallowed by any system of rules for the general regulation of behavior which no one could reasonably reject as a basis for informed, unforced general agreement” (Scanlon 1982, p. 110).

“Contractualism” is the word for the operative central notion of morality as a system of general norms: a system that treats each person as an equal justificatory authority and does not impose any norms on persons in weak bargaining positions, thus exhibiting an important parallel with discourse ethics. Scanlon stresses that moral norms cannot be the subject of actual consent but that there has to be – let us say – a dialectic between the real, personal interests

of people and general considerations of the value and generalizability of these interests. What is more, Scanlon has taught us that people can reasonably agree to a lot of things, including supererogatory courses of action, so that what is more important is what they can reasonably *reject* as an unjustified imposition. That is the deontological core of contractualism: that each person has a veto right, so to speak, against being left out of or marginalized within the moral justification community. Their standing as equals in this regard must not be violated, and thus the contractualist principle is itself a substantive principle of morality, not a mere methodological device. There is an important parallel to the Kantian categorical imperative in this respect, though not in others, because Scanlon is not convinced that morality can be grounded in a reflection on our rational agency.

Scanlon's suggestion that we should think about morality in such a reflexive and discursive way led to a great many debates in moral philosophy, and eventually he presented his views as a general theory of morality in his magnum opus *What We Owe to Each Other*, published in 1998. This book is too rich in ideas and important theoretical advances to even try to do justice to it here. Suffice it to say that there Scanlon isolates a core morality – of universally binding norms – that he distinguishes from other aspects of morality with a different binding force, which represents a very important move in clarifying our understanding of the term morality. And secondly, while in the original 1982 paper on contractualism the motivational basis of morality was located in a “desire” to act in a way justifiable to others, he now argues powerfully – in line with a more general critique of the role “desires” can play in our moral reasoning, developed, for example, in his more recent book *Being Realistic about Reasons* (2014) – that such desires must be based on a moral reason and thus cannot be basic; rather, the moral reason for that kind of respect towards others is fundamental.

Here we arrive at what I consider to be one of the most fascinating aspects of Scanlon's theory, the famous *Grund* question. If morality is not grounded in our rational agency, in our subjective desires or in cultural life forms, what else is the ground for respecting others as equal justificatory authorities? It is, Scanlon argues, a certain ideal of social relations with others – with all others, one should add. And when Scanlon spells out this ideal of relating to others as justificatory equals, his debt to John Stuart Mill once again becomes apparent in Mill's notion of “the desire to be in unity with our fellow creatures” (Scanlon 1998, p. 154). And although, in contrast to Mill, Scanlon does not speak in terms of sentiment or desire, he does think that what grounds morality is a social ideal, the “ideal of justifiability to others” (p. 155), of respecting them as equals with whom I share a community of moral membership, a truly universal community mediated by rea-

sonable justification – an ideal of “mutual recognition” (p. 162), as he also says. Let me quote the relevant passage from *What We Owe to Each Other*:

“Unlike friendship, morality is commonly seen as a form of constraint, not as a source of joy or pleasure in our lives. I am suggesting, however, that when we look carefully at the sense of loss occasioned by charges of injustice and immorality we see it as reflecting our awareness of the importance for us of being ‘in unity with our fellow creatures’” (p. 163).

Unpacking sentences like these poses a major and fruitful challenge, one which I believe takes us to the core of morality. Lots of things need to be said here, for example why reasonable justifiability is the medium of such unity, but I will leave this for another occasion. I just want to point out that what Scanlon is doing here is giving us a different version of deontological morality than many others which focus on its constraining character – stressing its enabling and, in a sense, liberating character. In an interview with Alex Voorhoeve, Scanlon explains this in a characteristic way, not just by rejecting relativism, but also by stressing the, if you excuse the easily misleading term, “communitarian” character of his thought: “I think this kind of relationship to others is something that people in all cultures and all times have reason to care about. [...] I am saying it is *the* fundamental value” (Scanlon 2009, p. 187).

And when he is asked in this connection about Kant’s kingdom of ends, he responds that, read as an ideal of moral community, Kant’s notion is “very appealing,” and adds jokingly: “So you could say I offer the kingdom of ends on the cheap!”

I find this revealing about Scanlon’s – how shall we say? – deontological communitarianism: it shows his moral ideal of a community with all other human beings to whom we owe respect as justificatory equals.

The same animating thought can be found in his other works in moral and political philosophy. I think it explains why in his *Moral Dimensions* (2008) he stresses that apart from the permissibility of an action as traditionally understood we need to be aware of its meaning “for the agent and others” (p. 4), pointing to the particular character of the reasons we have for treating others in a certain way. Doing the right thing for the right reasons constitutes our moral relationship with others.

Deontological communitarianism – not a good term, I admit, but it sounds catchier than deontological relationalism – is also at work in his important arguments about toleration and equality. For while many regard toleration as a cold virtue of not engaging with others and letting them be, Scanlon shows us that what is really at work here is a deep form of respect for others as social members equal to us even though we reject their beliefs or ways of life. And in his seminal work on equality – which he is currently developing into a book we all look for-

ward to (addendum: published in 2018 as *Why Does Inequality Matter?*) – he not only explains that many of our objections to inequality are really grounded in different concerns and values from equality. He also shows us that the concern that others are not being treated as justificatory equals is central to our notions of social and political equality, stressing in particular the “experiential evil” (Scanlon 2003, p. 212) of feeling inferior and without social worth, “what might be called the loss of fraternity.”

This, I think, is the voice of a philosopher who shows us that liberalism and egalitarianism go together if animated by the moral conviction of what we owe to others as autonomous persons *and* at the same time as members of a community we are all part of, simply in virtue of being human. This is the deep humanism of the philosopher whom we honor today, Professor Thomas Scanlon. Congratulations!

References

- Scanlon, Thomas M., “A Theory of Freedom of Expression,” in: *Philosophy and Public Affairs* 1 (2), 1972, pp. 204–226.
- Scanlon, Thomas M., “Contractualism and Utilitarianism,” in: Amartya Sen and Bernard Williams, eds., *Utilitarianism and Beyond*, Cambridge: Cambridge University Press, 1982, pp. 103–128.
- Scanlon, Thomas M., *What We Owe to Each Other*, Cambridge: Harvard University Press, 1998.
- Scanlon, Thomas M., “The Diversity of Objections to Inequality,” in: Scanlon, *The Difficulty of Tolerance: Essays in Political Philosophy*, Cambridge: Cambridge University Press, 2003, pp. 202–218.
- Scanlon, Thomas M., *Moral Dimensions: Permissibility, Meaning, Blame*, Cambridge: Harvard University Press, 2008.
- Scanlon, Thomas M., “The Kingdom of Ends on the Cheap,” Interview with Alex Voorhoeve, in: Voorhoeve, *Conversations on Ethics*, Oxford: Oxford University Press, 2009, pp. 179–194.
- Scanlon, Thomas M., *Being Realistic about Reasons*, Oxford: Oxford University Press, 2014.
- Scanlon, Thomas M., *Why Does Inequality Matter?*, Oxford: Oxford University Press, 2018.

T. M. Scanlon

Contractualism and Justification

I first began thinking of contractualism as a moral theory 38 years ago, in May of 1979. The idea was not entirely original. I was of course familiar with Rawls' theory of justice, and influenced by it. A more direct source for the idea of justifiability to others as the basis of *morality* lay in remarks of Tom Nagel's.¹ Rawls and Nagel had, perhaps, the wisdom not to try to work out a general moral theory based on this idea. But I naively rushed ahead, and even more naively thought that I could write a book developing such a theory over the course of the 1979–80 academic year, when I would be on leave and a visiting fellow at All Souls College, courtesy of Derek Parfit. All I managed to do, of course, was write one paper, "Contractualism and Utilitarianism."² When I showed an early draft to Parfit, his first response was, "Tim, this is not a moral theory. It is just an expression of your personality."

Judging from what Parfit says in his contribution to this symposium, and in *On What Matters*, his opinion of my project has gone up over the years, perhaps excessively, even though he still has criticisms.³ Partly as a result of these criticisms, my own opinion has become more qualified. I am still strongly attracted to the idea, but aware of faults in my original formulation and of serious difficulties facing any view of this general kind. In this paper I will examine some of these faults and difficulties, and consider how I might respond to them. The responses I survey will in many cases be only provisional.

Contractualism makes the rightness of an action or policy depend on whether it would be permitted by justifiable principles. And it makes the justifiability of principles depend on the reasons of certain kinds that individuals have to accept or reject them. I was drawn to this view because I saw it as a way of providing a general account of (at least a central portion of) non-consequentialist mor-

Acknowledgement: A revised version of the paper presented at the Lauener Foundation conference in September, 2016. Later versions were presented at the Oxford Moral Philosophy Seminar and the New York University Colloquium in Legal, Political, and Social Philosophy. I am grateful to participants at all of these occasions for their helpful comments and questions. Thanks also to Derek Parfit, Husain Sarkar, and Johann Frick for written comments on earlier version.

1 In Nagel 1979, p. 123.

2 Published in Sen and Williams, eds. 1982.

3 See Parfit 2011, Volumes One and Two.

<https://doi.org/10.1515/9783110733679-005>

ality—an account of the moral requirements governing the ways that individuals may be treated. The need for such an account seemed to me to be obvious, as illustrated by my Transmitter Room example⁴ among many others.

Here is a different example. In a famous paper about the justification of punishment, Jeffrie Murphy argued that deterrence-based justifications of punishment were unacceptable.⁵ Leaving aside the familiar problem that deterrence arguments might justify punishing innocent people, he said, even deterrence-based justifications for punishing the guilty are morally unacceptable because they involve treating those who are punished “as mere means.” Only a desert-based retributive justification, Murphy argued, can avoid this problem. He might have added that only a desert-based view can explain the limits on criminal penalties. If extremely severe punishment for some minor offense, such as overtime parking, would sufficiently reduce the occurrence of such crimes, then the relevant cost-benefit analysis might support it, especially since such penalties would rarely need to be enacted.

What such examples show, first and foremost, is just that there are moral limits to the ways that individuals may be treated in order to produce desirable results. This is the element of truth in the idea that individuals should not be treated “as mere means.” Something is a mere means if any way of treating it can be justified by the fact that it promotes some good. Something is not a mere means if it is not like this: that is to say, if there are limits on how it can be treated for the sake of such goods. The fact that something is not a mere means in this sense does not rule out the possibility that treating it in a certain way can be justified by the good effects doing so. It just means that there are limits on when this can be the case.

This idea of “treating someone as a mere means” tells us nothing yet about what these limits are.⁶ To say that an action is wrong because it involves treating someone as a mere means is at best elliptical. It is like saying that this act is wrong because it involves treating someone as if there were no limits on how a person may be treated for the sake of other goods, but without saying what those limits are, or why there are such limits.

Like other deontological theories, contractualism seeks to provide an account of these limits. In the case of punishment, it holds that justifiable criminal penalties are limited because there is a limit to what can be imposed on one person in order to reduce the likelihood of some harm to others. And it holds that

⁴ Scanlon 1998, p. 235.

⁵ Murphy, J.G. 1973.

⁶ I discuss ideas in this paragraph more fully in Chapter 3 of Scanlon 2008.

individuals can be punished only if they have chosen to violate a (justifiable) law, because people have strong reason to reject being subject to penalties that they have no opportunity to avoid.

These limits on morally permissible punishment might be specified in other ways, however. They might, for example, be specified by a system of individual rights, or other non-consequentialist principles, taken as foundational elements of the morality of right and wrong. But such a system of rights or principles needs itself to be justified in some way. An adequate moral theory should provide some systematic way of thinking about what these rights or principles are and why they are the ones that set the limits to how individuals may be treated.⁷ Contractualism aims to provide a way of answering these questions, although as we will see there are questions about how “systematic” its answers can be made.

Let me take a moment to expand on the word “why” in the previous paragraph. In Chapter 4 of Book I of the *Nicomachean Ethics* Aristotle stresses the importance of distinguishing between two directions of argument. We must be clear, he says, about whether we are trying to discover first principles or arguing from such principles. In doing the former, he says, we must start with what is already known to us—what would today be called, using a word I do not find particularly helpful, moral intuitions. That is to say, we discover first principles by employing the method of reflective equilibrium. Engaging in this kind of investigation is one way of “thinking about what the moral limits are on the ways individuals can be treated.” The fact that certain principles emerge from this process, when it is carried out well, can be a reason for thinking that they are correct. But the fact that they result from this process does not tell us *why* these principles are correct. Rather, the principles themselves are supposed to tell us why the judgments that follow from them are correct. This is Aristotle’s second direction of argument. When I said that we needed a way of thinking about the limits to the ways that individuals can be treated and *why* these are the limits, I was saying that we need something like first principles in Aristotle’s sense: a way of thinking about what these limits are that is grounded in an understanding of why these are the limits.

In developing my contractualist view, I was following Aristotle’s model. I was using the method of reflective equilibrium to identify the contractualist procedure of justification, which I thought gave the best account of (at least a portion of) morality, and then taking this procedure to be a way of reasoning “*from* first principles” about what the content of this morality is and why. A more purely intuitive method for investigating non-consequentialist morality

⁷ The need for such an account is emphasized by Samuel Scheffler (2003).

could follow this same path: it would employ the method of reflective equilibrium to identify a set of independent principles that seem to provide the best account of the moral limits on how individuals may be treated, and would then take those principles as explaining why forms of treatment that violate these limits are morally wrong. What this method would not provide would be a more general account of why these principles are correct, and a method of arriving at them. This is what contractualism and consequentialism seek to provide. But perhaps this is too ambitious. Any explanation has to end somewhere, and perhaps it has to end in some set of deontological principles.

In 1975, before I had thought of formulating contractualism as a general moral theory, I offered a different answer to the question of how rights should be understood and defended it in a paper called “Rights, Goals and Fairness” (Scanlon 1978). I wrote there that rights are restrictions on individuals’ discretion to act or not act that are justified by the fact that they are needed to protect important interests and do so at a tolerable cost. I thought of this at the time as a form of modified consequentialism. When I showed this paper to Rawls, he said, with characteristic modesty and gentleness, “What you say seems generally right to me, Tim. But I don’t think you should call it a form of consequentialism.”

Rawls was quite correct, although it took me another four years to figure this out. The account of rights that I put forward in “Rights, Goals and Fairness” shares a structure with the contractualist moral theory that I later adopted. It explained rights by appeal to the interests that rights are needed to protect (that is the interests that are grounds for objecting to principles that would permit actions or policies of the kinds that these rights rule out.) But it also required that the costs of these rights, in terms of other interests, should not be too great (that is to say, it required that insisting on these rights should be reasonable.)

My later contractualist view differs from a consequentialist reading of “Rights, Goals and Fairness” by interpreting these interests in a different way. The reasons for rejecting principles permitting rights violations, for example, is not that the effects of the actions they rule out would be objectively bad things to have happen, but rather that individuals have, personally, good reason to object to being affected in these ways. I suspect that this is close to what Rawls had in mind when he said that I should not call my earlier view consequentialist.

Despite this difference, my contractualist view shares with my earlier view of rights, and with many consequentialist views, what might be called a *normatively reductive* character. These views do not seek to reduce normative ideas to something non-normative. But they do seek to explain principles concerning moral right and wrong in terms of other normative ideas, held to be more fundamental. In the case of my contractualism, these are ideas about the reasons that partic-

ular individuals have to want certain opportunities and to want not to be affected in certain ways. In the case of consequentialism they are ideas about the goodness or badness of certain states of affairs, including, but not necessarily limited to, effects on particular individuals.

It is worth pausing to consider why this reductive strategy should be appealing—why it should seem to have explanatory virtues. Any explanation of right and wrong has to start somewhere, and no starting point can be entirely uncontroversial. But the reductive strategy seems to have two advantages. First the rights, or moral norms that are to be explained are complex: they have conditions of application and admit of various exceptions, which often can be only incompletely specified. The considerations that a reductive strategy appeals to, on the other hand, are comparatively simple. In my view, they are just reasons that individuals in certain circumstances have to object to being affected in certain ways. Second, the moral considerations that are to be explained are explicitly deontic—they are claims about what may or may not be done. The factors that the reductive strategy appeals to, by contrast, are not *in themselves* deontic. The structure within which they play a role (such as a process of contractualist, or rule consequentialist justification) assigns them what might be called minimally deontic status: they are considerations that are relevant for the justification of deontic principles.

A reductive strategy seeks to explain the deontic status of rights and other moral norms by the fact that they are required in order to respond properly to these more basic considerations (e.g., in order to avoid being reasonably rejectable on the basis of these reasons). But the main appeal of the method is not just that it explains this deontic status, but that it provides a way of thinking about the content of these deontic ideas, and a way of explaining and interpreting incompletely specified exceptions.

In my view, for example, the right of freedom of expression consists of a set of restrictions on the powers that governments can have to regulate expression. It is held that these restrictions are necessary in order to protect important interests that we have: interests in being able to express our views, having access to the views of others, and having the broader social benefits of free discussion, and that they provide this protection at a tolerable cost to other interests. In order to decide whether a kind of regulation is compatible with freedom of expression, we need to think about how that power is likely to be used, and how these interests are likely to be affected by its exercise. Similarly, the right to privacy is a set of norms that provide individuals with forms of control over the ability of others to observe them and to acquire information about them. In order to decide whether a given restriction is part of the right to privacy we

need to consider whether it is needed to provide this protection and whether it does so at tolerable cost in other terms.

This seems to me a more plausible and helpful a way of understanding our thinking about these rights than the idea that this simply involves thinking about them “intuitively” as basic deontic notions. This is not because I am suspicious of “intuitive” normative thinking. Quite the contrary, I believe that it is inescapable. The question is what kind of such thinking is most plausible and informative. In this case that seems to me to be “intuitive” thinking about reasons and “normatively reductive” thinking about rights and moral norms. The latter provides, among other things, a plausible way of understanding what may appear to be conflicts between different rights and norms, and of understanding why and how the content of rights and norms can change as circumstances change.⁸

I still find this reductive strategy very appealing. I have just completed a book on equality,⁹ and in the course of the seminar in which I discussed the chapters of this book last year, I recognized as a general theme of that book a strategy of “looking beneath” important moral concepts such as rights, but also liberty, equality, and coercion, in order to uncover the reasons that individuals have to want or to object to certain forms of treatment. It is these reasons that make these concepts important, and determine their content.

It might be said that even if this is the best way to understand political or legal rights such as privacy and freedom of expression, it does not apply as well to more personal moral rights, such as the rights a person has over his or her own body.¹⁰ Perhaps I was misled by the fact that I started off thinking about rights by thinking about freedom of expression. In partial response, I could say that I believe at least some “personal” moral rights, such as rights arising from promises, are best understood in this normatively reductive way.¹¹

There is, however, a serious question about how far this reductive strategy can be carried. One problem is that if the basic elements in this strategy are taken in the simplest form—for example, as reasons of a certain strength for objecting to being affected in a certain way—and if the process of determining normative principles is just a matter of comparing the relative strengths of the reasons that those who would benefit from certain actions have for wanting to be so benefitted and the reasons that those who would be burdened by this requirement have for objecting to it, this would seem to have results that are difficult

⁸ See Scanlon 2009.

⁹ Scanlon 2018.

¹⁰ As Nagel suggests in Nagel 1991, p. 141.

¹¹ As I argue in Scanlon 1998, Chapter 7.

to accept. Parfit points out, for example, that this idea, which he calls the “Greater Burden Claim”, would be extremely demanding. If, as in his Case One, another person would gain more years of life from having one of my organs than I would lose, the Greater Burden Claim would seem to imply that I must give it to him.¹²

Parfit poses this as a problem for my contractualism, taking the Greater Burden Claim as a thesis about when a principle is reasonably rejectable. But the problem would seem to arise for the reductive strategy more generally, whether this strategy is carried out as a form of contractualism or not. This would suggest that the overdemandingness that gives rise to some objections to consequentialism may be a feature of its reductive character, independent of any problems specific to consequentialism, or having to do with aggregation.¹³

One conclusion might be that the reductive strategy cannot be fully carried out, and that any defensible theory will have to presuppose rights or other deontological limits on what can be demanded of individuals. This would leave the problem of explaining how these rights or norms are themselves to be interpreted and justified. Another conclusion would be that either the considerations that are the basic elements of a reductive strategy, or the way in which these are taken to interact in the process of justification that that strategy describes, need to be more complex than in the simplest version described by the Greater Burden Claim. I will return to this question.

First, I need to consider the relation between the reductive strategy and the contractualist idea that actions are morally permissible if principles that allow them can be justified to those who would be affected. As I said earlier, I still find this idea appealing. But it needs to be made clearer what role, if any, the idea of justifiability plays in this account.

A moral theory needs to answer two questions. One is the question of *content*: a theory should provide a characterization of the subject matter of (at least a part of) morality, and a method for thinking about its content. Second, a theory should provide, or at least fit with, some answer to the question of *acceptance*—an explanation of why people should care about morality as the theory has characterized it. Why should they regard the conclusions reached through the reasoning that the theory describes as normally settling the question of what to do? The idea that what morality (or at least a certain central part of morality) requires is that our actions should be justifiable to others seemed to me promising because it could play a role in answering both of these questions: we have reason to care about morality because we have reason to care about the justifi-

¹² Parfit 2011, Volume Two, pp. 192ff.

¹³ See Ashford 2003.

ability of our actions to others, and the idea of justifiability to others can guide us in finding an answer to the question of content.

I will begin with the answer that contractualism gives to the question of content and then return to the question of acceptance, and the relation between the two. Like other versions of the reductive strategy, contractualism seeks to explain the content of rights and moral norms in terms of the considerations that count in favor of having these norms and the costs that must be taken into account in defending them. What is distinctive about contractualism lies in the way it characterizes these basic elements and the account it offers of the way they are combined to yield deontic conclusions.

According to contractualism, the basic elements are reasons that individuals have, such as reasons to want certain opportunities and reasons to object to being treated in certain ways. More specifically, these elements are reasons people have for objecting to principles because of the ways their lives would be affected by having to act in accord with these principles and having others feel free to act in the ways that these principles would allow. This formulation, in terms of reasons for accepting or objecting to principles, is the relevant way to state the view, because the general patterns of action that principles allow for and require have both benefits and burdens that go beyond the effects of individual actions that they involve. Individuals have strong reasons not only to object to specific instances of interference but also to want reasonable assurance that they will not be interfered with in these ways. And it is one thing to be required to provide aid of some kind on a particular occasion, and something else to be required to stand ready to do so on any occasion, should the need arise.

So we have the contractualist formula: an action is wrong if any principle that permitted it would be one that some affected person could reasonably reject. I have said something about the role of principles in this formula, and said something preliminary about the possible bases for rejecting principles, namely the reasons that individuals have to want not to be affected in certain ways. These are what I called “personal” reasons, and I will say more about such reasons later. First I need to say something about the idea of reasonableness that is used to characterize the process through which these reasons are to be combined. Claims about reasonableness are claims about what someone has sufficient reason to do given a certain circumscribed set of considerations. So the claim that a principle could reasonably be rejected is a claim that the way that certain parties would be affected by that principle constitutes sufficient reason for them to reject it, taking into account the relevant range of other considerations. These considerations include, I wrote, “the aim of finding principles

that others also could not reasonably reject.”¹⁴ (Rather than “aim” I probably should have said “the reasons they have to find principles that others also could not reasonably reject,” since aims provide reasons only in virtue of the reasons to have those aims.) This reason brings others in its train, namely the reasons that others have to object to principles on the ground of how they would be affected by them. A reason makes rejection of a principle reasonable only if it constitutes *sufficient* reason to reject that principle taking these other reasons into account.

I have assumed that claims about reasonableness in this sense are not moral claims. That is to say, they are not claims made using moral concepts of right, wrong, obligation, etc., but general normative claims about the strength of reasons. (I will return to the question of how the idea of “strength” is to be understood.) Even if claims about what is reasonable in this sense are not explicitly moral (deontic) claims, the fact that they require the reasons of different individuals to be taken into account and treated symmetrically gives them the right kind of content to have moral significance. For this reason, these claims about reasons seemed to me to be a plausible “starting place” for a reductive strategy. As I have said above in discussing the reductive strategy, however, this assumption may be mistaken. Perhaps some moral notions need to be relied on even at the most basic level. I will return to this question below.

In the version of contractualism I defended in *What We Owe to Each Other*, reasons for rejecting a principle were limited to personal reasons—reasons having to do with the way that an individual in that position would be affected by complying with and having others act in accord with a principle. Contractualist justification as I described it thus excluded what I called *impersonal* reasons—reasons that do not derive from the way in which individuals would be affected if that principle were generally followed. For example, we have reasons not to fill the Grand Canyon with trash. There are, of course, personal reasons not to allow such things, flowing from the reasons various individuals have to be able to experience such wonders in their unspoiled state. But it seems to me that there are also impersonal reasons not to act in these ways: the qualities of these natural wonders themselves give us reason to object to actions that would diminish or destroy them. My exclusion of such reasons from contractualist justification is what Parfit calls my *Impersonalist Restriction*.¹⁵ In addition, the version of contractualist justification that I developed only takes account of the personal reasons of individuals, considered one by one, excluding aggregation of the person-

¹⁴ Scanlon 1998, p. 213.

¹⁵ See Parfit 2011, Volume Two, p. 214.

al reasons of many individuals. This is what Parfit calls my *Individualist Restriction*.

These restrictions were distinctive features of contractualism as I first presented it, and some of its most controversial aspects. Excluding aggregative reasons allows contractualism to avoid conclusions that I, at least, regard as implausible, such as that it is permissible, or even required to let one person suffer in order not to deprive a very large number of people of some very small benefit. But excluding these reasons also appears to lead to other conclusions that may seem implausible, such as that it is permissible to save one stranger from drowning rather than five, or ten. Parfit has argued that allowing appeal to impersonal and aggregative reasons would avoid these consequences, and would in general strengthen the contractualist view without depriving it of its distinctive character.¹⁶ In order to assess this proposal, I need first to consider what led me to exclude reasons of these kinds to begin with.

Here I need to turn to the question of acceptance, the problem of explaining the reasons we have to accept moral requirements and the significance that is properly attached to them. In “Contractualism and Utilitarianism” I said that my aim was to provide an account of “morality” and its motivational basis. But by the time I wrote *What We Owe to Each Other* I had become convinced that the realm of morality, as that term is commonly used, is not a unified normative domain. So the question of acceptance does not have a single answer when asked about morality in general. Failure to take proper care of one’s children, and willingness to despoil the environment, are both properly called moral faults. But the reasons lying behind them are quite different, and also different from the reasons supporting general duties to keep one’s promises and refrain from acts that would harm strangers. (They have different “motivational bases,” as I would have put it in “Contractualism and Utilitarianism”.)

This pluralist view of morality as a whole seems to me to be supported by what I call the *remorse test*. This is the idea that the particular way in which it is appropriate to feel bad about doing something that is wrong indicates the distinctive kind of wrongness that is involved, and indicates in turn the kind of factors that make an action wrong in that particular way.

The kind of remorse made appropriate by wrongs of the kind I had in mind in developing my contractualist view is grounded in the sense that an individual who is affected by our action has a reason for objecting to it that cannot be an-

¹⁶ Parfit 2011, Volume Two, pp. 231–243.

swered satisfactorily.¹⁷ As an example, I cited in my book my reaction to Peter Singer's argument about the moral case for contributing to famine relief (Singer 1972). What affected me so strongly in this case was the fact that these starving people had reasons for wanting help that I was failing to respond to rather than just the objective badness of the fact that they were starving.

These different forms of remorse may both be appropriate. But they reveal different understandings of the basis of the relevant moral requirements—that is to say, different answers to the question of acceptance.¹⁸ On the latter view, because human beings have a particular kind of value, it matters what happens to them. In particular, it is a bad thing if they suffer, and we have reason to prevent that. Non-human animals, and perhaps works of nature, are also valuable, and there are reasons not to harm them as well. If the case of human beings is distinctive, on this view, it is because human beings are more valuable, or valuable in a different way than these other creatures and things.

On the alternative view, indicated by the kind of remorse I described myself as feeling in response to Singer's article, what is special about human beings is that they themselves have reasons, including reasons to want not to be treated in certain ways. The fact that human beings have reasons opens up a further dimension of our relations with each other—namely the degree to which we are treating each other in ways that are in accord with principles that none of us has sufficient reason to object to.¹⁹ We have reason to be concerned with this domain of interpersonal morality because we have reason to want to be related to other rational beings in a way that is responsive to the reasons they have. Realizing that we have not done this is what triggers the distinctive kind of remorse I was feeling.

I described this kind of remorse as the sense that one's actions cannot be justified to those whom they affect, and I took this to be a distinctively "contractualist" answer to the question of acceptance. This answer also seemed to me to indicate a distinctive answer to the question of content, according to which the

17 As Frances Kamm suggests, this makes it plausible to say this person has been *wronged* by such an action, not just that the action is wrong. See Kamm 2002, p. 336, and Frick 2016, p. 262, note 48.

18 I am indebted to Victor Tadros for a conversation in which he urged on me the following way of seeing the issue, on which he takes the opposite view to mine. Frances Kamm may also be taking that position when she recommends that I adopt "the simple idea that the value of persons is objectively great." See Kamm 2002, p. 329.

19 It was because of this emphasis on the reasons individuals have for accepting or rejecting principles that I called my view "contractualism," a name that may in some ways have been misleading.

rightness or wrongness of an action depends on whether any individual has sufficient reason to reject a principle that would permit it, taking the corresponding reasons of others also into account.

The idea of justifiability to a person thus plays a role both in answering the question of acceptance and in the contractualist answer to the question of content. But the forms of justifiability involved are different. The contractualist answer to the question of content is that in order not to be wrong an action must be allowed by a principle that is justifiable to every person in a specific way: the (personal) reasons that any individual has for objecting to this principle must not be sufficient reasons for rejecting it. This form of justifiability is thus a stylized one, specified by the distinctive contractualist idea of reasonable rejection. The question of acceptance is the question of whether this standard (of acting only in ways allowed by principles that are justifiable to all in this special sense) is one that everyone has sufficient reason to accept, taking into account all the reasons that a person has.²⁰

The idea of reasonable rejection as an answer to the question of content is incomplete unless the relevant range of reasons is specified. For example, such an account would be trivial if it allowed that principles could be rejected on the ground that they would permit actions that are morally wrong. As I said earlier, the version of contractualism stated in my book restricted these to what I called “personal” reasons, excluding impersonal reasons and aggregative reasons.

In taking this position, I was guided (perhaps misled) by my understanding of the remorse test. Actions are wrong in the particular way I was concerned with—and thus trigger a distinctive kind of remorse—because of the reasons that individuals have to object to being treated in that way. These cases are quite different from ones in which the reason against acting in a certain way is an impersonal reason, such as the kind of reason that we have not to fill the Grand Canyon with trash, a reason grounded in the value of such natural wonders themselves (and going beyond the personal reasons that individuals have for wanting to be able to experience such natural wonders in their unspoiled state).

If there are such impersonal reasons for preserving natural wonders, then it would seem that there are also impersonal reasons to object to others’ acting in

20 This may be what Frances Kamm has in mind when she says that contractualism rests on an “objective theory of value.” See Kamm 2002, p. 329. If what she means is that the adequacy of its answer to the question of acceptance depends on the facts about all the reasons that individuals have, then I agree. This dependence is not something distinctive about contractualism, however, but rather a general truth about the question of acceptance that any moral theory needs to answer. The contractualist answer to the question of content also depends on an objective understanding of (a certain range of) the reasons that individuals have.

these ways that would destroy them, and therefore reasons of this kind to object to a principle that would permit such actions. It might even be said that these reasons would make it reasonable to reject such principles, and that the actions that they would allow could not be justified to others, since they would have sound (impersonal) reasons for objecting to them.

A person who raised this objection, however, would merely be calling attention to reasons provided directly by the grandeur of the Grand Canyon, reasons which have nothing to do with him or her. Failure to respond to these reasons would be a fault, and would properly trigger a kind of remorse. But it is different from the kind of remorse that is triggered by the realization that one has treated a person in a way that he or she has personal reason to object to.

It might be said, however, that the fact that the person has a reason to reject a principle permitting a certain action is redundant. If the person has such a reason, it is only because the effects of actions of that kind would be bad for him, say by harming him in some way. But, it might be said, the badness of these effects is, in itself, a reason not to act in such a way. That the person has reason to object to a principle that would permit such actions is just another implication of the badness of these effects. This badness, it might be said, is normatively basic and is sufficient to explain the wrongness of the actions in question. But this account leaves out something important. The badness of a person's being affected in a certain way may itself count against treating him in that way, but the fact that he has sufficient reason to object to a principle that would permit this treatment is crucial to the kind of wrongness that triggers the remorse that I am describing.

The idea that only personal reasons can make an action wrong in this way was, so to speak, the theoretical basis of my decision to exclude impersonal reasons in general and aggregative reasons in particular from the class of reasons for rejecting a principle. (I refer to this as my "theoretical" basis to contrast it with the aim of avoiding implausible conclusions that aggregative reasons might lead to in cases like my Transmitter Room example.) There are, however, other cases in which it seems clear, intuitively, that it can be morally wrong to fail to save a greater number of people. So I need to consider how these cases can best be explained, and whether this can be done in a way that is consistent with this theoretical basis for holding that an action is wrong in the sense that contractualism is describing only if an individual in some position has sufficient personal reasons to object to a principle that would permit it.

One possibility, considered by Parfit, is a kind of moral pluralism, according to which actions that are not wrong in the contractualist sense might be wrong in

some other sense.²¹ Suppose I can rescue either, but only one, of two groups of people, one group containing 5 people, the other containing 20. Perhaps either course of action would pass the contractualist test. Since individual members of the two groups have identical reasons for wanting to be saved, and the fact that there are more individuals in one group does not give any of the individuals in that group a stronger reason for demanding to be saved, it appears that it would not be reasonable for any of them to reject a principle requiring one to save the other group. It thus seems that saving either group would be permissible according to contractualism. But even if this is so, and no individual would be wronged by either course of action, it could be added that it would show a lack of regard for the value of human life to save the smaller number. To do so would thus be morally wrong in a different way than the one that contractualism describes.²²

I am not opposed to the possibility of moral pluralism of this kind.²³ There are different conceptions of morality, revealed by applications of the remorse test. The question to ask about them is not “Which one gets it right about what *morality* is?” but rather “Which of them do we have reason to treat as guides of conduct, and in what way?” This leaves open the question of what one should do when two forms of reasoning yield conflicting directives. This would not be a *moral* question in the sense of “moral” described by either of the two conceptions. It would, rather, be a question about what we have most reason to do in a more general sense, not itself backed by any distinctive form of remorse.

This pluralistic explanation of the case for saving the greater number might, however, lead to a kind of implausible asymmetry between what might be termed subset cases and disjoint set cases. If I could save a whole group of people, it would seem clearly wrong to choose to save only a subset of them, when it would be just as easy for me to save them all. Contractualist reasoning could explain why this would be wrong: someone in the position of those who are not saved would have reasons to reject a principle permitting this, and neither the agent nor anyone else would have any reason to object to a principle requiring one to save all. Contractualist reasoning would not, however, explain why it would be wrong to save the smaller number of people in a case that was like this except that the two groups were disjoint, and it may seem odd to say that

²¹ See Parfit 2011, Volume Two, pp. 213 ff.

²² This possibility is discussed in different ways by Elizabeth Anscombe (1967) and Veronique Munoz-Dardé (2005). A stronger kind of pluralism has been suggested more recently by Johann Frick near the end of Frick 2015. I will discuss Frick’s view below.

²³ See also Scanlon 1998, pp. 342 ff.

these two actions are wrong “in different ways,” as the pluralist explanation would imply. A more unified account would be more plausible.

It now seems to me on reflection that what I referred to as my theoretical reason for excluding aggregative reasons from the process of contractualist justification was too hasty. The remorse test suggests that when actions are wrong in the way that contractualism aims to describe they are wrong because of reasonable objections that some individuals would have to principles permitting them, objections based on personal reasons having to do with the way they would be affected by living under those principles. I concluded, further, that whether such an objection is reasonable depends *solely* on the relative strengths of the personal reason supporting that objection as compared with the strongest personal reason that any other individual has for objecting to alternative principles. The latter step led me to the Individualist and Impersonalist restrictions. But this step was too quick.

Even if impersonal reasons, or aggregative reasons, cannot be grounds for rejecting a principle, these reasons might play a role in determining whether such rejection is *reasonable*. Return to the example of the choice between rescuing a group containing 5 people, and rescuing a group containing 20. It might be said that while it would be reasonable for a person in the larger group to reject a principle permitting one to save the smaller number, it would not be reasonable for a person in the smaller group to reject a principle requiring one to save the greater number. This is not because an individual in the larger group has a stronger reason to be saved than the corresponding reason of someone in the smaller group, but because it is unreasonable to insist on this given the many individuals with strong reasons against it. (I leave aside for the moment whether this is a kind of impersonal reason or not.) Even if these reasons did not provide grounds for rejection, they could be considerations that can make such rejection *unreasonable*. At least the remorse test does not rule this out.

So the line of thinking that seemed to me to provide theoretical grounds for excluding impersonal and aggregative reasons from the contractualist justification procedure appears to be flawed. The remorse test led me to focus on reasons for rejecting a principle, neglecting the fact that other reasons are relevant to the reasonableness of rejection. In particular, the number of individuals that are affected could make an objection to a principle unreasonable. This raises the question of how this might be so, and raises in turn a larger question about the idea of the strength of a reason and the role that this idea plays in the idea of reasonable rejection.

In the examples that Parfit presents, we are asked to choose between benefitting (or burdening) a person in one position and benefitting (or burdening) a person in another position, and these benefits and burdens are described in

terms of quantifiable changes in welfare, such as years of added life or hours of pain. These cases invite the idea that the reasonableness of rejecting a principle depends simply on the comparative strength of individuals' reasons for and against it, measured in these terms.

This way of looking at the matter is misleadingly simple, however. Whether one party in such a case has sufficient reason to object to a principle can depend on many factors. It can depend not only on the change in welfare involved but also, as Parfit notes, on how badly off each party would be if not aided. It can also depend on the opportunities each person has to avoid bearing these burdens. A relevant notion of the strength of personal reasons would need to incorporate the normative significance of diverse factors of these kinds.

I have come to believe that the relative strength of reasons is not a basic notion that can play this role.²⁴ The fundamental notion is the idea of being a *sufficient* reason for a given action or attitude in certain circumstances. Judgments about the “strength” of reasons should be understood simply as claims that certain reasons are or are not sufficient in cases in which certain conflicting reasons apply, or often as claims about whether they would or would not be sufficient in certain counterfactual circumstances in which there were such reasons. There is no independent normative property of strength that provides the basis for arriving at conclusions about sufficiency.

The question at issue in the contractualist idea of reasonable rejection is whether the way in which a person in a certain position would be affected by a principle—either by what it would require such a person to do or by what it would allow others to do—is a sufficient reason for rejecting it, taking into account the other factors that the contractualist procedure identifies as relevant. It follows from the general point I have just made about strength that conclusions about reasonable rejectability cannot be based on some supposedly independent idea of the relative strength of opposing reasons.

This view of reasonable rejection allows for the possibility that in some cases what makes it unreasonable to reject a principle is just the costs that alternative principles would involve for an individual in some other position. But the view also allows for the possibility that the sufficiency of a reason for rejecting a principle can depend on a broader range of factors, not just on the particular reason some other individual has for objecting to alternative principles.

Consider again cases in which an agent could, with the same ease, save either a larger or a smaller group of people, and suppose that the two groups are

²⁴ I defend the following view about the strength of reasons in Chapter 5 of *Being Realistic about Reasons*.

disjoint. It seems to me plausible to say that a member of the smaller group could not reasonably reject a principle requiring the agent to save the larger number. This would not be because it is impersonally better for more lives to be saved, but rather because the reasonableness of rejecting the principle must take into account the reasons that all the individuals involved have to want to be saved, and the fact that there are more of these reasons counting in favor of the other action.

Can we also say, on the other hand, that members of the larger group *could* reasonably reject a principle permitting the agent to save the smaller number? Their personal reasons for objecting to such a principle are identical to the reasons that members of the smaller group have for favoring that principle. So, comparing these reasons one by one, it might seem to be a standoff. But it would be unreasonable, for the reasons just surveyed, for members of the smaller group to insist on a principle that would favor them. So it would be reasonable for members of the larger group to reject a principle allowing the agent to save the smaller number, even taking into account the reasons those in the other group have to object. If this is correct, it follows that saving the smaller group would be wrong, in the way that contractualism describes.

This argument would not apply in cases like my Transmitter Room example, in which the reasons of those in the larger group are very different. If, as I assumed, the pain of the person trapped behind the transmitter is very great, it does not seem unreasonable for a person in that position to reject a principle requiring that his rescue be delayed until the game is over. The fact that many others have reason to object to their enjoyment of a sporting event being interrupted does not make this unreasonable.

This way of interpreting the idea of reasonable rejection avoids the implausible aggregative conclusion by rejecting the idea that the reasons that individuals have to want the broadcast not to be interrupted can be identified with the degree of pleasure or well being involved, and that the strength of these reasons “adds up,” so that the collective strength of these reasons can become arbitrarily large, as the numbers of people watching the game increases. My readiness, in *What We Owe to Each Other*, to exclude aggregative considerations altogether was due in large part to the mistaken supposition that if the number of people affected by a principle did make a difference it would have to be in this additive way.

Interpreting reasonable rejection in terms of a more abstract idea of the sufficiency of a reason under certain conditions allows for the possibility that in some cases the fact that many people have opposing reasons can make a given reason insufficient ground for rejecting a principle although in other cases, such as Transmitter Room, this is not the case. Whether the fact that

many people have certain conflicting reasons makes a single person's reason insufficient grounds for rejecting a principle is a question that is settled, as it were, at the level of reasons. This seems to me the correct place to locate the question. For one thing, the question in this form (as a question about reasons) cannot be evaded. If it does not arise within contractualism, as a question about the reasonableness of rejecting a principle, then it will arise externally, within a pluralist view, as a question about what to do when conclusions about reasonable rejection conflict with reasons of other kinds. And my main reason for preferring the latter alternative, deriving from the remorse test, now seems to me mistaken.

This shift comes at the price of putting a great deal of weight on judgments about the sufficiency of reasons, in a way that weakens the explanatory character of this contractualist version of the reductive strategy. If we say in some cases, such as Transmitter Room, that certain conflicting reasons are, in Frances Kamm's phrase²⁵ "not relevant" to a certain decision, what kind of relevance is in question? Judgments of this kind look more like moral judgments than (apparently) simple comparisons of the strength of reasons did.

Here is one way to bring out this explanatory loss. If we stick with the idea that the reasonableness of rejecting a principle depends on whether a reason for objecting to it is stronger than the strongest reason any other individual has for objecting to alternative principles, then the question of whether this is so might be answered by invoking the *Intrapersonal Test*, according to which the reason provided by benefit B is stronger than that provided by avoiding cost C if gaining B would be sufficient (prudential) reason for a single individual to choose a certain option even if it involved cost C.²⁶ This test might be appealing because it indicates that the comparisons involved are not moral judgments.

Something like this test may be at work behind many of the examples that Parfit discusses, involving years of life, loss of limbs, and hours or days of pain. The test could allow for the significance of the "base line" of welfare at which a

²⁵ See Kamm 2007, pp. 33–37.

²⁶ Since I am not endorsing the *Intrapersonal Test*, I will not explore the question of exactly how it should be formulated. But the version I have just given does not, it seems to me, capture a sufficient condition for drawing a morally significant conclusion. Gaining a certain increase in expected life expectancy might be a sufficient reason for a person to join a "body lottery scheme." But even if this is so, it would not follow that the donation of vital organs that such a scheme would involve could be morally required. One possible explanation for this is that the question of whether an individual would have sufficient reason to *choose to join* and to accept the terms of such a scheme build in an element of voluntary acceptance that would be missing from the moral requirement. It is one thing to give up a vital organ because one agreed to do so, and something else to be morally required to do this whether one would agree to or not. Johann Frick makes this point in Frick 2013.

person would be left if unaided, as in Parfit's *Contractualist Priority Rule*, since this base line may be relevant to the strength of an individual's purely prudential reasons. Moreover, the test does not depend on whether the strength of a reason is taken as a basic notion or as dependent on a prior idea of a sufficient reason. It could be stated in either of these ways. But it seems not to allow for the possibility that the number of people whose lives would be saved by following a principle could be relevant to the sufficiency of a reason for rejecting that principle. As far as I can see there is no way of representing this factor in intrapersonal terms.

If this is so, the interpretation of reasonable rejection in terms of a broad notion of sufficiency weakens the explanatory power of this version of the reductive strategy in two ways. First, for the reasons just mentioned it makes judgments about the reasonableness of rejecting a principle less obviously "non-deontic." Second, the resulting view seems to offer less clear guidance than a view that made reasonableness depend on an independently understandable notion of the strength of individuals' reasons. These changes do not, it seems to me, render the view empty. This version of contractualism still has the virtue of the reductive strategy in general: it offers a distinctive characterization of the basic determinants of right and wrong—namely the reasons individuals have to reject principles that would affect them in certain ways—even if it does not provide a clear standard for assessing these reasons. Perhaps no view can do that. But the decrease in explanatory strength should be acknowledged, even if it is a price that must be paid.

Parfit argued that dropping both the *Individualist* and *Impersonalist* restrictions would strengthen my contractualist view. This modification, he thought, would avoid implausible conclusions without sacrificing the basic idea that makes the view appealing. In my earlier thinking about the matter, I took aggregative reasons to be a species of impersonal reasons. It seemed to me that in order to "make the numbers count" it was necessary to place value on the *sum* of individual benefits, and that a reason for promoting such a sum had to be impersonal. Parfit maintained that this was a mistake, and I have just suggested, in a different way than Parfit, that the number of people who are affected by a policy can be relevant in a way that relies only on the personal reasons that they have.

Whether or not this amounts to "dropping the individualist restriction," it is a way of making the numbers count that is compatible with the basic ideas of contractualism. This raises the question of whether something similar could be done with regard to impersonal reasons. The idea would be to continue to hold that the reasons that make an action wrong in the way contractualism describes must be personal reasons to reject any principle that permits such ac-

tions, but to allow that personal reasons can fail to be sufficient grounds for rejecting a principle because of countervailing *impersonal* reasons.

This strategy is not, however, fully successful. Suppose that something of impersonal value would be destroyed if many people engage in some activity that it would be mildly inconvenient for them to avoid. For example, if many people walk over the dunes near a beach, this may cause erosion, destroying dunes that are the habitat of many plants and birds. Because of the impersonal value at stake, the inconvenience involved might be insufficient reason for rejecting a principle that forbids people from doing this. In order to explain why walking over the dunes would be wrong, however, we need to cite some reason for rejecting a principle what would permit this. In the cases of aggregative reasoning discussed above, some individuals had strong personal reasons for rejecting a principle permitting one to save the smaller number, reasons which were made sufficient by the fact that opposing reasons were not reasonable. In the present case, however, if the impersonal reasons for protecting the dunes are only reasons for rejecting a principle that prohibits walking over them, they do not suffice to make dune walking wrong in the way that contractualism describes. Allowing that impersonal reasons could be what makes it reasonable to reject a principle permitting dune walking would thus involve a more significant departure from contractualism. Moreover, the remorse test seems to me to indicate that the wrongness of acting in this way flows directly from the disvalue of destroying the dunes, not from a failure of justifiability to others. So a form of pluralism still seems to me a more plausible way of accounting for such cases.

This analysis sheds some light on a different class of cases discussed by Johann Frick.²⁷ These are cases in which we face a choice between a policy that brings certain benefits to a small number of people and one that benefits a larger number of people by reducing their risk of some harm. Frick considers cases in which the choice is between rescuing a small number of miners who are already trapped in a mine and investing the same resources in improved mine safety, which would save the lives of miners in the future. He argues convincingly that, contrary to what I maintained in *What We Owe to Each Other*, the benefits of improved safety should be assessed *ex ante*, as a reduced likelihood, for current miners, of being trapped in the mine in the future. If this reduction in risk is quite small, say from 3% to 1%, and the number of affected miners not very large, then the reasons that the trapped miners have for being saved seem sufficient grounds for rejecting a principle requiring, or even permitting, investment in mine safety rather than on rescuing them.

27 In Frick 2015, pp. 219ff.

Frick points out, however, that as the number of miners becomes larger it becomes extremely likely that a larger number of miners will be saved by the latter policy. It may, he says, be difficult to accept a policy of saving one miner now if this means 20 deaths later. If we consider only *ex ante* reasons, as Frick argues we should, then the great likelihood of these deaths does not alter the personal reasons that current miners have for favoring investment. The reduction in their risk of death remains very small. So if the likelihood of these deaths counts in favor of that policy, and against rescuing the miner or miners who are currently trapped, it can only be as a form of impersonal value. Frick concludes that if failing to prevent these future deaths would be an interpersonal wrong this could best be explained by a form of pluralism within interpersonal morality, one part of which is responsive to the reasons that individuals have, the other responsive to the impersonal value of their lives.

An alternative would be to say, parallel with the case of dune walking just discussed, that the impersonal value of the 20 “statistical” lives that would be saved when the number of future miners is larger is relevant because it makes it unreasonable for the trapped miner or miners to insist on being saved. However, in contrast to the dunes case, in which no personal reasons counted in favor of prohibiting dune walking, in this case many miners have personal reasons to want their risk of being trapped to be reduced (even slightly). Perhaps these reasons (even if they are no different from the reasons they would have if their numbers were smaller) could be sufficient reasons for them to reject a policy requiring us to save those who are currently trapped, given that the reasons of those currently trapped to insist on rescue are insufficient. It is not *obvious* to me that this is so (or obvious what the right thing to do is in such a case). But it is a possibility opened up by the interpretation of reasonable rejection that I have been discussing.

With this as background, let me return to the question of overdemandingness as a problem for contractualism and for the reductive strategy more generally. Parfit presented the problem as it arises from the Greater Burden Claim, which holds that “it would be unreasonable ... to reject a principle because it imposed a burden on you when every alternative principle would impose much greater burdens on others.”²⁸ Parfit presents as an example his Case One, in which the Greater Burden Claim would seem to require a person to give up one of his vital organs if that organ would give some other person more years of added life. But there is a much wider range of cases in which the Greater Bur-

28 Parfit 2011, Volume Two, p. 192, citing “Contractualism and Utilitarianism”, p. 110.

den Claim would seem to require implausible levels of sacrifice, and this makes the problem particularly threatening.

I have already explained why it would be overly simple to think that the reasonableness of rejecting a principle is always determined only by the comparison of individual benefits and burdens in this way, although there are cases in which this appears to be true. It appears to be true, for example, in Peter Singer's pond example, in which a lone passerby can save a child from drowning at the cost of wet clothing and perhaps a missed appointment (see Singer 1972).

But what is right or wrong in this case is not settled simply by a comparison of these particular benefits and burdens (some soiled clothing and shoes and a missed appointment vs. a child's life.) First, there is no alternative way in which the child's life might be saved in cases of this kind other than by the passerby's making this sacrifice. (The relevance of this fact is recognized in the Greater Burden Claim: any alternative principle would lead to the child's death.) Second, as I mentioned earlier, the reasonableness of rejecting a principle depends on the burdens that would be involved in complying with it *in general*, not simply on the sacrifice it would involve in a particular case. Cases like the pond example are rare. So the costs of complying with a principle requiring one to rescue a person in every such case one encounters are not very great—probably not greater than those involved in a single case. There is no “intrapersonal aggregation” of these burdens.

Things are very different in the case of what might be called the Greater Burden *Principle*, which would require us to decide what to do in every instance by comparing the benefits and burdens to everyone affected by our choice. This principle could reasonably be rejected because complying with it in general would have a paralyzing effect on our lives.²⁹ If we had always to take the interests of other potential users into account, giving them the same weight as our own comparable interests in each individual case, we would not be able to give any preference to individuals toward whom we have special relations and would not be able to plan effectively to carry out any project, since other more pressing uses would be very likely to arise.³⁰ Moreover, in many cases there will be other, less costly, ways of promoting the benefits in question.

²⁹ This response is obviously similar to the objection Bernard Williams raised to utilitarian or Kantian morality in Williams 1981. One difference is that I am raising it as a point about the content of contractualist morality, whereas his point was a negative answer to what I am calling the question of acceptance, given what he took the content of these forms of morality to be.

³⁰ Parfit offers a somewhat similar response on pp. 209–212, focusing on what it would be like to have these demands on our resources and our bodies be enforceable by law. Like the argument I have just sketched, his relies on considering the general consequences of living under

The relevance of alternative principles is a way of capturing Liam Murphy's idea that preventing harms of the kind in question may be a collective project—that is to say, that others may also have obligations to contribute to doing this.³¹ This leads to the question, which Murphy focuses on, of what must be done in cases in which others are not complying with the “first-best” principles for dealing with a problem. (In my terms, the question is which principles specifying such “second-best” duties could not reasonably be rejected.) It seems to me that the consideration of “intrapersonally aggregated” cost of general compliance plays the primary role here. For example, in Murphy's pond case in which there are two drowning children and two passers-by one of whom refuses to help because he hates children, the other person could not reject a principle requiring him or her to save both children. This involves greater sacrifice than what the fairest principle would require, but not a great enough sacrifice to ground reasonable rejection.

As I said earlier, although Parfit presented the problem of the overdemanding nature of the Greater Burden Claim with reference to his Case One, in which it seemed to require giving up one's vital organs in any case in which they would provide more years of life to some other person, the problem is much broader, since there are many cases in which that claim, if interpreted as what I have called the Greater Burden Principle, would require implausible levels of sacrifice. The two factors I have just mentioned—the need to consider the burden of complying with a principle *in general*, not just in a particular case, and the relevance of alternative principles that would provide the needed aid—seem to me to be the main bases of a defense of contractualism against the charge that it is overly demanding in this way. But these factors do not fully account for Parfit's Case One.

If giving up a vital organ involves dying immediately, then one cannot be required to do this more than once. So we cannot appeal to the costs of repeated application of a principle requiring this sacrifice in order to explain why it could be rejected. Nor does the possibility of a less burdensome alternative seem to avoid the problem. Being required to make this sacrifice seems objectionable even if the principle requiring it spreads this burden as widely and fairly as possible, say by a lottery among possible donors. The case thus raises a question about whether contractualism, or any form of the reductive strategy, which focuses solely on the benefits and burdens that a principle involves, can give a

a principle, rather than considering each case separately. But I think that once one sees that this is the way to look at the matter the point holds at the level of moral requirements, without taking enforcement into account.

31 See Murphy, L. 1993.

full account of the limits within which constrained versions of the Greater Burden Claim might apply.

This problem is brought out by the apparent asymmetry between the following two kinds of cases. In the case that Parfit presents, the principle in question would require a person to give up one of his or her vital organs for donation if this organ would give some other person more years of continued life. I assume that in this case donating this organ means dying immediately. Suppose, in a second kind of case, that a person will die of organ failure if he does not receive a transplant immediately. Some suitable organs are available, from people who have donated their organs after dying from other causes. These are, however, in short supply. It seems to me that a person could not reasonably reject a principle according to which in such a case the scarce organs should be allocated to those in immediate need whose lives will be extended the most. But the benefits and burdens on the various sides in this case (in terms of years of future life) may be exactly the same as in cases of the first kind, like the one Parfit presented. The question is why a comparison of the benefits can settle the matter in a case of this latter kind but not in one of the former kind, which Parfit described. The difference between the two suggests that, intuitively, what matters is that, in Parfit's case, the organs are "his" or "hers" and that a person has a right to his or her own organs. This is a problem for contractualism and for reductive views in general since the significance of this fact is not captured by the benefits (extended life or quality of life) that the person in a case of the first kind gains by keeping his own organs.

This background assumption is clear in something Parfit says in presenting another of his examples. He writes, "... we can suppose that certain people have painful diseases and that as doctors who have scarce medical resources we must decide which of these people we must treat. None of these people has any special claims, nor do they differ in any other morally relevant way."³² This presents the case within a particular moral background, in which there are "resources" on which none of the possible recipients have any special claims. In situations of this kind, the Greater Burden Principle has considerable plausibility. The question is when is this the case: when are certain "resources" ones that no one has any special claim to? Intuitively, it seems that Parfit's Case One is not a situation of this kind: people do have "special claims" to their own vital organs. The challenge to contractualism and to reductive views in general is whether they can explain when these conditions apply, or whether such views are applicable only within moral limits set in some other way. What I called above the

³² Parfit 2011, Volume Two, p. 196.

Greater Burden Principle would treat our time and energy as resources on which no one has any special claim. So if I was correct in claiming that this principle could reasonably be rejected, then this is one claim about these limits that contractualism can explain. The question is whether a contractualist view (or any reductive view) could explain these limits more generally, without an appeal to independently given rights or other deontic norms.

One possible explanation would be that a person has stronger reason to object to losing some benefit he or she presently has than to object to not receiving a similar benefit (at least as long as the way the person came to have that benefit was not itself subject to some objection). The idea would be that this makes it reasonable for the person in Parfit's case to reject a principle requiring him or her to give up the organ. In cases of the other kind, however, neither of the affected parties has the organ to begin with. This explanation would amount to a kind of status quo bias at the level of reasons (as long as the status quo did not come about in a way that someone has sufficient reason to object to). A second possibility would be to appeal to the fact that giving up the organ in the first case would require cutting into the person's body, and to claim that each person has sufficient reason to reject a principle that required him or her to submit to this.

It may be said that these explanations are insufficient, and that the only adequate explanation of the fact that a person's organs are not resources that should be distributed to those who would benefit the most from their use lies in the fact that individuals have a right to their own organs. But that leaves us with the questions of how the claim that there is such a right is to be defended, and in particular whether this claim does not depend on claims about reasons of the two kinds just mentioned. Are rights just a way of building in a "status quo bias"? If not, why not? But if so, why is it an objection to a reductive view that it forces us to make this bias explicit?

In this paper I have tried to clarify the line of thinking that led me to my contractualist theory of right and wrong, to identify some problems for that view, and to consider some ways of responding to these problems. I have described contractualism as an instance of a more general normatively reductive strategy in moral philosophy, a strategy that seeks to explain more complex deontic notions in terms of simpler ones. Contractualism is distinctive in taking these most basic elements to be the reasons individuals have to object to or to favor principles because they would affect their lives in certain ways.

Actions are wrong in the way that contractualism describes if the objections of this kind to principles that would permit such actions cannot be answered adequately. I put this by saying that such actions are not justifiable to those whom they would affect. They therefore make appropriate a distinctive kind of remorse. But I erred in concluding, from the fact that actions are wrong in this way only if

individuals have sufficient personal reasons for objecting to principles that would permit them, that the reasonableness of rejecting a principle depends only on the comparison of individual personal reasons.

Taking account of the fact that this does not follow, and of the more general fact that the strength of reasons is not a basic notion, but depends on the idea of a sufficient reason, I have described a revised version of the contractualist idea of reasonable rejection, which allows for the numbers of individuals affected by an action to be relevant to its permissibility. Finally, in the light of this modification, I considered the resources that a contractualist view, or any normatively reductive view, has to set the limits to the sacrifices that individuals can be required to bear.

References

- Anscombe, Elizabeth, "Who is Wronged?", *Oxford Review* 5 (1967), pp. 16–17.
- Ashford, Elizabeth, "The Demandingness of Scanlon's Contractualism," *Ethics* 113 (2003), pp. 273–302.
- Frick, Johann, "Uncertainty and Justifiability to Each Person: Response to Fleurbaey and Voorhoeve", in Nir Eyal, Samia Hurst, Ole Norheim and Dan Wikler (eds.), *Inequalities in Health: Concepts, Measures, and Ethics*, New York: Oxford University Press, 2013.
- Frick, Johann, "Contractualism and Social Risk", *Philosophy & Public Affairs* 43, 2015, pp. 175–223.
- Frick, Johann, "What We Owe to Hypocrites," *Philosophy & Public Affairs* 44, 2016, pp. 223–265.
- Kamm, Francis, "Owing, Justifying, and Rejecting", *Mind* 111, No. 442 (2002), pp. 323–354.
- Kamm, Francis, *Intricate Ethics*, Oxford: Oxford University Press 2007.
- Munoz-Dardé, Veronique, "The Distribution of Numbers and the Comprehensiveness of Reasons," *Proceedings of the Aristotelian Society* 105 (2005), pp. 207–233.
- Murphy, Jeffrie G., "Marxism and Retribution," *Philosophy & Public Affairs* 2, 1973, pp. 217–243.
- Murphy, Liam, "The Demands of Beneficence," *Philosophy & Public Affairs* 22, 1993, pp. 267–292.
- Nagel, Thomas, *Mortal Questions*, Cambridge: Cambridge University Press 1979.
- Nagel, Thomas, *Equality and Partiality*, Oxford: Oxford University Press, 1991.
- Parfit, Derek, *On What Matters*, Volumes One and Two, Oxford: Oxford University Press, 2011.
- Scanlon, Thomas M. "Rights, Goals and Fairness." In Stuart Hampshire, ed., *Public and Private Morality*, Cambridge: Cambridge University Press, 1978, pp. 93–125
- Scanlon, Thomas M., "Contractualism and Utilitarianism" in *Utilitarianism and Beyond*. Edited by Amartya Sen and Bernard Williams. Cambridge: Cambridge University Press, 1982, pp. 103–128.
- Scanlon, Thomas M., *What We Owe to Each Other*. Cambridge, MA: Harvard University Press, 1998.

- Scanlon, Thomas M., *Moral Dimensions: Permissibility, Meaning, Blame*, Cambridge, MA: Harvard University Press, 2008.
- Scanlon, Thomas M., "Rights and Interests," in *Arguments for a Better World: Essays in Honor of Amartya Sen: Vol. I: Ethics, Welfare, and Measurement*, edited by Kaushik Basu and Ravi Kanbur (New York: Oxford University Press, 2009).
- Scanlon, Thomas, M., *Being Realistic about Reasons*, Oxford: Oxford University Press, 2014.
- Scanlon, Thomas M., *Why Does Inequality Matter?*, Oxford: Oxford University Press, 2018.
- Scheffler, Samuel, "Rawls and Utilitarianism" in Samuel Freeman, ed., *The Cambridge Companion to Rawls*, Cambridge: Cambridge University Press, 2003, pp. 426–459.
- Singer, Peter, "Famine, Affluence, and Morality," *Philosophy & Public Affairs* 1, 1972, pp. 229–243.
- Williams, Bernard, "Persons, Character, and Morality" in *Moral Luck*, Cambridge: Cambridge University Press, 1981, pp. 1–19

Rainer Forst

Justification Fundamentalism: A Discourse-Theoretical Interpretation of Scanlon's Contractualism

In my brief remarks, I develop a particular, discourse-theoretical reading of Thomas Scanlon's theory of contractualism with respect to moral justification and then suggest a conception of political justification as following from it. The latter provides the core for a conception of political and social justice that I find close to Scanlon's own views. So what I present here is less of a critique than it is an invitation. But Scanlon may not enthusiastically welcome that invitation, as it makes him appear more of a *justification fundamentalist* rather than a reasons fundamentalist, and in addition, his theory moves closer to a certain version of Kantianism. Which is, as I think, where it belongs.

1 Moral Justification

I would like to start my reflections with a quote from an interview of Scanlon with Alex Voorhoeve, where he says that the very structure of the contractualist test of reasonable rejectability is "already a moral principle" – such "that everybody counts – that we should be able to justify our norms to everyone. The reason is that we are all reasonable creatures, to whom it makes sense to want to justify our norms to everyone."¹ I agree with this, and also with Scanlon when he continues by saying "I find it hard to accept a notion of wrongness that is prior to or independent of justification [...]."² He furthermore suggests the contractualist test of the generalizability of moral norms as an alternative to Kant's test of contradiction in conception or in will and emphasizes the relational, intersubjective character of this test, which involves regarding every person as an equal au-

Acknowledgment: Thanks to the participants of the Lauener Symposium on Thomas Scanlon's work in Bern in September 2016 for a productive discussion, especially to Tim Scanlon himself for a kind, spontaneous reply. I also owe thanks to Felmon Davis and Isaac Taylor for their helpful remarks on an earlier version of this text.

1 Scanlon 2009b, p. 182.

2 Ibid., p. 185.

<https://doi.org/10.1515/9783110733679-006>

thor of moral claims and justifications, i.e., as in that sense an autonomous member of a universal moral community of respect and mutual justification.

In light of that, I would call his approach “justification fundamentalism” rather than “reasons fundamentalism,” the term that Scanlon chooses in *Being Realistic About Reasons*.³ I think he is a justification fundamentalist in not accepting any notion of wrongness or rightness prior to a certain form and procedure of intersubjective justification; thus he takes, as he says in *What We Owe to Each Other*, “the idea of justifiability to others [...] to be basic”⁴ since it is these others that one primarily owes appropriate, mutually justifiable moral reasons to. So they are equal justificatory authorities when it comes to evaluating good moral reasons. Yet at the same time, the demand of *reasonable* rejectability or justification implies another authority to be in play here, namely that of reason as a faculty that we exercise *together* when we determine which reasons are good ones and which are not.

According to Scanlon’s contractualism, moral persons have a basic moral claim to be respected as equal justificatory authorities, and this kind of respect requires such persons to follow the reason-based contractualist test of justification to determine moral rightness or wrongness. There are two main components of this view – to regard persons as equal moral authorities, and to regard a certain procedure of justification (or of reasonable rejectability) as required to exercise that authority. Both are components, I think, of what I call “justification fundamentalism”: First, the main point – and ground – of morality (in the narrow meaning of what we owe to each other) is to respect others as moral equals to whom I owe a proper justification for my moral actions; and second, the proper justification is to be found by way of a certain procedure of intersubjective (or: discursive), reasonable justification.

These are very Kantian formulations, but I think they are in line with Scanlon’s views – at least as long as we don’t imply some kind of transcendental rational agency in a Korsgaardian fashion to explain the ground of the moral ought here.⁵ For I agree with Scanlon that it is not the constitution of our rational agency that grounds morality but a reflection on us as equal justificatory authorities who have a basic moral responsibility to exercise the faculty of justification in a respectful way, i.e., to respect every other human being as a moral equal. That kind of respect – or the “idea of justifiability to others”⁶ – is the “motivational basis” and normative core of contractualism, and if we understand Kant’s

³ Scanlon 2014, Lecture 1.

⁴ Scanlon 1998, p. 5.

⁵ Korsgaard 2009. See Scanlon’s critique of her approach in Scanlon 2011.

⁶ Scanlon 2011, p. 139.

moral philosophy not primarily as a philosophy of rational action but as one of morally *responsible* – i. e., justifiable – and in that sense *autonomous* and reason-guided action, the same idea lies at the core of Kant’s project.⁷

The basic kind of contractualist respect I would therefore not call, as Scanlon does, “avowedly heteronomous,”⁸ though it is based on an ideal of certain social relations with others. Yet these relations are always mediated by reasonable justification, and thus they are themselves based on the proper exercise of practical reason-seeking principles that “govern” our social relations morally. Here is a Kantian-sounding quote by Scanlon expressing what I have in mind:

“[R]especting the value of human (rational) life requires us to treat rational creatures only in ways that would be allowed by principles that they could not reasonably reject insofar as they, too, were seeking principles of mutual governance which other rational creatures could not reasonably reject.”⁹

That social version of the “kingdom of ends,” as I would call it, expresses not just a social ideal of mutual recognition, but more than that an ideal of mutual justification guided by principles relating to the use of reason – which is what Kantians call an “ideal of practical reason”: practical as it applies to norms of action, reason as these norms have to go through a process of mutual justifiability guided by criteria of reason, ideal as it spells out an ideal way of relating to oneself and others as justificatory equals. I will come back to this point below.

I stress *justifiability*, as like Scanlon I do not think that an actual or some imagined hypothetical consent can deliver the moral justification we are seeking but that a judgment based on the best reasons is needed. This is why Scanlon thinks reasons, rather than imperatives of rationality, are fundamental. But what makes these reasons justifiable, I take it, is the proper exercise of practical reason respecting every moral person as an equal moral authority deserving proper reasons and (in the standard case) being able to respond with reasons. So the alternative between a view based on *reasons* and one based on *reason* seems to disappear: reason operates through reasons, but it is not guided by them; rather, reason is the faculty used to test reasons, question them, and evaluate them in the light of *principles* suitable to determine justifiable reasons in the various domains of justification – which is how I interpret the notion of “domains.”¹⁰ Otherwise, such domains would have no unity, and we would not know how to use reason to distinguish better from worse reasons within such

⁷ See Forst 2012, chs. 1 and 2.

⁸ Scanlon 1998, p. 6.

⁹ *Ibid.*, p. 106.

¹⁰ Scanlon 2014, pp. 19–26.

contexts. Reasoning guided by principles of reasoning determines good reasons, not the other way around. Thus, if the contractualist test is one of reasonable justification, it expresses principles and criteria of practical reason which *constitute* good reasons. And that is why reasons fundamentalism has to give way to justification fundamentalism; in other words, if the contractualist test tests something, it is reasons of rightness or wrongness – and thus it cannot be guided by such reasons. A justification fundamentalist does indeed believe that reasons can be explained in terms of something more fundamental: principles of reasoning, i. e., of justification.

Why would Scanlon disagree? He defines a reason as “a consideration that counts in favor of”¹¹ something, and the faculty for arriving at such considerations is that of reflective judgment as guided by “standards” of rationality.¹² Similarly, he describes the process of arriving at truths about reasons as a process of “thinking carefully about what seem to us to be reasons, considering what general principles about reasons would explain them, what implications these would have, and considering the plausibility of the implications of these principles.”¹³ Across various domains, some general principles like those of logic apply, but there are domain-specific principles of judging reasons – yet all of these have to be principles of reasonable justification. They *constitute* good reasons.

It might be helpful at this point to briefly mention my particular rendering of what mutual justifiability means. I suggest to interpret “reasonable rejectability” as “reciprocal and general rejectability,” assuming, as contractualism does, that the realm of what people can reciprocally and generally accept may be broader than what they can reciprocally and generally reject – and that the latter realm thus circumscribes the narrow realm of morality. The first criterion implies that one must not make normative claims one denies to others (reciprocity of claims) or impose one’s own interests or particular views on others who could reject them as partial (reciprocity of reasons), while the second implies that one must not exclude others from the moral justification community (generality). To reject a norm or a principle by showing that it violates reciprocity or generality is a *reasonable rejection* (while one could also accept it as an act of supererogation). Thus it is not consensus that determines moral rightness but the quality

11 Scanlon 1998, p. 17.

12 “[I]n order for judgments about reasons to be taken to be about some subject matter independent of us in the sense required for it to be possible for us to be mistaken about them, what is necessary is for there to be standards for arriving at conclusions about reasons.” *Ibid.*, p. 63.

13 Scanlon 2014, p. 102.

of the reasons given – yet that quality is determined by the discursive use of the two criteria, in practice as well as in one’s judgement.¹⁴ These two criteria constitute non-rejectable reasons.

It would be a mistake to read the word “constitute” in metaphysical terms and to think that this is the terrain at which the debate between realism and constructivism takes place.¹⁵ For a justification fundamentalist can remain agnostic about the question of whether contractualist justification discovers “real” moral truths or produces them in any way that would be inconsistent with such a form of realism.¹⁶ For the view that a justification procedure generates good reasons by way of certain criteria of justification need not take a stance on that issue. All that matters is that the normative criteria of reason are decisive for us to be able epistemically to distinguish good from bad reasons. Whether these reasons are “created” or “discovered” in a metaphysical or ontological sense is a question that does not matter from a moral or from an epistemic or normative point of view. I call such a version of constructivism *pragmatic* rather than metaphysical.¹⁷

At least with respect to the domain of morality narrowly understood, pragmatic constructivism suggests itself as the proper view. And indeed, when it comes to questions of social and political justice and to individual morality, Scanlon thinks a constructivist account of objectivity – i.e., judgment- and choice-independence as he defines them¹⁸ – has “considerable plausibility”¹⁹, as judgments about rightness or wrongness depend on certain procedures of justification properly performed.

The reason why Scanlon does not accept constructivism all the way down is that he believes that it cannot deliver a general account of reasons for action. A Kantian view, for example, cannot explain why certain ends should be valuable only if I have chosen them autonomously according to the Categorical Imperative.²⁰ This objection is correct if, for example, we think of many values and ideals that persons care deeply about apart from the (narrow) moral realm, such

14 I take this to be in line with Scanlon when he writes: “When we think of those to whom justification is owed, we naturally think first of the specific individuals who are affected by specific actions. But when we are deciding whether a given principle is one that could reasonably be rejected we must take a broader and more abstract perspective.” Scanlon 1998, p. 202.

15 See, for example, Enoch 2009.

16 Scanlon 2014, Lecture 2.

17 Forst 2012, p. 50.

18 Scanlon 2014, pp. 93f.

19 *Ibid.*, p. 98.

20 *Ibid.*, pp. 98–100.

as religious values or particular conceptions of the good. Such (in my terminology)²¹ “ethical” ends are seen as valuable for particular reasons, but they are not made valuable through adopting them autonomously. But Scanlon does not just criticize such an overstretch of Kantian accounts of morality into the realm of the ethical; rather, with respect to morality narrowly understood, his view is that while all the moral reasons that go through the contractualist test can be seen as constructed by that test, the very reason to engage in the test in the first place cannot be constructed in that way. There has to be an independently valid reason to accept the contractualist moral principle which cannot be reduced to reasons of rational agency.²²

This problem is notorious for a constructivist view, as no constructivist theory can deliver the basic reason to be moral (and to use a certain justification procedure expressing equal respect) in a purely constructivist fashion. A constructivist view (of justice or morality) entails two different kinds of normative arguments, or two kinds of normativity: *first*, the normativity of the principles and ideas of practical reason, as Rawls phrases it in his version of constructivism²³ – or, on Scanlon’s account, the principle of reasonable rejectability and the substantive claim that every moral person owes following this principle to every other moral person, or in my view, the principle of reciprocal and general justification and the moral notion of free and equal persons as equal normative authorities with a right to justification. Then there is, *second*, the normativity of the norms (or “laws”) generated by the constructivist, contractualist or discursive procedure, be it the categorical imperative, the contractualist test or a notion of free and equal discourse. In a Kantian view, it is essential that practical reason (on my understanding, justificatory reason) provides the basis for the principles and ideas used, where practical reason is understood as a rational and, at the same time, *moral* capacity, that is, not just as a matter of knowing *how* to justify norms, but also of knowing *that* one is under a *duty* to do so.

One cannot infer, however, that since the second kind of normativity is based on the first, which is not constructivist in the same way as the second, that constructivism fails. All one needs to say is that the moral ground of the constructivist procedure is a substantive moral ground, but not one that is separate from a reflection on us as reasonable justificatory beings who have a moral duty to respect others as justificatory equals by using the faculty of justification

²¹ Forst 2012, ch. 3.

²² Scanlon 2014, p. 100.

²³ See Rawls 1980, and Rawls 1993, ch. 3. For a Kantian interpretation of Rawls’s constructivism, see Forst 2017b.

in the right way. Thus while I think, as I tried to explain above, that the notion of “reasons fundamentalism” does not do justice to the priority of justification and the criteria of justification that the contractualist test implies, I also think, as I briefly suggested, that the “ideal of justifiability to others”²⁴ contractualism is based on indeed is an *ideal of reason* insofar as its ideal of moral community is one of a community ruled by reasonable norms of “mutual governance” – and thus also of moral autonomy in a Kantian sense of the term. And the idea of moral persons as having a right to and a duty of justification seems to be a good way of expressing the basic imperative of moral respect contractualism emphasizes: “Human beings are capable of assessing reasons and justifications, and proper respect for their distinctive value involves treating them only in ways that they could, by proper exercise of this capacity, recognize as justifiable.”²⁵

Thus I think that Scanlon is right to say that a constructivist account of morality is based on a substantive answer to what grounds the moral duty of justification, but I think that it is not correct to say that this is independent from our faculty of practical reason. For the moral duty of justification is a moral duty that reflects our status as moral authorities in the space of justifications who are equal to all other humans, and it further entails that the way to exercise that duty is through a proper procedure of justification respecting us and others as equals. This is still justification fundamentalism, but in a different key: the basic contractualist ideal of justifiability to others is not just one normative ideal among others that we have good reason to accept. Unlike other ethical or religious ideals, it is an ideal of reason, as it expresses what it means to be (noumenally, in traditional terms) part of and to help create in practice a “kingdom of ends,” or a moral community of mutual respect and *reasonable* justifiability. Reason is a faculty that binds us in this sense; it contains a categorical imperative to use it in the morally right way.

2 Political Justification

According to my discourse-theoretical, Kantian interpretation of contractualism, human beings have a basic moral right to justification which is so to speak a veto right against reciprocally and generally non-justifiable normative claims. To respect that right means to respect others as ends in themselves, to use Kant’s

²⁴ Scanlon 1998, p. 155.

²⁵ *Ibid.*, p. 169.

term and give it a (hopefully) more precise meaning. I think this accords with Scanlon's way of understanding and respecting the moral value of human persons as the individuals they are and as moral equals at the same time.

Like Scanlon and Rawls, I believe that our basic moral standing as equals travels from the moral realm to that of political and social justice, i.e., to that of our basic standing as members of a legal and political normative order who justifiably demand legal, political and social justice. From that perspective, I see many parallels with Scanlon's arguments about human rights, toleration and equality. But in the realm of justice, I think the part of the contractualist formula that requires norms to be a "basis for informed, unforced general agreement"²⁶ becomes more relevant for a conception of reciprocal and general justification as a social *practice* of justice – the practice of a society whose members have a basic standing as justificatory equals such that they are not subjected to a normative order (of institutions and laws) that cannot be justified to them in a practice we call *democracy* (in the ideal sense of the term). Here the ideal of justifiability towards others has a concrete, institutional meaning calling, among other things, for a legally, politically and socially secure status of non-domination, i.e., of not being subject to a normative order without proper justification and, especially, of not being subjected to a normative order without proper procedures and institutions of justification in place. To have such a status of non-domination or, positively speaking, of being a legally, politically and socially equal authority of justification, is the basic demand of justice, and thus it is the task of what I call *fundamental justice* to secure this status – with the help of basic rights, democratic procedures and powers and social relations which enable all members to use their rights and avoid being forced to accept social structures and relations of domination.²⁷

On my discourse-theoretical understanding (which differs substantially from neo-republican versions)²⁸ domination does not primarily mean being denied equal status in the sense of no longer enjoying personal freedom of choice protected from arbitrary interference; rather, it means in a more basic sense being disrespected in one's basic claim to be a free and equal normative authority within the order one is subject to. This implies the basic right to co-determine the structure of that society. Basic rights are not just rights to be protected in one's status as a legally, politically and socially non-dominated person; they are, in a reflexive sense, also rights to determine the rights and duties that define

²⁶ Ibid., p. 153.

²⁷ See Forst 2012, part II, and Forst 2013a, esp. ch. 1.

²⁸ See Forst 2017a.

this status.²⁹ Thus, the authority to determine your rights must reside in a discursive procedure of reciprocal and general justification in which all participants are justificatory equals. This is what justification fundamentalism means in the political realm.

I think this is not just a notion of justice close to Scanlon's views; it is also the best extension of moral contractualism to the realm of social and political life. Let me explain. In my view, following Rawls's and close to Scanlon's, I believe, the concept of justice possesses a core meaning whose essential contrasting concept is that of *arbitrariness*.³⁰ Arbitrariness can assume the form of direct arbitrary rule over others by individuals or by a part of the community (for example, by a class), or it can involve the acceptance of social contingencies that lead to asymmetrical social positions and relations of domination as if they were unalterable or beyond justification, even though they are nothing of the sort. Arbitrary rule is the rule of some people over others without legitimate reason, and where social struggles are conducted against injustice, they are directed against forms of domination of this kind. The underlying impulse that opposes injustice is not primarily that of wanting something or more of something, but of no longer wanting to be dominated, harassed or overruled in one's claim and basic right to justification. Herein resides the profoundly *political* essence of justice that a purely goods- or recipient-oriented view fails to grasp: justice concerns *who determines who receives what* and not only or primarily who should receive what.

That is also why a well-meaning authoritarian regime that provides its citizens with a decent package of basic goods of housing, health and income does a lot to improve its citizens' lives – but serves no justice. Those who dissect justice into social or distributive justice as a matter of the provision of goods, whether sufficientarian or egalitarian, on the one hand, and political justice or legitimacy, on the other, would disagree. But I think that is a mistake: if there were an ideal Cohenite “egalitarian distributor”³¹ who had figured out the right metric of justice and instantiated it by authoritarian rule, justice would not have been done, neither politically nor socially. People would just be better off but still dominated and not treated as autonomous subjects of justice, because

²⁹ See Forst 2016.

³⁰ Rawls 1999, p. 5: “Those who hold different conceptions of justice can, then, still agree that institutions are just when no arbitrary distinctions are made between persons in the assigning of basic rights and duties and when the rules determine a proper balance between competing claims to the advantages of social life.”

³¹ Cohen 2011, p. 61.

that autonomy implies that they determine the metric of justice (in a procedure on the basis of fundamental justice).

It is on this basis that we can construct a comprehensive theory of political and social justice, though in the present context I can only hint at what such a theory would entail. First, we must distinguish between fundamental and full justice. Whereas the task of fundamental justice is to construct a basic structure of justification, the task of full justice is to construct a fully justified basic structure. In order to pursue the latter, the former is necessary, that is, a “putting-into-effect” of justification through constructive, discursive democratic procedures in which justificatory power is distributed as evenly as possible among the citizens. Fundamental justice guarantees all citizens an effective status as justificatory equals. This would still fall short of guaranteeing “informed, unforced general agreement,” as the political contractualist would wish for, but it would still aim at establishing a basic structure in which asymmetries of justificatory power that keep reproducing themselves can be overcome.

I think that such a conception of fundamental justice is much in accord with Scanlonian contractualism. Let me first point out the parallels and then also reflect on the reasons why Scanlon still might not follow me on this path. To start with, a conception of basic rights as expressing respect for others as justificatory equals in legal, political and social life is in line with contractualism going political, but it differs from Scanlon’s conception of rights which is closer to an interest theory, arguing that rights protect (or promote) central valuable interests of human beings.³² According to the notion of basic rights I have in mind, these rights are status- rather than interest-based, i.e. they are justified by a consideration of a secure status of non-domination and an inquiry into the relevant contexts of such non-domination and the possibilities of being a justificatory subject that is able to co-determine the normative order he or she is subject to. Sure, we can say that basic interests in non-domination are protected by this, but this is not our main argument, for the status of being a justificatory equal carries the justificatory weight, not the idea of certain interests, whether real, ideal or objective (and Scanlon taught us a lot about the problems in defining such interests). Such a conception of basic rights explains best, I think, why human rights are a “neutral concern,”³³ as Scanlon argues, and how a core of them remains firm while interpretations of that core may differ among societies and cultural contexts.

³² Scanlon 1978.

³³ Scanlon 1979.

That political justice demands a secure status of being a justificatory equal also fits with Scanlon's important reflections on toleration. For he rightly argues that toleration "involves 'accepting as equals' those who differ from us"³⁴ and who are nevertheless equally entitled to co-determine social life. I may reject their views but have to accept other persons as equal members, as long as they do not deny my standing of an equal or that of others. So my ethical objection to their, say, religious views is reasonable but need not be shared by all such that it can ground a moral rejection; if it grounds such a rejection, toleration is out of place.³⁵ Tolerated beliefs and practices are reasonable to hold and reasonable to reject as a general norm (a case of what Rawls called "reasonable disagreement"); and intolerable beliefs and practices are those that violate reciprocally and generally non-rejectable norms.

Finally, I think that the notion of equal respect and of standing as non-dominated justificatory equals in legal, political and social contexts is very close to the important reasons that Scanlon isolates for equality, especially the three reasons he mentions of preventing the stigma of inferiority, avoiding domination and preserving procedural fairness.³⁶ My justification fundamentalist reconstruction of contractualism can show, I think, that these three powerful considerations have a common source, i.e., our status as justificatory equals with a right to justification. I think that this notion of moral equality is the core of any justified claim to legal, political and social equality, and that greater equality of outcomes or strict equality of social standing and resources is only justified if it can be justified on these grounds of justificatory equality. This establishes a normative order among the equality considerations Scanlon distinguishes, something missing in his account (so far).

All of these reflections about rights, about toleration and about equality presuppose, in my view, the notion of the authority of persons to co-determine the structure of their normative order as democratic justificatory equals. And that is no add-on or separate consideration of political justice; rather, it lies at the heart of a just regime. So I wonder why democracy as the practice of political justification plays no bigger role in Scanlon's political philosophy. I don't want to speculate about this, but it might have something to do with certain aspects of his thought which are less justification fundamentalist and more reasons fundamentalist. It may be a remnant of a value-based approach to political philosophy, where it is not through justificatory discourse that autonomous members of a

³⁴ Scanlon 2003a, p. 190.

³⁵ See Forst 2013b.

³⁶ Scanlon 2003b. See also Scanlon 2018.

normative order determine their society, but where they are called upon to order and arrange their society according to an order of values they can find through impersonal reflection. That, however, is an approach that I find hard to square with the interpretation of contractualism I laid out.

References

- Cohen, Gerald A., "Afterword to Chapters One and Two," in Cohen, *On the Currency of Egalitarian Justice, and Other Essays in Political Philosophy*, Princeton: Princeton University Press, 2011, pp. 61–72.
- Enoch, David, "Can There Be a Global, Interesting, Coherent Constructivism About Practical Reason?," in: *Philosophical Explorations* 12 (3), 2009, pp. 319–339.
- Forst, Rainer, *The Right to Justification: Elements of a Constructivist Theory of Justice*, New York: Columbia University Press, 2012.
- Forst, Rainer, *Justification and Critique: Toward a Critical Theory of Politics*, Cambridge: Polity Press, 2013a.
- Forst, Rainer, *Toleration in Conflict: Past and Present*, Cambridge: Cambridge University Press, 2013b.
- Forst, Rainer, "The Justification of Basic Rights: A Discourse-Theoretical Approach," in: *The Netherlands Journal of Legal Philosophy* 45 (3), 2016, pp. 7–28.
- Forst, Rainer, *Normativity and Power: Analysing Social Orders of Justification*, Oxford: Oxford University Press, 2017a.
- Forst, Rainer, "Political Liberalism: A Kantian View," in: *Ethics* 128 (1), 2017b, pp. 123–144.
- Korsgaard, Christine, *Self-Constitution: Agency, Identity, and Integrity*, Cambridge: Cambridge University Press, 2009.
- Rawls, John, "Kantian Constructivism in Moral Theory," in: *The Journal of Philosophy* 77 (9), 1980, pp. 515–572.
- Rawls, John, *Political Liberalism*, New York: Columbia University Press, 1993.
- Rawls, John, *A Theory of Justice*, rev. ed., Cambridge: Harvard University Press, 1999.
- Scanlon, Thomas M., "Rights, Goals and Fairness," in: Stuart Hampshire, ed., *Public and Private Morality*, Cambridge: Cambridge University Press, 1978, pp. 93–125.
- Scanlon, Thomas M., "Human Rights as a Neutral Concern," in: Peter Brown and Douglas Maclean, eds., *Human Rights and U.S. Foreign Policy*, Lexington: Lexington Books, 1979, pp. 83–92.
- Scanlon, Thomas M., "Contractualism and Utilitarianism," in: Amartya Sen and Bernard Williams, eds., *Utilitarianism and Beyond*, Cambridge: Cambridge University Press, 1982, pp. 103–128.
- Scanlon, Thomas M., "The Difficulty of Tolerance," in: Scanlon, *The Difficulty of Tolerance: Essays in Political Philosophy*, Cambridge: Cambridge University Press, 2003a, pp. 187–201.
- Scanlon, Thomas M., "The Diversity of Objections to Inequality," in: Scanlon, *The Difficulty of Tolerance: Essays in Political Philosophy*, Cambridge: Cambridge University Press, 2003b, pp. 202–218.

- Scanlon, Thomas M., *What We Owe to Each Other*, Cambridge: Harvard University Press, 1998.
- Scanlon, Thomas M., *Moral Dimensions: Permissibility, Meaning, Blame*, Cambridge: Harvard University Press, 2008.
- Scanlon, Thomas M., "Rights and Interests," in: Kaushik Basu and Ravi Kanbur, eds., *Arguments for a Better World: Essays in Honor of Amartya Sen*, Vol. I: *Ethics, Welfare, and Measurement*, New York: Oxford University Press, 2009a, pp. 68–79.
- Scanlon, Thomas M., "The Kingdom of Ends on the Cheap," Interview with Alex Voorhoeve, in: Voorhoeve, *Conversations in Ethics*, Oxford: Oxford University Press, 2009b, pp. 179–194.
- Scanlon, Thomas M., "How I am Not a Kantian," in: Derek Parfit, *On What Matters*, Vol. 2, Oxford: Oxford University Press, 2011, pp. 116–139.
- Scanlon, Thomas M., *Being Realistic about Reasons*, Oxford: Oxford University Press, 2014.
- Scanlon, Thomas M., *Why Does Inequality Matter?*, Oxford: Oxford University Press, 2018.

Susanne Mantel

On How to Explain Rational Motivation: Where Internalism and Externalism Meet

Normative cognitivism – the view that normative attitudes are beliefs – is often confronted with the challenge that it seems to be incompatible with explaining how normative attitudes motivate agents. In his book *Being Realistic about Reasons* Tim Scanlon provides a rational explanation of why agents act in accordance with normative beliefs. I discuss an interpretation of this explanation according to which it relies on an enkratic requirement, and point out that this version of the explanation faces two problems: It is unacceptable to instrumentalists and does not meet the entire scope of the motivational challenge, which comprises actions performed for optional reasons. I suggest an alternative version of the explanation which relies on a dispositional structure and is detached from rationalist theory. Thereby, the cognitivist's answer to the motivational challenge steers clear of disputes between rationalist internalists and externalists and accommodates actions performed for optional reasons. The motivational challenge can thus be answered in its entire scope and the defense of normative cognitivism bears fewer theoretical commitments.

In his book *Being Realistic about Reasons*, Tim Scanlon advocates normative cognitivism – the view that normative judgments, e.g., judgments about normative reasons, are *beliefs*.¹ Scanlon defends normative cognitivism against metaphysical, epistemological, and motivational objections. I will focus on the third challenge, the motivational one. Scanlon's answer to this challenge points in the right direction, I think. However, it seems to rely on a view about the requirements of rationality which is controversial and which some cognitivists might want to avoid. More importantly, it does not seem to cover the motivational challenge in its entire scope, since his solution does not seem to apply to judgments about *optional* reasons, for instance. Therefore, I provide a dispositional interpretation of Scanlon's answer which, I hope, allows me to do without any contro-

Acknowledgment: I am very grateful to the participants and the organizers of the Lauener Symposium 2016 in Bern, as well as to the participants of the Doctoral Colloquium of Practical Philosophy at Saarland University in 2016.

¹ More precisely, normative cognitivism states that normative judgments are mental states with a *mind to world direction of fit*, or *representational* beliefs.

<https://doi.org/10.1515/9783110733679-007>

versial assumption about the requirements of rationality. My proposal is meant to be compatible with rationalist motivational internalism as well as with externalism. Furthermore, I suggest some amplifications such that we may explain the motivational impact of the full variety of normative judgments – even of judgments about optional reasons. To sum up, I hope to improve the cognitivist’s answer to the motivational challenge by reducing its burdens and broadening its scope.

1 The Motivational Challenge

The motivational challenge can be put as the claim that a cognitivist account of judgments about reasons must explain how these judgments motivate agents to act accordingly.² As van Roojen puts it, judgments about reasons “are supposed to be practical in a way that judgements about theoretical physics or agriculture or kitchen appliances are not” (van Roojen, 2015, 54). One way to mark the difference is that judgments about physics, agriculture or kitchen appliances motivate an agent only if he or she happens to have an independent desire – to test a physical theory, to reap a lot of grain, to cook with less effort, or the like. By contrast, judgments about reasons seem to motivate many people no matter what they happen to desire. How can these judgments do so, given that they are beliefs?³

Cognitivism seems to be in a less comfortable position than expressivism when it comes to explaining the practicality of normative judgments. If judgments about normative reasons are not beliefs, but are desires or plans, then it is easy to explain how the agent is motivated to act in accordance with them – after all, these judgments constitute a motivational state, e.g., a desire,

2 The motivational force of normative judgments plays an important role in *acting for* normative reasons, since we often act for reasons with the help of our *normative* representations of them. That is, it is often our belief *that p is a normative reason to do a* which makes us do a. Elsewhere, I have called this phenomenon “acting for normative reasons in the *de dicto* way”. However, I believe that we may also act for normative reasons without representing them as normative, i.e., we may act for the normative reason that p if it is merely our belief *that p*, not *that p is a normative reason to do a*, which moves us to act in the right way. This can be referred to as “acting for a normative reason in the *de re* way” (compare Mantel 2018, chap. 4). The motivational challenge is one that besets the *de dicto* way, since it is about judgments with an explicitly normative content.

3 An early version of the motivational challenge is found in Mackie 1977, 40, and discussed, for example, in Olson 2014, chap. 5.3.

that the agent already has, in virtue of so judging, and this motivational state may explain his behavior.

I will focus on this problem with respect to judgments about *conclusive* reasons until I consider judgments about inconclusive reasons in section 4. Until then, when I speak about judgments about reasons without specifying them further, I meant to refer to judgments about conclusive reasons.

Scanlon distinguishes between two versions of the challenge for the normative cognitivist. Take the normative judgment *that the injury of this person is a conclusive reason to help*. The normative cognitivist claims that such a judgment is a belief (rather than a conative attitude like a desire to help the injured).⁴ This claim raises both the question of how this belief *causes* action and of how it *rationalizes* and thereby *explains* the action.⁵ Scanlon argues that the causal version presents no deep philosophical problem, and that the explanatory version presents a philosophical problem which can be solved. I agree with him, but I will suggest some modifications of his solution.

1.1 The Causal Aspect of the Motivational Challenge

If beliefs about reasons *cause* action, Scanlon suggests, this will be due to some neural mechanism and does not pose any specifically philosophical problems for cognitivism. The causal aspect of the challenge, that is, is not so much a challenge for philosophy but rather for neuroscience (Scanlon 2014, 54–55).

It might be objected that, contrary to this view, there are specifically philosophical problems pertaining this causal relation. In particular, one might think that the Humean Theory of Motivation established that beliefs cannot cause motivation. This objection, however, seems to be grounded in a misunderstanding of the Humean Theory of Motivation, which merely says that beliefs do neither *constitute* nor *entail* motivation, or that beliefs and desires are “distinct existences” (cf. Smith 1994, 7).⁶ The expressivist may claim that it is part of a desire’s *functional role* to cause action in suitable circumstances, but not so for beliefs.

⁴ By the term ‘judgment’ I refer to a mental state, not an utterance.

⁵ One can also understand the question how normative judgments motivate as the question how they cause and explain *motivational states* like intentions. For my purposes, the difference is not important and I suppose that they explain action, when they do, via explaining motivational states.

⁶ In giving this minimalist interpretation of the Humean theory, I reject the quite common view that the Humean theory precludes that beliefs motivate without independent desires, cf., e.g., Parfit 1997.

Desires are, on this view, complex dispositions to deliberate and to be motivated in certain instrumental ways (although the agent may not intend or act in accordance with a desire if stronger desires overrule it).⁷ In this sense, desires are motivational by their very nature, whereas beliefs are at best contingently connected to motivation.

Nevertheless, Scanlon is right that causal mechanisms may connect a belief to a motivational state like a desire or intention, and thereby ultimately to action, even if this is not essential to beliefs.⁸ Contingent causal connections are not metaphysically strange and thus do not as such challenge cognitivism. This is of course not a special observation about the mental but holds for causes in the non-mental sphere as well. Suppose exposure to sunshine *essentially* causes heat (if it has any essential effect) and not fissures or happiness. Still, there is nothing strange about the fact that it may also, contingently, cause an object to have fissures or a person to be happy – while it may leave many objects intact and annoy someone who forgot to use sun blocker.

Of course, it is to be expected that contingent connections do not hold universally – contingent causal chains may be absent with respect to some objects and people. This suggests that people sometimes fail to be motivated to do what they believe they have reason to do. The position that they are *necessarily* motivated by their normative judgments, often called strong internalism (e.g., Brink 1997, 7), would turn out to be false.⁹ This consequence does not yield a *reductio* but, to the contrary, it supports cognitivism, I believe, since it is rather plausible to interpret at least some agents as acting against their normative judgments.¹⁰ It seems that this form of akrasia is not uncommon, and it is especially vivid in

⁷ This functionalist view of desires is the expressivists', not Scanlon's. If the functionalist view is false and desires are not dispositions to act the expressivist must look for another story of how desires cause actions.

⁸ Cf. Parfit 1997, 105 for the view that normative beliefs can cause motivation without the help of independent desires. He attributes the thought to Nagel, although Nagel stresses that normative judgments are not "merely classificatory" (1970, 109), which can be interpreted as stressing that they are not strictly speaking beliefs. Cf. Broome 1997, 142 for the view that the explanation relies on the "natural" disposition to do what you believe you ought to do.

⁹ Some take this strong internalism to be an *a priori* claim (cf. Björnsson, Strandberg, Olinder, Eriksson, and Björklund 2015). My intuitions, as Scanlon's, allow akrasia and thus rather support, at most, a defeasible internalism (see section 3).

¹⁰ Expressivism does not merely have difficulty explaining why akratic people fail to be motivated in accordance with their judgments about what they themselves have conclusive reasons to do, but also why even rational agents do not seem to be motivated in any particular way by their judgments about what *other* people have conclusive reasons to do (Thomson 2006, Gregory 2017).

depressive agents (Stocker 1979, 744). Further examples cite addiction (Frankfurt 1971) and emotional strain (Hempel 1961, 297) as factors that impair the connection between normative judgments and action. Arguably, especially depressive agents seem not even to experience a weak rise in motivation due to normative judgments, since they may be unmoved even if they have no contrary desires. If this is right, normative judgments not only fail to motivate them sufficiently, but lack any motivational impact. Being unmoved by one's judgment about conclusive reasons is *possible*, even if, we hope, it is not too common. If the connection between normative judgments and action, as well as motivation, can break down, it is all the more plausible to conceive of it as a contingent causal connection. While my main aim is to show that cognitivism is not inferior to externalism with respect to answering the motivational challenge, these thoughts might seem to indicate that cognitivism is even superior to expressivism in explaining why many people are not as ideal and enkratic as one might wish they were.¹¹

1.2 The Rational-Explanatory Aspect of the Motivational Challenge

According to Scanlon, the alleged problem for cognitivism is primarily explanatory. It lies in the question how normative judgments *rationalize* and thereby *explain* action. To accommodate rationalization, the explanation in question must be located at the mental level, not the neuronal one.

On a Davidsonian interpretation, a rationalizing explanation lets us “see something the agent saw, or thought he saw, in his action – some feature, consequence, or aspect of the action the agent wanted, desired, prized, held dear, thought dutiful, beneficial, obligatory, or agreeable” (Davidson 1963, 3).

As it seems to me, Scanlon acknowledges this Davidsonian sense of the term “rationalization” (Scanlon 2014, 53) but then adopts a stronger sense of rationalization according to which the rational explanation implies not only that the action seems appealing to the agent (in whatever way), but that the agent is rationally *required* to perform the action. He states that his answer to the motivational challenge “relies on the idea of a rational agent” according to which it is true that “if a rational agent believes that *p* is a conclusive reason to do *a*, she

¹¹ The expressivist may respond that in these cases, contrary to appearance, the agent does not have a normative attitude in the full sense. I am not convinced by this move, but do not have space to argue against it here. Read this passage as outlining my view on the causal aspect of the motivational challenge, which is similar to Scanlon's, rather than fully defending it. My main focus in this paper lies on the rational-explanatory aspect of the motivational challenge.

generally will do *a*, and do it for this reason” (Scanlon 2014, 54, his italics).¹² I suppose that his view is that rationality *requires* such motivation and action, and that agents are not fully rational unless they conform to this requirement. As Scanlon points out: “None of us is perfectly rational, but it is appropriate to call us rational agents just in case we come sufficiently close in meeting these standards. When a rational agent does something that he or she judges him or herself to have reason to do, this judgment makes sense of the action in normative terms and explains it, because the action is what one would expect of a rational agent who accepted that judgment” (Scanlon 2014, 55).

I will, for the sake of convenience, refer to this view as the view that rationality requires *enkrasia* – although the term ‘*enkrasia*’ is often used for the requirement to be motivated in accordance with one’s beliefs about what one *ought* to do. I suppose, however, that the two requirements are close enough to let me apply the term ‘*enkrasia*’ to the former requirement pertaining to beliefs about conclusive reasons in this paper.¹³ I will not be concerned with conformity to ought beliefs in the following, but merely with conformity to beliefs about normative reasons.

The view that rationality requires *enkrasia*, and the psychological setup of the rational agent, is now employed in the explanation of why the agent performed the action. Note that this step is far from being trivial – at least if, as I just assumed, agents may in certain circumstances be indifferent and not perform the action that is rationally required on behalf of their beliefs about reasons. In other words: Rationalization, in this sense of being rationally required, does not guarantee motivation, since agents can act irrationally. How, then, does it contribute to the explanation?

When an action is explained in terms of a normative belief, this is because the agent is rational rather than irrational. It is thus a fact about the agent – namely, that she is rational – that helps us to explain the action by reference to the normative belief. The agent’s rational setup is what the motivational im-

12 Compare also Scanlon 1998, 23–24 and 33–34.

13 Since conclusive reasons “settle the matter of whether to do a” (Scanlon 2014, 105) they seem to correspond to oughts. There are many versions of *enkratic* requirements in the literature. Broome gives a rough description in terms of ought: “*Enkrasia*: Rationality requires of you that you intend to *F* if you believe you ought to *F*” (2009, 96) which is later supplemented to some extent (p.97). Kolodny gives a similar principle that refers to a belief about conclusive reasons (instead of an ought belief): “I+: Rationality requires one to intend to *X*, if one believes that there is conclusive reason to *X*” (2005, 521). Smith (1994, 148) writes: “C2: If an agent believes that she has a normative reason to ϕ , then she rationally should desire to ϕ ”. Gaut claims: “(...) plausibly it is analytic that reasons motivate an agent in so far as she is rational and aware of the reasons” (1997, 161).

pact of her belief about reasons come from. The impact is lacking, however, if an agent is behaving irrationally. Since some agents are quite irrational and since even quite rational agents behave irrationally now and then, the status of the condition that the agent is rational is that of a contingent and, arguably, empirical matter. Maybe, extreme irrationality would be inconsistent with being an agent at all, but it does not follow that agents are necessarily rational, in the required enkratic sense, since the rationality at issue must be more substantial if it is to help explain the particular action.

In short, the explanation of action seems to follow the schema:

Rational Explanation

- (1) Subject S believes that p is a conclusive reason to do a.
- (2) Rationality requires that an agent does a if he or she believes that p is a conclusive reason to do a. Thus, an agent is *rational* only if the agent does a if he or she believes that there is a conclusive reason to do a.¹⁴
- (3) S is a rational agent.
- (4) Thus, S does a.

Claims (1) to (3) jointly form the explanans of this explanation, and claim (4) gives the explanandum. This is a version of Hempel's schema of a rational explanation (but one in which, unlike in Hempel's own version, a belief about reasons rather than a pair of a desire and a means end belief plays the major role).¹⁵ In short, acting against one's belief about conclusive reasons is possible, but incompatible with rationality. If someone is rational, we can thus explain her action by her belief that there is a conclusive reason.

Although I am sympathetic with Scanlon's answer, I fear that it will not convince those who are skeptical with respect to the claim that rationality requires

14 If the rational requirement has wide scope, the rational agent might as well abandon the belief that p is a conclusive reason to do a. This is a problem for the rational explanation, since S might comply with the wide scope requirement without doing a. However, abandoning the belief that p is a conclusive reason would not seem to count as *manifesting or exercising* enkrasia (rather as escaping its demands), and it seems truly appropriate only if there is some evidence that the belief that p is a conclusive reason is false. In this case, the agent manifests another rational trait by abandoning the belief, such as the rational capacity to reconsider one's beliefs in the light of countervailing evidence. Let us ignore these problems for present purposes, assuming there is no sufficient evidence against the belief and thus no way of rationally abandoning the belief about conclusive reasons in the circumstances. This assumption will not be relevant later on, since rational requirements will be detached from the explanation.

15 Hempel thus might not have accepted this enkratic version if he rejected the view that any rational agent is motivated by normative beliefs – he might have accepted only means-end rationality, for instance; cf. Hempel 1961, 291–293.

enkrasia (e. g., Smith 2015). One might, for instance, accept only instrumental rationality and thus deny the alleged conceptual truth that rationalist motivational internalism relies on. Therefore, the argument would be more powerful if it could be detached from controversial assumptions about rationality. I believe that this is possible. Scanlon seems to think that we would then need to identify the neural mechanism that is operating in the rational agent. This is of course something that I cannot offer. However, there might be a different option which draws on dispositions.

More importantly, I think that the explanation needs to be broadened such as to explain also the motivational relevance of belief in reasons that are *not* conclusive, for instance, in optional reasons. These are reasons that do not suffice for placing the agent under a rational requirement (although they generate a rational permission). Their motivational impact needs to be accommodated, since beliefs in such optional reasons often seem to explain our actions, for instance when we decide what to do on a free evening. Rationality seems to permit going to the theatre, as well as going to the party, and the explanation of why the agent went to the party and not to the theatre will not be that doing so was to be expected of a rational person nor that it is rationally required to act in accordance with the reasons for going to the party. Such actions are part of our everyday lives and often arise from the motivational impact of normative beliefs (e. g., beliefs about the reasons for going to the party), even though this is not due to a rational *requirement* to act in accordance with these normative beliefs.

In the following I will try to point out how, by invoking *dispositions* to be motivated by beliefs about normative reasons, normative cognitivism can improve its answer to the motivational challenge in three ways: first, the mechanism that operates when normative beliefs motivate can be sketched a little bit more informatively without diving into neuroscience (sec. 2), second, the strategy can be detached from controversial views about rationality without losing any of its substance or explanatory power (sec. 3), and third, the story of the impact of normative judgments can be extended even to situations in which the action under consideration is not rationally required (sec. 4). We can thus accommodate acting for reasons that are not conclusive.

2 Elucidating the Rational Explanation as a Dispositional Explanation

I will start with the first point and attempt to clarify how we need to understand the claims about the rational agent if they are to do explanatory work. This is

necessary, I believe, since it might be complained that the rational explanation is quite mysterious as long as there is no further specification of how people actually can be rational and about what is happening inside the particular agent who is said to be a rational agent. Is an agent rational if, by mere chance, his normative beliefs happen to cause motivation although they might as well not have done so? Mere causal relations can be due to quite peculiar or extraordinary circumstances. We philosophers are well acquainted with the many sophisticated examples of deviant causal chains in the literature. Even in the unlikely case of an agent's normative attitudes being frequently connected to motivation via various deviant causal chains, these chains would not seem to establish that the agent is rational. So, it might seem as if being rational requires almost magical connections between one's mental states. How can we understand the rational agent in a more plausible and naturalistic way?

The generalization about the rational agent states: "if a rational agent believes that p is a conclusive reason to do a , she generally will do a , and do it for this reason" (Scanlon 2014, 54, his italics). This claim is formulated as a conditional claim, but it should best be interpreted as dispositional if we want to answer the worry about deviant causal chains that I just mentioned. The conditional is prone to counterexamples: it could be true even of an irrational agent if there is a constant source of deviancy. For instance, the conditional may be true of Asya if a neuroscientist manipulated her brain every time she believes that p is a conclusive reason such that she ends up doing a .¹⁶ Conversely, an agent would seem to be rational if she were disposed to do a when she believes that p is a conclusive reason for doing so, even if, for some extraordinary circumstances, she never managed to do a – for instance, if a neuroscientist prevented her from doing so. In other words, the conditional and the disposition can come apart. I gave two examples of this phenomenon, the first involving a mimic of the disposition and the second involving a masked disposition.¹⁷ In both

16 This counterexample is excluded by the supplement "and do it for this reason", but the supplement excludes it precisely by implying that the agent manifests a disposition – at least this is what reflection on what it is to act for a reason tells us, or so I have argued (Mantel 2016 and 2018). Therefore it is more helpful to use a dispositional interpretation from the start (since the disposition is what is really doing the work), such that the counterexample does not even arise and thus not call for a supplement.

17 Since Scanlon speaks of how an agent "generally" (2015, 54) or "normally" (1998, 33) reacts to beliefs about reasons, he might be read as being concerned with the disposition, not the counterfactual, and as hinting at its defeasible relation to the counterfactual. This applies also to the following passage, in which he seems to highlight intrinsic masks of the disposition: "It is an obvious and familiar fact that one's state of mind, the state of one's body, and the content of

cases, the agent's rationality (in the relevant sense of being enkratic) seems to go hand in hand with the disposition – whether or not the conditional is true.

I will therefore assume that the statement is best interpreted as concerned with the *disposition* to do what one believes one has conclusive reasons to do.¹⁸ This disposition is implied by enkratic rationality and explains why rational agents usually – or under normal circumstances – perform the action when the belief is present, since the conditions are not usual or normal in the relevant sense if the disposition is masked etc.

Now, turn to the claim that the particular subject S is a rational agent. This statement must now be interpreted as ascribing the disposition (with a sufficient strength) to a particular subject.¹⁹ However, the ascription of the disposition does not suffice for the explanation to work. As just mentioned, dispositions may be masked from time to time, or not be manifested under certain conditions. A rational subject who has the enkratic disposition may fail to manifest it in a particular situation. This does not imply that the subject is not rational, but it poses a problem for the rational explanation nonetheless. To exclude this problem, the third claim should not only say that S is rational or enkratic, but that S also *manifests* her rationality or enkrasia on this occasion. We need to read the third claim in what Smith calls the “capacity-plus-exercise” sense (Smith 2009, 62). In my own terminology, the claim must state that the agent has the disposition and manifests it with respect to the particular belief, i. e., the belief that p is a conclusive reason to perform the action. If the agent manifests this disposition, then he does so by performing the action, and (provided he also manifests epistemic competence with respect to the reason) he performs the action for this reason (Mantel 2018, chap. 4.1).

Let me point out why this manifestation reading is necessary by considering two cases in which two alternative explanations compete. First, assume that Mila believes that her workload is a conclusive reason not to go to the party in the

one's immediate experiences strongly affect the reasons one attends to: when I have not eaten for some time I just can't keep my mind off food...” (1998, 34).

For discussion of the conditional analysis and problem cases like maskers or antidotes, see, e. g., Johnston 1992 and Bird 1998. For mimics, see Lewis 1997. Further problem cases concern finks (e. g., Martin 1994). For intrinsic masks, compare Ashwell 2010 and Clarke 2010.

18 A brief discussion of this dispositional account of enkrasia is found in Williams 1979, 109. Dispositional accounts of rationality of different kinds are defended, e. g., in Wedgwood 2014, 322, Broome 2009, section 7 and 8; and Broome 2013.

19 As Broome says: “Enkrasia explains why a rational person intends to do what she believes she ought to do: she would not count as rational if she did not. It remains to be explained how she can come to be rational in that respect. (...) First, we may say that a rational person has a disposition to satisfy Enkrasia.” (Broome 2009, 97)

evening, although she is tempted to go to the party anyway. Furthermore, she is angry at the person who throws the party, and thereby somewhat inclined to ignore her invitation (although she does not believe that her anger, or whatever it is that she is angry about, provides a conclusive reason not to go, since the party would certainly be fun). There are thus two dispositions – *enkrasia* and anger – that might be manifested by not going to the party. However, it is possible that Mila stays home by manifesting merely the second disposition, not the first: Her anger, without any help of her normative judgment, was sufficient to make her stay home, and had she not been angry with the host of the party, she would have gone to the party despite her normative judgment that she has conclusive reason to work. Maybe her usually reliable disposition to do what she judges she has conclusive reasons to do was malfunctioning for a moment, or even masked by certain circumstances. If these details are known, we would intuitively reject the explanation that Mila stayed home *because* she judged that she had conclusive reasons to stay home and work – even if she is disposed to conform to such judgments very regularly, and thus has the disposition. This indicates that the explanation requires the *manifestation* of the disposition – not merely its presence, as when it is present but masked.

Second, take a case in which Mila believes that there are two conclusive reasons not to go to the party – her work-related obligations and her duty to watch over her sleeping kids. It might be that she manifests her disposition to do what she believes she has conclusive reason to do only with respect to the second of these judgments but not with respect to the first. If it had not been for her duty to look after her children, she would have gone to the party anyway. In this case, too, the explanation must pick out the judgment with respect to which the disposition is actually manifested. It would be quite misleading to say: “She stayed home because she is rational and judged that she had a conclusive work-related reason to stay home, and rational agents act in accordance with their judgments about conclusive reasons.”

Thus, the explanation should be interpreted as stating not only that the agent has the disposition and manifests it, but also that she manifests this disposition *with respect to the specific belief that is mentioned*. This provides us with a *dispositional analysis* of the explanation at issue.

Rational Explanation as Dispositional Explanation

- (1) Subject S believes that p is a conclusive reason to do a.
- (2') Rationality requires that an agent does a if he or she believes that p is a conclusive reason to do a. Thus, an agent is rational only if the agent is *disposed* to do a if he or she believes that p is a conclusive reason to do a.

- (3') S is *disposed* to do what S believes there is conclusive reason to do and *manifests* the disposition with respect to the belief referred to in (1).²⁰
- (4) Thus, S does a.

No doubt there is more to be said on dispositional explanations, especially on their relation to causal explanations, which I cannot elaborate on here.²¹ At any rate, I take it that this dispositional approach has some intuitive appeal, even if it could be further explored, and that it plausibly and informatively explicates the form of explanation that Scanlon invokes.²² Notably, it does so without diving into the neuronal level (although we would need to descend to this level if we wanted to say *how the disposition is realized* in the particular agent). To the contrary, the disposition in question might be realized by *different* neuronal structures and is thus compatible with several ways in which the agent's neurons are wired up. Similarly, when we say that a person is irascible or that a vase is fragile, we need not be aware of the, possibly multiple, ways in which these dispositions may be based on the neuronal or the atomic level.

3 Detaching the Dispositional Explanation from the Rational Explanation

In this section I will advocate the thought that, contrary to appearance, rationality – understood as a set of valid rational requirements – does not play any essential role in the rational explanation. Rather, all the explanatory work is done by the disposition to conform to one's beliefs about conclusive reasons. In other words, claim (2') is dispensable for the explanation, since claim (3') does the work on its own.

Scanlon is right that, in addition to a causal explanation that might be given at the neuronal level, we need another form of explanation that is more essential to agency, that is concerned with the agent's perspective, and that is therefore given at the *mental level* – at the level, that is, at which rational assessments

20 The more specific description of *p*, if it exists, will distinguish the manifestation of *enkrasia* with respect to one believed reason by contrast to another. In the case of Mila, both claim (1) and (3') will refer to the same belief, i. e., the belief that her obligation to watch over her kids is a conclusive reason not to go to the party, if the explanation is to be correct.

21 For my view on these matters, cf. Mantel 2018, chap. 8.

22 It is informative, for instance, insofar as it excludes the possibility that the agent is manipulated by a neuroscientist and the possibilities that the agent merely manifests anger or manifests *enkrasia* with respect to another conclusive reason, see above.

are made. How to construe this explanation is the most interesting aspect of the motivational challenge. But I will argue that it merely appeared as if this *mental explanation*, as we may call it, was *essentially rational*. Instead, it is essentially concerned with mental dispositions – which may or may not be fully rational ones. Nevertheless, this mental explanation might be described as ‘making the action intelligible’, given the agent’s mental states and *character traits*, if dispositions of (rational or irrational) reasoning be counted among these. But making the action ‘intelligible’ with respect to the agent’s mental states should be understood as showing it to be manifesting the agent’s character or mental dispositions – and nothing explanatory is added, or is missing, whether or not these mental dispositions correspond to (or are said to correspond to) a rational requirement like *enkrasia*.

If these thoughts are on the right track and the dispositional explanation can be given independently of the theory of rationality, this would have an important advantage for normative cognitivism. It would allow to formulate the explanation of the action less controversially, since the explanation would not rely on the assumption that practical rationality goes beyond instrumental rationality. Instrumentalism, the view that practical rationality consists merely in means-end rationality, is often viewed as the “default view” (Millgram 2001, 4). The view that practical rationality might demand anything more than instrumental reasoning seems unpalatable to some, for instance because it is incompatible with the Humean theory of rationality.²³ While, I suppose, Hume does not deny that (normative) beliefs may contingently *cause* desires, he denies that it is *irrational* if they do not. The cognitivist’s solution to the challenge would be rendered more widely acceptable if it could do without the claim that rationality requires *enkrasia*.

It is quite obvious how to do without this claim. On the dispositional interpretation, the explanation of the action works just as well without claim (2’). The rational explanation, as I interpreted it, presupposes that some individuals are disposed to do what they believe they have conclusive reasons to do, and that the agent has this disposition and manifests it with respect to the belief in question. This form of explanation does not seem to depend on how we label these agents – as the *rational* agents, or just as the ones who are so disposed. In fact, arational and irrational dispositions of agents may also explain their arational and irrational behavior. Therefore, the explanation could be stated simply in

²³ An example is Smith, who rejects Scanlon’s version of internalism (*reasons* judgment internalism) on these grounds – even though he hopes that another form of rationalist motivational internalism, called *evaluative* judgment internalism, might be more defensible (Smith 2015).

terms of ascribing an agent a disposition – whether or not this disposition corresponds to a rational requirement. Call this the Dispositional Explanation.

Dispositional Explanation

- (1) Subject S believes that p is a conclusive reason to do a.
- (3') S is disposed to do what she believes there is conclusive reason to do and manifests the disposition with respect to the belief referred to in (1).
- (4) Thus, S does a.

Note that I do not make this suggestion because I, personally, doubt that it is rationally required to conform to one's beliefs about conclusive reasons. By contrast, I sympathize with Scanlon's view on this matter (and think that it would be desirable to defend this view of rationality further). However, I think that this thesis about rationality is *not needed* for the explanation to work, and that the explanation will be more acceptable to Scanlon's critics if it is reformulated without the commitment to his views on rationality. My point is that the normative cognitivist can give an account of the way in which beliefs about reasons motivate even without being committed to particular views about rationality.

Are we missing out on anything important without alluding to a rational requirement in the explanation? I don't think so. The rational requirement concerns merely the labelling of the agent's disposition, not its nature and explanatory power. It could be responded that it would seem to be expected that many have the disposition in question if the disposition is a rational one, but that without the commitment that the disposition is rational we need a further explanation about why at least some agents have this disposition. According to this objection, claim (2') is to be understood as (partial) support for why claim (3') is true. The agent has the disposition in question *because* it is a rational disposition to have.²⁴ However, I am not persuaded that there is an extra explanatory burden when the commitment to rationality is given up. We may ask what function the disposition plays in our lives or how it evolved in evolutionary processes. However, we may ask these questions just as well about irrational dispositions as about rational ones. If rationality consists in conforming to one's beliefs about conclusive reasons, how come that we, or sufficiently many of us, evolved to be (sufficiently) rational rather than irrational, and that we make use of our ra-

24 It is not clear that this explanation makes much sense. It resembles the claim that the agent is disposed to conform to normative judgments because the agent *should* be so disposed – although it does not seem to be true that agents are always, or even mostly, as they should be. At most, if minimal forms of rationality are constitutive of being an agent, then agents are at least minimally as they should be. This does not explain much.

tional capacity in our everyday lives? Which function does it serve? These questions apply whether or not the disposition is (correctly or incorrectly) labeled a rational one.

I concede that there is some sort of special explanatory relevance that claim (2') might have – it might perhaps explain why S is not *functioning properly* if S does not have the disposition in question, or why S *should* have the disposition (at least from the perspective of rationality). But this explanatory import is directed at explaining why S acted as he *should* when performing the action, not at why S performed the action rather than not doing so. It is thus directed at an altogether different explanandum.²⁵

Interestingly, by detaching the explanation from the theory of rationality the explanation of the rationalist motivational internalist becomes available to the motivational externalist as well. Whereas Scanlon's formulation invoked a claim about rationality that rendered it (defeasibly) internalist, the new formulation is neutral when it comes to the debate between internalists and externalists. It may or may not be combined with the view that rationality requires enkrasia – and thus requires having the disposition mentioned in claim (3'). Depending on the theory of rationality it is combined with, the dispositional explanation may be a component of rationalist motivational internalism (as in Scanlon's work) or of externalism (as it is understood, e.g., by Dreier 2000).

Rationalist motivational internalism, as I understand it here, states that it is a conceptual truth that a fully rational agent who believes that there is a conclusive reason for her to perform action a is sufficiently motivated to perform action a and hence performs action a (or at least tries to perform it).²⁶ Externalism de-

²⁵ Dreier (2015) takes the explanandum to be why it is *irrational* not to be motivated in accordance with one's beliefs about what one ought to do. This is the question which the realist does not answer. I agree that this question is important in its own right, but I would like to disentangle it from the question why the action is performed (which seems to be the question with which Scanlon is concerned). The motivational challenge, as Scanlon understands it, is concerned with the question of why the agent performed the action (rather than not), and not with the quite different question of why, if the agent performed the action, he acted as he should have. If the motivational challenge were about the latter question, then it would require an argument for why rationality requires conformity to judgments about reasons, which Scanlon does not try to provide in his discussion of it.

²⁶ A rational agent who justifiedly believes that there is a conclusive reason to leave the room might of course fail to leave the room if, upon trying to leave, he finds out that he is locked in. Such failure is not due to the motivational setup of the agent but to external circumstances.

This formulation of internalism which holds that beliefs with this specific content motivate rational agents is to be distinguished from the formulation in Tresan 2006, 147 (where the content does not need to employ the concept of a normative reason but may instead employ a descriptive concept that *de facto* picks out a normative reason) and Zangwill 2015, 46 (where inter-

nies this connection. Since this conceptual connection seems to hold especially due to the conception of rationality in play, our verdicts about the conceptual connection will vary with our conception of rationality. It therefore seems misleading to describe rationalist motivational internalism as concerned with a conceptual connection between two things, *judgments about conclusive reasons and motivation*, but one which holds merely for the rational agent. Rather, rationalist internalism is concerned with a conceptual connection between three things, it seems, namely the aforementioned two *and rationality*.²⁷ Depending on the theory of rationality that is combined with the dispositional explanation, the explanation will or will not seem to vindicate this conceptual link. It thus seems that the very same form of dispositional explanation yields an externalistic or an internalistic picture, depending on whether it is paired with a certain theory of rationality.

This observation has implications for the so-called fetishism objection to externalism. According to Smith's objection to externalism, externalism inevitably leads to moral fetishism, namely to acting in accordance with normative reasons in a merely instrumental fashion, whereas a rationalist internalism avoids these problems (Smith 1994, 71–76). If my observations are correct, it cannot be true that rationalist internalism and externalism fare differently with respect to moral fetishism, since they may appeal to the same disposition. Nothing precludes that this disposition issues in non-instrumental motivation to perform an action. The rationalist internalist is committed to a dispositional (and arguably non-instrumental) story to make sense of rational explanations, but this dispositional story can be used by the externalist as well – without the rationalist

nalism is understood as the claim that the normative judgment causes motivation without the help of any *desire*).

²⁷ In van Roojen's terminology, this is Defeasible Reasons/Motives Judgment Internalism (van Roojen 2015, 58) – defeasible in the sense that the connection holds only under the condition that the agent is rational. I do not find it very illuminating to allow for defeasible internalism since the statement about the defeasible conceptual connection in the rational agent sounds, to my ears, as misleading as the statement that there is a conceptual relation between sunshine and fissures, though it is defeasible and holds only for certain materials, i.e., for materials disposed to display fissures when exposed to sunshine. By contrast, it is clear what is meant by saying that there is a conceptual connection between three things: the disposition to display fissures when exposed to sunshine, sunshine, and fissures. The conceptual relation the rationalist motivational internalist is after will not be quite as straightforward, maybe, but should nevertheless be viewed as a three-place relation.

superstructure. Both theories may appeal to the very same psychological mechanisms and will thus fare the same with respect to moral fetishism.²⁸

If this is true, then internalists and externalists are equipped with the same resources to deal with the motivational challenge. More precisely, a *rationalist* version of defeasible internalism and a *dispositional* externalist approach offer the same solution in different appearance. They appeal to the same disposition and its manifestation, where this disposition is or is not described as a rational one.²⁹

4 Expanding the Dispositional Explanation to Inconclusive Reasons

I conceived of the motivational challenge as the challenge to understand how beliefs about conclusive reasons motivate agents to act. However, agents are motivated not merely by the belief that something is a *conclusive* reason (i. e., a reason that settles the matter of whether to do a) but quite often they are motivated by the belief that something is an inconclusive reason. The motivational challenge thus has a broader scope. How do beliefs about reasons issue in action even if the agent takes the reasons to be inconclusive?³⁰

Take an example. When deliberating about what to do in the evening, Jennifer thinks that there is a reason to go to the party, namely that she will enjoy being there. She does not believe that this is a conclusive reason, i. e., a reason

28 For a discussion of dispositional externalism, see Broome 1997, esp. 144, and Dreier 2000, who points out that the externalist might appeal to a disposition but does not seem to see that rationalist internalism shares these dispositional commitments, and thus the costs, of externalism. Dreier furthermore helpfully points out that the disposition to be motivated by beliefs about reasons and a desire to do what reasons require seem to have different features (see also Zangwill 2015, 54). If these differences could be denied (which I do not think), however, my arguments would indicate that the rationalist motivational internalist was committed to the view that the rational capacity of *enkrasia* consists in the *desire* to do what reasons require. The rationalist motivational internalist would thus have to accept the influence of this desire just as the externalist does.

29 Only strong internalism provides a different story, but one which is implausible since it does not allow for *akrasia*, see section 1.

30 This question has been formulated as a rationale for skepticism with respect to dispositional accounts by Bridges (2011, 199–203). An inconclusive reason for action a is a reason that is not conclusive, i. e., that does *not* settle the matter whether to do a, cf. Scanlon 2014, 105. An inconclusive reason may be sufficient, however, it may make the action reasonable and rationally permitted.

that settles that she *ought* to go to the party. She believes that staying home and watching TV would also be okay, and going to the cinema might even be as enjoyable as going to the party. There might be a further activity – although she isn't sure what it would be – that she would enjoy even more than going to the party. But she doesn't mind whether going to the party is the optimal course of action for tonight. There is a reason to go to the party – albeit an inconclusive reason – and no conclusive reason against going. So, she goes to the party.

As the example shows, when deliberating about how to spend the evening, people often manifest a tendency to perform the first activity that comes to their mind as an activity that is favored by reasons that are good enough – going to the theatre, or to the movies, or reading a book. But they might not even try to decide whether the activity is favored by the balance of all reasons that apply to them, or whether it is true that they *ought* to engage in it.³¹ How does the belief about inconclusive reasons explain their actions?

Note that this question is especially urgent for someone who, like Scanlon, believes that some reasons are optional, such that one is not open to rational criticism even if one does not act on them without having countervailing reasons against doing so (Scanlon 2014, 106). These reasons “render an action rationally eligible without making it rationally required in the absence of some countervailing reason” (Scanlon 2014, 107). If beliefs in optional reasons motivate rational agents, which they seem to do, we must explain how they do so. Since there is no valid rational explanation based on a *requirement* to act on specific optional reasons, a dispositional explanation must come to the rescue.

Scanlon's rational explanation – or, at any rate, my reconstruction of it – does not seem to fit for these cases because it is not as if we could find a principle akin to principle (2') in his Rational Explanation, for instance: “(2'') Rationality requires that an agent does a if he or she believes that p is an inconclusive reason to do a and he or she is in circumstances C, e.g., is deliberating about how to spend the evening. Thus, an agent is *rational* only if the agent is *disposed*

31 Maybe many people are usually disposed to do what is favored by a reason with sufficient strength *only if* they do not at the same time believe that they ought not to do so. That is to say, if an agent is also *enkratic*, as we would hope she is, then her disposition would either be such that it is triggered only in (or partly by) the absence of a belief that there is a conclusive reason not to act accordingly, or the disposition would hopefully at least be overruled by the more powerful disposition of *enkrasia* if the agent believes there is a conclusive reason not to perform the favored action. However, we should also allow for *akrasia* even when it comes to being motivated by beliefs about reasons. Some agents may be disposed to choose an activity for the evening when they believe that there are some reasons of entertainment for doing so, and they may manifest this disposition despite their belief that there are conclusive reasons to work, or to get engaged in moral projects instead, or the like.

to do a if he or she believes that there is an inconclusive reason to do a and he or she is in C.”

Rationality does not *require* that agents have such dispositions to conform to inconclusive reasons – they are merely rationally permitted to have them, but all the same permitted not to have them. This means that agents may be fully rational without the disposition, and that we cannot explain that they are motivated by their belief in inconclusive reasons, if they are, by appealing to the fact that they are rational, for they would be just as rational if they were not so motivated.³²

However, it should be clear from the previous sections how to construe a dispositional explanation of the action. Agents like Jennifer seem to manifest a disposition, albeit a disposition which might be relevant only in certain circumstances, such as when deliberating about how to spend one’s spare time or about practical questions that are taken to be of little importance. They are disposed to be motivated to do what they believe they have a reason to do (at least with respect to reasons of a certain kind), even if they do not believe that the reason is conclusive.

The disposition may explain why the agent acts for one particular optional reason, rather than for another, by referring to some feature F that distinguishes this optional reason from others. This feature will be insignificant from a rational point of view, of course, since all optional reasons share the same rational status. The feature will, however, make a difference to how the agent attends and reacts to the optional reason, and will lead the agent to prefer reasons of a certain kind.

Thereby, the belief that there is an inconclusive reason explains the action all the same, in the following way:

32 It might be said that we may explain their action by their rationality and by appeal to a rational permission rather than to a rational requirement. Thereby, we would merely explain why they act in accordance with at least one belief in a sufficient (or optional) reason, but not why they act in accordance with the specific normative belief in question. We would explain, that is, why they refrain from the actions that are not, by their lights, supported by sufficient normative reasons, but we would not explain why someone goes to the party rather than reading a book, if the agent believes that both options are supported by optional reasons. I think that we thereby do not explain the particular action of going to the party – at least not when contrasted with the action of reading a book – but merely explain why other actions are ruled out. It is also not plausible that a rational permission gives at least a probabilistic explanation of the action that is being performed, since the pool of permissible actions may be huge and the probability of going to the party given that one chooses a rationally permitted action might be tiny, and no greater than the probability to choose any other permissible action.

Dispositional Explanation – Inconclusive Reasons

- (1') Subject S believes that p is an inconclusive (e. g., optional) reason to do a, where this reason has feature F (e. g., it is the first reason that comes to mind, or it is a reason that is related to a certain project) and S is in circumstances C (e. g., S is deliberating about how to spend the evening).
- (3'') S is disposed to do what, in circumstances C, she believes there is an inconclusive reason with feature F to do and S manifests the disposition with respect to the belief referred to in (1).
- (4) Thus, S does a.

To sum up, a rational explanation, as I understand it here, would use the claim that rationality requires the action, given the agent's normative judgment, and that the agent is rational. Yet the action cannot be explained by saying that the agent is rational – even if she is – since rationality is compatible just as well with not doing what is favored by the optional reason. Therefore, the only explanation of actions performed for optional reasons is one that refers to a disposition which is not required by rationality. Such a disposition might be referred to as arational, e. g., the disposition to be motivated in circumstances C by the first optional reason that comes to mind that is manifested by Jennifer. Jennifer may be just as rational without having this particular disposition. At most, it may be required for her full rationality that she has some tie-breaking disposition that helps her to decide when it comes to optional reasons that pull in different directions, lest she should be paralyzed by such reasons as Buridan's ass by two equal stacks of hay. But this shows merely that, in order to be fully rational, she must have some such disposition or other, while her rationality does not require her to have any particular instance of such a disposition.

Going one step further, I see no problems for the view that even clearly *irrational* dispositions to conform with certain reasons (e. g., with certain reasons that are believed to be outweighed) may be of a similar explanatory power as rational or arational dispositions. Neither arationality nor irrationality impugns on a disposition's explanatory power. We need arational dispositions to explain why agents act on beliefs in inconclusive reasons, and this fact yields further support for the view that the dispositional explanation implied by Scanlon's rational explanation can do its job just as well if it is detached from Scanlon's theory of rationality. In such a detached form, the explanation should be all the more convincing even in the eyes of Scanlon's critics, and it should be accepted as a non-mysterious, uncontroversial, and broad answer to the motivational challenge.

Conclusion

Often people do what they judge there is a reason to do. The motivational challenge is to explain their motivation and action if this judgment is a belief. I am sympathetic to Scanlon's view that when the agent judges that the reason is conclusive, the agent is *rationally required* to form a motivation (although I did not argue for this view here). I do not agree, however, that this motivation calls for an essentially *rational explanation* – rather, it can be fully explained by a *mental and dispositional explanation* that is not committed to views about rational requirements. In order to arrive at the characterization of this explanation, I gave Scanlon's rational explanation a metaphysical interpretation according to which rational capacities imply dispositions that play an explanatory role. Subsequently, I removed the theory of rationality from the dispositional explanation, and expanded the dispositional explanation to judgments about inconclusive reasons. I argued that the resulting dispositional explanation is more acceptable to Scanlon's critics and therefore more useful when defending normative cognitivism, and that it can be extended to cover the true scope of the motivational challenge, which encompasses actions performed for inconclusive reasons.

References

- Ashwell, L. 2010. Superficial Dispositionalism. *Australasian Journal of Philosophy* 88 (4): 635–653.
- Bird, A. 1998. Dispositions and Antidotes. *The Philosophical Quarterly* 48: 227–234.
- Björnsson, G., Strandberg, C., Olinder, R. F., Eriksson, J., and Björklund, F. 2015. Motivational Internalism: Contemporary Debates. In: *Motivational Internalism*, eds. G. Björnsson, C. Strandberg, R. F. Olinder, J. Eriksson, and F. Björklund, 1–20. Oxford: Oxford University Press.
- Bridges, J. 2011. Dispositions and Rational Explanation. In: *The Possibility of Philosophical Understanding: Reflections on the Thought of Barry Stroud*, eds. J. Bridges, N. Kolodny, and W. Wong, 182–216. Oxford: Oxford University Press.
- Brink, D. O. 1997. Moral Motivation. *Ethics* 108: 4–32.
- Broome, J. 1997. Reasons and Motivation: II. *Proceedings of the Aristotelian Society* 71 (1): 131–146.
- Broome, J. 2009. Motivation. *Theoria* 75: 79–99.
- Broome, J. 2013. *Rationality through Reasoning*. Wiley-Blackwell.
- Clarke, R. 2010. Opposing Powers. *Philosophical Studies* 149: 153–160.
- Davidson, D. 1963. Actions, Reasons, and Causes. Reprinted in *Essays on Actions and Events*, ed. D. Davidson, 3–20. Oxford: Clarendon Press, 1980.

- Dreier, J. 2000. Dispositions and Fetishes: Externalist Models of Moral Motivation. *Philosophy and Phenomenological Research* 61 (3): 619–638.
- Dreier, J. 2015. Can Reasons Fundamentalism Answer the Normative Question? In: *Motivational Internalism*, eds. G. Björnsson, C. Strandberg, R. F. Olinder, J. Eriksson, and F. Björklund, 167–181. Oxford: Oxford University Press.
- Frankfurt, H. G. 1971. Freedom of the Will and the Concept of a Person. *The Journal of Philosophy*, 68 (1): 5–20.
- Gregory, A. 2017. Are All Normative Judgments Desire-Like? *Journal of Ethics and Social Philosophy* 12(1): 29–55.
- Hempel, C. G. 1961. Rational Action. Reprinted in *Readings in the Theory of Action*, eds. N. S. Care and C. Landesman, 285–286. Bloomington, Indiana: Indiana University Press, 1968.
- Johnston, M. 1992. How to Speak of the Colors. *Philosophical Studies* 68 (3): 221–263.
- Kolodny, N. 2005. Why Be Rational? *Mind* 114, 509–563.
- Lewis, D. 1997. Finkish Dispositions. *The Philosophical Quarterly* 47: 143–158.
- Mackie, J. L. 1977. *Ethics: Inventing Right and Wrong*. Harmondsworth, Middlesex: Penguin Books.
- Mantel, S. 2016. Three Cheers for Dispositions: a Dispositional Approach to Acting for a Normative Reason. *Erkenntnis* 82 (3): 561–582.
- Mantel, S. 2018. *Determined by Reasons: A Competence Account of Acting for a Normative Reason*. New York: Routledge.
- Martin, C. B. 1994. Dispositions and Conditionals. *Philosophical Quarterly* 44 (174):1–8.
- Millgram, E. 2001. Practical Reasoning: The Current State of Play. In *Varieties of Practical Reasoning*, ed. E. Millgram, 1–26. Cambridge, Massachusetts: MIT Press.
- Nagel, T. 1970. *The Possibility of Altruism*. Princeton, New Jersey: Princeton University Press.
- Olson, J. 2014. *Moral Error Theory: History, Critique, Defence*. Oxford: Oxford University Press.
- Parfit, D. 1997. Reasons and Motivation: I. *Proceedings of the Aristotelian Society*, 71 (1): 99–130.
- Scanlon, T. M. 1998. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.
- Scanlon, T. M. 2014. *Being Realistic about Reasons*. Oxford: Oxford University Press.
- Smith, M. 1994. *The Moral Problem*. Oxford: Blackwell Publishing.
- Smith, M. 2009. The Explanatory Role of Being Rational. In *Reasons for Actions*, eds. D. Sobel and S. Wall, 58–80. Cambridge: Cambridge University Press.
- Smith, M. 2015. Evaluative Judgements, Judgements about Reasons, and Motivations. In: *Motivational Internalism*, eds. G. Björnsson, C. Strandberg, R. F. Olinder, J. Eriksson, and F. Björklund, 27–43. Oxford: Oxford University Press.
- Stocker, M. 1979: Desiring the Bad: An Essay in Moral Psychology. *Journal of Philosophy*, 738–753.
- Thomson, J. J. 2006. The Legacy of *Principia*. In *Metaethics After Moore*, eds. T. Horgan and M. Timmons, 233–254. Oxford: Oxford University Press.
- Tresan, J. 2006. De Dicto Internalist Cognitivism. *Noûs* 40 (1):143–165.
- van Roojen, M. 2015. *Metaethics: A Contemporary Introduction*. New York and London: Routledge.
- Wedgwood, R. 2014. Rationality as a Virtue. *Analytic Philosophy* 55 (4):319–338.
- Williams, B. 1979. Internal and External Reasons. Reprinted in *Moral Luck*, ed. B. Williams, 101–113. Cambridge: Cambridge University Press, 1981.

Zangwill, N. 2015. Motivational Externalism. In: *Motivational Internalism*, eds. G. Björnsson, C. Strandberg, R. F. Olinder, J. Eriksson, and F. Björklund, 44–60. Oxford: Oxford University Press.

Thomas Nagel

Moral Reality and Moral Progress

How should we think of moral progress? The history of humanity includes progress (often interrupted or reversed) of multiple kinds: scientific, technological, artistic, legal, social, economic. I am interested in the kind of moral progress that can be regarded as an advance in understanding, or knowledge, not just in behavior. That assumes a realist position about moral truth – that moral propositions can be true or false independent of what we believe. So one kind of moral progress would be to give up a false moral belief in favor of a true one. Another would be to arrive at a true belief about something about which one did not have an opinion before, perhaps because the question had not arisen. Such changes can occur either in the attitudes of an individual or in the shared attitudes of a community. But what does it mean to describe such a change as progress?

The question can be sharpened by comparing moral progress with scientific progress. On a realist understanding of science, scientific knowledge tells us the truth about a natural world that exists independently of us and our beliefs about it. A scientific discovery in physics, chemistry, or biology reveals something about the world that was true long before we discovered it, and would have been true even if we never discovered it. If humans had never developed chemistry – indeed even if humans had never existed – it would still have been true that salt is sodium chloride. Scientific progress, at least in the basic natural sciences, consists in the discovery of truths that have been true all along, because the world has existed with these same basic characteristics all along.

Moral truth, however, if there is such a thing, is not about the natural order – about the structure of the world, or its composition, or what happens in it. It is about what we have certain kinds of reasons to do and not to do. These reasons are not objects in a moral region of the universe with which we interact: we can't think of them as like chemical elements waiting to be discovered. Sometimes we will want to say that moral progress consists in the coming to acceptance of a moral belief that was true all along, or the abandonment of a belief that was false all along; but our understanding of how this is so must be different from the scientific case. Further, I think that sometimes a claim about truth will depend on a claim about progress rather than the reverse. That is, there may be no more to the claim that a newly adopted moral position was true even before anyone adopted it than that its adoption is a case of moral progress by comparison with prior attitudes.

<https://doi.org/10.1515/9783110733679-008>

The key to exploring this topic is that the normative domain is a domain of reasons. Moral progress must be identified with a change in moral outlook that there is reason to adopt. Such a reason is tied to the time in moral history to which it is assigned. It need not imply that the same reason existed at all times; whether it existed in the past is a further question, whose answer may depend on the circumstances then, and what considerations were available or accessible.

The contrast with science is stark. In 400 B.C. no one could have even understood, let alone had reason to believe, that salt was sodium chloride. It would require a long path of scientific theory and experiment to reach the conceptions of elements and chemical compounds that allowed such a proposition to be formulated and confirmed. Yet it was as true then as it is now. Its truth is in no way dependent on there being any reason to believe it. (Let me say parenthetically that the same is true of mathematics. Mathematical truth will always vastly outrun what we have reason to believe.) By contrast, the truth of a moral proposition cannot be distinguished from there being a reason for people to conduct themselves in the way it prescribes. And I would claim that there cannot be such a reason if there is no reason for anyone to believe there is. It is of course possible for some people to have reasons not to believe in the existence of moral reasons that actually exist, or to believe in the existence of moral reasons that do not actually exist – reasons of religious authority, for example. But when such a situation obtains, the morally correct position must be supported by reasons that are accessible – reasons that those same people have some reason to believe in, even if they are not acknowledging them at the moment. They cannot be completely hidden, as was the chemical composition of salt in the distant past.¹

How might an improved moral outlook become accessible that was not accessible before? One possibility is that the non-normative facts, or knowledge of those facts, may change in a way that poses new moral questions that have not had to be addressed before: for example, it becomes clear that the burning of fossil fuels is contributing to a potentially dangerous increase in global temperatures; or advances in medicine pose questions about end-of-life decisions or parental surrogacy; or the development of a modern market economy poses questions about the legitimacy of agreements that block competition. A simple account of what happens in such cases could be given if there were a single governing moral principle, such as utilitarianism, from which all moral reasons were derived, and changes in the known consequences of what people were able to do generated corresponding changes in what they ought to do. It might

1 See Raz (1994).

be difficult to do the cost-benefit calculation, and there might be great uncertainty, but the answer would not depend on discovering new moral reasons of a kind that hadn't been thought of before. It would depend only on what principles of conduct, or policies, or institutions were justified by the single standard of impartial welfare maximization.

If moral reasons were like that, moral progress would in a sense resemble scientific progress, because whatever is the right answer to a newly posed question, about how to respond to climate change or medical advances, would have been true hypothetically long in advance. That is, it would have been true before anyone thought of it that *if* certain forms of artificial life support were to become available, then it would/would not be permissible to turn off the respirator/remove the feeding tube from a patient in a vegetative state under such and such conditions of prior consent. In other words, all moral progress would be the discovery of moral truths that in a sense had been true all along, like the chemical composition of salt – because they followed from a premise that any rational being could have recognized. But since I do not believe that all moral reasons derive from a single utilitarian axiom, I will leave this possibility aside.

I am interested in the interpretation of moral progress if we assume instead a more pluralistic and complex conception of moral reasons, though nonetheless a realist one. I believe there will be cases where reasons come to be recognized that could not have been recognized much earlier – either in response to the presentation of new choices by new non-normative facts, or as a result of moral reflection that revises or extends existing moral attitudes. I believe it makes sense to be a realist about such reasons simply by holding that judgments made in such circumstances can be correct or incorrect, and that their correctness or incorrectness consists in their appropriateness as responses to those very circumstances. The most basic truth here, in other words, may be local – that this is how we should now move forward. Whether it shows that it was always wrong to think anything else is another question, the answer to which depends on what there was reason to think in other circumstances, in the past. Sometimes moral progress will be presentable as the discovery of what was “true all along”; sometimes it won't. In the latter case, it will be because the recognizability of such truth (like the progress of scientific knowledge) is path-dependent: that a certain policy or practice would be an improvement may be understandable on reflection only by those who have already passed through certain prior stages of moral thought and practice.

We have seen many historical examples of moral progress, usually connected with political and institutional progress: the abolition of slavery, the replacement of aristocracy by popular sovereignty, the growth of religious toleration and freedom of expression, the elimination of cruel punishments, the emancipa-

tion of women, the abandonment of racial discrimination, the attempt to create equality of socio-economic opportunity, the defense of sexual freedom, improvement in the treatment of animals. Change involves both the work of conscious moral reformers and moral or political theorists and disseminated alteration in the attitudes of large populations, at various velocities. To call these changes progress is to make a normative claim: that there were moral reasons to replace the practices prevailing at the time with something else. In some cases, like the abolition of slavery, those reasons imply that there was never any justification for the practice being replaced; but not always. For example, I believe the moves to popular sovereignty and equality of opportunity between the sexes required certain levels of political, economic, and educational development before they were even imaginable as options.

The clearest cases in which moral progress is the recognition of something that was true all along are cases where a great injustice is overthrown. The reasons for the change are not subtle, and typically they have long been obvious to some people (and not only to the victims of the injustice), even if most have been oblivious to them. An example that we have lived through is the recent dramatic and rapid change in the status of homosexuality in most Western countries. This began with a campaign to get the state out of the business of enforcing standards of personal morality – as opposed to preventing conduct that harms others – and to persuade people that voluntary sexual conduct is a private matter. This argument deliberately abstained from requiring those who thought homosexuality immoral to change their opinion: it asked them only not to use law to enforce that opinion. The strategy had some success in decriminalizing homosexual conduct, but the appeal to mere toleration that was ostensibly behind it was not satisfactory, and it came under assault with the movement for gay rights and the massive refusal of homosexuals to hide any longer. In spite of opposition by some religious communities, the view that there is nothing wrong with homosexuality made extraordinarily rapid progress, driven by the discovery, once the closet was flung open, that almost everyone had friends, relatives, and colleagues who were gay, and whose full humanity they could not deny. Heterosexuals who had been formed in a homophobic culture were able, amazingly, to make a rapid switch from being unable to imagine the erotic lives of homosexuals without fear and disgust to recognizing this as another form of human love. And the young coming of age in the new climate often couldn't even understand the old taboos.

This is a case of moral progress that thoroughly discredited the old outlook. The biological function of sex as the means of procreation provided no reason to condemn or be ashamed of sexual desires that could not result in procreation, and the reasons of individual freedom and happiness against blocking the fulfil-

ment of those desires were overwhelming. So I think it would be correct to say that this alteration in moral opinion is an advance because the formerly prevailing opinion had been wrong all along. There was never a justification for suppressing homosexuality, however many people believed there was. The reasons against doing so were in principle accessible to anyone.

It may be that the next stage of reform, the legalization of same-sex marriage, is an example of the second type of moral progress – an advance at a particular historical time whose identification as an advance depends on reasons that did not exist, and that could not have been recognized, very much before the issue was raised in the context of the gay emancipation that preceded it. But I won't take up that question, and will instead illustrate the second type of progress with a different example, taken from the writings of T. M. Scanlon.

In his 1972 paper, "A Theory of Freedom of Expression," Scanlon defends what he calls the "Millian Principle," because he takes it to be a natural extension of the thesis Mill defends in chapter 2 of *On Liberty*. The principle reads as follows:

"There are certain harms which, although they would not occur but for certain acts of expression, nonetheless cannot be taken as part of a justification for legal restrictions on those acts. These harms are: (a) harms to certain individuals which consist in their coming to have false beliefs as a result of those acts of expression; (b) harmful consequences of acts performed as a result of those acts of expression, where the connection between the acts of expression and the subsequent harmful acts consists merely in the fact that the act of expression led the agents to believe (or increased their tendency to believe) these acts to be worth performing."²

Natural extension or not, it is an original proposal, and Scanlon's normative defense of it is grounded in a particular conception of political legitimacy. He says it "is a consequence of the view, coming down to us from Kant and others, that a legitimate government is one whose authority citizens can recognize while still regarding themselves as equal, autonomous, rational agents."³

I shall not set out in detail the argument connecting the principle with the conception. It is subtle and ingenious, a superb example of moral reflection that leads to a new understanding of the scope and grounds of freedom of expression. The point I want to make is that these reflections are possible only

² Scanlon (1972); p. 14 in Scanlon (2003). When I presented this paper at the Lauener Symposium, Scanlon reminded me that in a subsequent paper, Scanlon (1979), he had abandoned the Millian Principle in favor of a rule-consequentialist theory of freedom of expression. However, since I continue to find the Millian Principle plausible, I will persist with the example.

³ *Ibid.* pp. 14–15.

for someone who understands from the inside the conception of political legitimacy on which they depend. Someone who believed that the authority of rulers depended on a divine source, or on dynastic right, or who believed with Hobbes that those who command a monopoly of force deserve our allegiance as the only protection against the horrors of anarchy, would not be able to understand this kind of restriction on the grounds for the exercise of their power. Such a pre-liberal subject might still find fault with legal restrictions of expression on the ground that they have bad consequences or are based on false beliefs, but he would not be in a position to criticize them for relying on justifications that undermine the acceptability of the state to autonomous subjects. He would not have reached the stage that enables him to think, as Scanlon puts it, that the state lacks “the right to deprive citizens of the grounds for arriving at an independent judgment as to whether the law should be obeyed.” (18)

I am claiming, in other words, that one needs to have already arrived at a modern understanding of the conditions of political legitimacy and the autonomy of the individual in relation to the state, in order to be able to engage in the reasoning that allows one to see what it entails with regard to freedom of expression. This is moral progress, and it should be understood realistically – i.e. as the discovery of objective reasons to adopt this new principle. But the reasons did not always exist, because the conditions for seeing them were not always present.

There is a natural objection to this claim: Couldn't someone with highly developed rational faculties living in the England of Henry the Eighth, for example Sir Thomas More, have come to see that the regime lacked legitimate political authority because that authority could not be recognized by citizens who regard themselves as equal, autonomous, rational agents? And couldn't he then have gone further to draw the conclusion, using Scanlon's arguments, that a government that was legitimate in this sense would not have the authority to restrict freedom of expression on grounds that violate the Millian Principle (so that his own persecution of heretics was impermissible)?

I do not think so. Such thoughts do not depend only on the exercise of a universal faculty of reason, but require that the thinker have a conception of the relation of the individual to the state that has been shaped by actual historical developments and the experience of living under institutions that claim secular authority of the kind in question. Thomas More did of course challenge the authority of the king, but only on the basis of a higher authority – not in light of his status as an equal, autonomous, rational agent. I also think that More could have come to recognize that his persecution of heretics was wrong – by recognizing that his grounds for certainty in his religious beliefs were insufficient to warrant sending people to the stake. But the Millian Principle was out of his reach.

In a way I am conceding a criticism that Bernard Williams made against me, among others, regarding the universalist pretensions of political liberalism – though I do not accept the “relativism of distance” that he favored as an alternative. He memorably quipped:

“Must I think of myself as visiting in judgment all the reaches of history? Of course, one can imagine oneself as Kant at the Court of King Arthur, disapproving of its injustices, but exactly what grip does this get on one’s ethical thought?

In particular, is it really plausible that one makes this imaginary journey only with the minimal baggage of reason? Granted the notable fact that no one had the liberal world view then, the ethical time-traveler must take with him implicitly the historical experience which has made him the liberal he is, and that experience does not belong to the place he is visiting.”⁴

Rather than relativism, I am offering a conception of the relation between moral realism and moral progress that requires realism only about the reasons for a change in moral outlook at the time it occurs. Sometimes progress will occur as the result of the recognition of reasons that have existed unrecognized for a long time; but not always. The point is that the claim of moral progress is itself a moral comparison between two available alternatives, not a comparison of both of them with some independent moral reality. The question is, “Where should we go from here?” I believe this is how we should think of the moral problems facing us in the domains of socioeconomic inequality, global justice, and our obligations to the future in the light of climate change. One can be as much of a realist about emerging moral reasons as about timeless ones.

But if it is true that we should now accept the Millian Principle, is it not also true that five hundred years ago it was already the case that *if* a future society were to reach our stage of moral and political development, it would be right for them to adopt this standard? I believe it follows only in the trivial sense that it follows from the occurrence of a sea fight on Tuesday that on Sunday it was already true that there would be a sea fight on Tuesday.⁵ The truth of the judgment depends on reasons that people can have, and recognize, only in the circumstances, external and internal, that obtain now. In a sense, those reasons did not exist five hundred years ago, and would not exist even for someone to whom the present circumstances were described in detail. He would have to be transformed and transported into those circumstances to be capable of understanding, and therefore recognizing them. In other words, I am denying that the judgments that can be made in those actual circumstances are the manifestation

⁴ Williams (1998); p. 384 in Williams (2014).

⁵ See Aristotle, *De Interpretatione*, chap. 9.

of a completely general practical reason that could be applied with the same result to the hypothetical description of those circumstances, by any rational being, from any point of view – in a way analogous to the way the utilitarian standard could be applied to any describable circumstances from any point of view, however far away.

Some of the most important forms of moral progress consist in the acknowledgment and implementation of moral reasons that have existed for a long time – the righting of ancient wrongs. But I believe we can also look forward to progress of the other kind, which depends on the appearance of reasons that did not exist before, but that emerge through the development of human and social possibilities and the process of moral reflection in historical time. It seems clear that as we try to deal with problems of global equity and governance, modification of the human genome, or new possibilities for the extension of the human lifespan, our efforts will have to take this form, even though the appeal to timeless moral truths will never become obsolete.

References

- Aristotle, *De Interpretatione*, in: J. L. Ackrill (ed.), *Aristotle, Categories and De Interpretatione: Translated with Notes and Glossary*. Oxford: Clarendon Press, 1963.
- Mill, John Stuart (1859) *On Liberty*, in: J. M. Robson (ed.), *The Collected Works of John Stuart Mill, Volume XVIII – Essays on Politics and Society Part I*, Toronto: University of Toronto Press, London: Routledge and Kegan Paul, 1977.
- Raz, Joseph (1994) “Moral Change and Social Relativism,” *Social Philosophy and Policy*, No. 1.
- Scanlon, T. M. (1972) “A Theory of Freedom of Expression,” *Philosophy & Public Affairs* I, No. 2, reprinted in Scanlon (2003).
- Scanlon, T. M. (1979) “Freedom of Expression and Categories of Expression,” *University of Pittsburgh Law Review* 40, no. 4, reprinted in Scanlon (2003).
- Scanlon, T. M. (2003) *The Difficulty of Tolerance: Essays in Political Philosophy* (Cambridge University Press).
- Williams, Bernard (1998) “Review of The Last Word by Thomas Nagel,” *New York Review of Books*; reprinted in Williams (2014).
- Williams, Bernard (2014) *Essays and Reviews 1959–2002* (Princeton University Press).

Serena Olsaretti

Equality of Opportunity and Justified Inequalities: How the Family Can Be on Equality's Side

Introduction

Equality of opportunity – the principle that requires that jobs and other positions of advantage be awarded to those who are best qualified for them and regardless of their starting point in the social system – seems to sit awkwardly within liberal egalitarianism. Insofar as liberal egalitarianism rejects the importance of rewarding merit or desert, equality of opportunity may appear, at best, to be nothing but an expression of a demand of efficiency rather than justice;¹ at worst, it may turn out to be an unstable ideal that pushes liberal egalitarians towards equality of outcomes. After all, if “[w]e do not deserve our place in the distribution of native endowments, any more than we deserve our initial place in society (...)” (Rawls 1999: 89) then it should be unjust to permit unequal access to jobs and other positions on the grounds of unequal qualifications, given that what qualifications individuals have must necessarily reflect their unequal natural endowments, even if we neutralised the influence of social factors.² At the same time, equality of opportunity will always be frustrated by, and will always militate against, a canonical liberty to which liberal egalitarianism is committed, the freedom to have and raise children, because differences in family values and resources will invariably affect what qualifications children come to acquire. Liberal egalitarianism thus faces two sources of tension in its endorsement of equal opportunity.

Acknowledgment: This paper was presented at the Lauener symposium in honour of Thomas Scanlon in Bern, September 2016; at the UPF Family Justice workshop in Barcelona, February 2017, and at the Stockholm Institute for Futures Studies Internal Seminar in June 2017. I am grateful to participants of those events for their comments and to Paul Bou-Habib and Andrew Williams for discussions. Work for this article has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 Research and Innovation Programme (Grant Agreement Number: 648610; Grant Acronym: Family Justice).

¹ For this worry, see Daniels 1978.

² See Williams 1962.

<https://doi.org/10.1515/9783110733679-009>

Thomas Scanlon's discussion of equality of opportunity sheds much light on how liberal egalitarianism can handle the first of these sources of tension. Equality of opportunity, Scanlon believes, is part of a justification of socio-economic inequalities that appeals to the value of harnessing the talents and skills of individuals in a way that benefits everyone, especially the least advantaged members of society; for this reason, equality of opportunity demands that positions of advantage be assigned to those whose talents and skills can be used in ways that benefit the least well off, not that those talents and skills have no impact on how well off people end up, as equality of outcomes does. But this demand is not merely a demand of efficiency, since equality of opportunity also requires that no one have reason for complaint for lacking the talents and skills that give people access to positions of greater advantage. Scanlon's treatment of equality of opportunity thus offers an understanding of how to interpret this ideal in a way that duly recognises the need to justify socio-economic inequalities to the least advantaged members of society.

Scanlon's discussion also attempts to defuse the worries raised by the second source of tension surrounding equality of opportunity's place within liberal egalitarianism, regarding the conflict between equal opportunity and the family,³ and specifically, parental partiality, that is, parents' disposition to confer benefits on their children. Here, the main point Scanlon makes is that it is often possible (without sacrificing efficiency unduly) to undo the link between, on the one hand, many of the benefits that wealthier families confer on their children, and on the other, access to positions of advantage. Cases in which this is possible include those in which the benefits that wealthier families confer on their children are not, strictly speaking, needed qualifications for the positions of advantage that these children will compete for. In these cases, the competition for positions should be set up so that it does *not* adjudicate between candidates on the basis of whether candidates possess these benefits. For example, children of wealthier families gain enriching experiences by travelling abroad, doing volunteer work, and learning foreign languages which make them more interesting job candidates, rather than ones who are better able to perform the job well. Competitions for jobs should not choose between candidates on these grounds. Furthermore, we can undo the link between parental conferral of benefits and better access to positions of advantage in a second type of cases, namely, those in

3 By "the family", throughout, I mean parent-child relationships characterised by the fact that the adults who are entrusted with the role of parents have substantial rights and duties over, and vis-à-vis, the child/children they are the parents of. I also assume that some parental partiality – broadly, a disposition to favour one's children – is among the things these rights and duties protect or mandate.

which, although what wealthier parents confer on their children *are* needed qualifications for certain positions of advantage, it is possible to structure those positions and the competition for them so that those qualifications are provided on the job for all successful candidates, rather than being presupposed. Competence in a certain computer programming language that now can only be obtained through expensive programmes, for example, could be provided to all successful job candidates who have been selected on the basis of their knowledge of other, more basic, programming languages. In all these cases, undoing the link between parents' benefiting of their children and access to positions of advantage is what equality of opportunity requires. This way, heeding the demands of equality of opportunity does not require preventing parents from benefiting their children, and some of the tension between equality of opportunity and the family is defused.

This paper does not take issue with Scanlon's understanding of equality of opportunity and its role in the justification of inequality, nor with his suggestion of how in certain cases we can reconcile respect of parental partiality and equality of opportunity. Instead, taking Scanlon's discussion as a starting point, this paper examines a different, hitherto neglected, aspect of the relationship between equal opportunity, the justification of socio-economic inequalities, and parental partiality. My central claim is that once we acknowledge that in a just society equality of opportunity must be satisfied for inequalities to be justified, and given a certain view of the special obligations of parents to their children, appealing to the latter can ground an argument for *reducing* socio-economic inequalities. In a just society, parents may not support institutions that create substantial socio-economic inequalities while also *both* upholding equality of opportunity *and* heeding their parental obligations. So, a society whose institutions are regulated by equality of opportunity, and in which parents act in line with their parental obligations, is one which will, other things equal, be more egalitarian than a society whose parents are neglectful of their parental obligations.

The discussion proceeds as follows. Section 1 presents in a bit more detail Scanlon's discussion of the relation between equal opportunity and the justification of socio-economic inequalities. Section 2 briefly sketches an egalitarian critique of the argument for inequality-generating incentives which that justification supports, formulated by G. A. Cohen. Doing so is helpful as a preliminary to my own analysis. Section 3 presents my central claims about the reconciliation of parental partiality and socio-economic equality in a just society, and section 4 concludes by teasing out some central ways in which those claims are significant.

1 Scanlon on equality of opportunity and the justification of inequalities

As mentioned earlier, Scanlon conceives of the ideal of equality of opportunity as part of a justification of socio-economic inequalities, or as part of an answer to a complaint against inequality.⁴ In this respect, as well as in terms of the other components of the justification of inequality it endorses, Scanlon's view shares much with John Rawls's. Socio-economic inequalities can be just only if it is true of them both that they benefit the least well-off members of society and that the offices and positions to which they are attached are open to all.⁵ More specifically, the justification of such inequalities must invoke two principles:

i) *An equality of opportunity* principle, which regulates *access* to jobs and other positions of advantage. These include offices, jobs, and higher education positions but also other inequality-generating institutions like rules regulating the set-up of limited liability corporations or the acquisition of patents. The principle requires, in Scanlon's terms, both that positions of advantage be allocated through fair procedures, on the basis of relevant qualifications – i. e., that “procedural fairness” be respected – and that everyone have a fair chance of acquiring those qualifications, so that no one can complain that she was not placed in a sufficiently good position to acquire the relevant qualifications – this is a demand of “substantive fairness”.

ii) *The Difference Principle*, which offers a justification of inequality-generating institutions by reference to the beneficial consequences of having institutions that generate these inequalities for the worse-off members of society. We can characterise this principle as one that concerns not access but the *payoffs* of jobs. The Difference Principle is forward-looking, identifies the interests of a particular social group – the least advantaged – as relevant, and eschews appeal to notions of desert and entitlement. If attaching unequal financial rewards to positions of advantage were not necessary to benefit the worse off, then inequalities in economic rewards would be unjust: the putative facts that the best qualified would be deserving of greater rewards, or that they are entitled to them if

⁴ See Rawls 1999; Scanlon 2018.

⁵ These are only necessary, not sufficient, conditions. Liberal egalitarians also endorse a principle of equal basic liberties as demanded by justice. On Scanlon's view, there may be further objections to inequalities that are not captured by these principles, for example if they are generated in ways that fail to give due considerations to individuals or if they are stigmatising. Throughout, I focus only on the two principles I have just mentioned and assume throughout that other demands of justice are respected.

others are willing to offer them their justly held resources by way of reward, are irrelevant. Instead, the Difference Principle justifies those inequalities in the pay-offs of jobs and other positions of advantage that make the least advantaged members of society better off than they would be if these inequalities were eliminated or reduced.

For Scanlon, as for Rawls, meeting the demands of either of these principles on its own would not be enough for inequalities to be just. A society in which socio-economic inequalities were permitted so as to please the best off (so the creation of the inequalities lacks a justification) would not be just, even if perfectly fair access were guaranteed to the best-off positions. Similarly, a society in which institutions were so structured as to ensure that socio-economic inequalities benefited the worse-off members of society, but in which access to the unequally rewarded positions were assigned nepotistically, would be unjust even if (as is unlikely) the nepotistic arrangements were not inefficient.⁶

The fact that equality of opportunity is a necessary part of a justification of inequality matters for my argument in what follows, since the question I examine is what implications the fact that equal opportunity must be satisfied has for which inequalities are justified. This question has not been raised in critical analyses of the Rawlsian view of justified inequalities; nor does Scanlon himself address it, although he, more explicitly than Rawls, makes perspicuous that the point and purpose of equal opportunity is to help silence objections to socio-economic inequalities and hence, that inequalities that arise against a background in which equality of opportunity is not realised are unjustified. When we focus on this fact, I argue, a hitherto unnoticed aspect of parental partiality comes to light: in a just society, that is, one that is regulated (at least to a substantial degree) by fair equality of opportunity, parents' fulfilling their obligation to their children requires that parents support fewer inequalities than those which, in the absence of considerations about what parents owe their children, can be justified by the Difference Principle. In other words, in a society regulated by the principles mentioned above, parental partiality serves as a *constraint* on which inequalities can be justified. If this argument is right, there is at least one respect in which the family can be seen to be aligned, rather than at odds, with the demand for the reduction of socio-economic inequality.

⁶ In his discussion of equality of opportunity, Scanlon brings out, more than Rawls does, the internal connection between the two principles, showing that the rationale for the inequality-engendering institutions must inform, at least in part, what the demands of equality of opportunity are. But even on his view, equality of opportunity also expresses independent concerns (e.g. with due and with equal consideration, with non-discrimination, and with substantive fairness).

As background to this argument, it is helpful to begin by noting that the Difference Principle provides support for inequalities in income and wealth insofar as these are necessary incentive-payments, that is, insofar as inequality-generating rewards provide those who have productive talents (henceforth, “the talented”) with incentives to use those talents productively.

According to this “incentives argument”, which Rawls’ Difference Principle provides normative support for (Rawls 1999: 69), and which Scanlon’s view presumably also allows for, inequalities created by the high rewards that are necessary to motivate the talented to take up productive occupations, and to work productively in those occupations, are just in that they help to maximally benefit the worse-off members of society. The incentives argument goes as follows:

1. Inequalities that are necessary to benefit the worse-off are just;⁷
2. The inequality-creating incentive payments demanded by the talented members of society are necessary to benefit the worse off;

Therefore,

3. The inequality-creating incentive payments demanded by the talented are just.

In what follows I suggest that, once we keep in view what Scanlon emphasizes, namely, that any justification of inequalities – and hence, this incentives argument, too – assumes that equality of opportunity is satisfied, then it is not an argument that could be proffered by talented individuals who are parents and who heed their parental obligations. So, even if these inequalities could be declared to be just by the lights of the Difference Principle, parental partiality – by which I mean, specifically, parents’ special obligations to benefit their children in certain ways – would condemn them. Before I formulate these points, in the next section I sketch an egalitarian critique of the incentives argument, which denies that the Difference Principle, correctly understood, sanctions incentives inequalities. Doing so, as will become apparent, helps bring to view how my argument differs from, and can contribute to, existing liberal egalitarian treatments of which inequalities are justified.

⁷ Assuming that other principles of justice, including the principle of equality of opportunity, are also respected.

2 G. A. Cohen's critique of the incentives argument

The incentives argument for inequalities has been the target of the egalitarian critique of John Rawls formulated by G. A. Cohen (2008). The core element of Cohen's critique is the contention that talented individuals in a just society may not, compatibly with respecting justice (and in particular, the spirit behind the Difference Principle, which Cohen identifies with ideals of community, fraternity and civic friendship), be self-seeking income maximisers in their economic choices (the choices of what occupation to work in, and how hard to work) in the way the incentives argument assumes, leaving only institutions to do work in realising justice. Contrary to Rawls and Scanlon, Cohen believes that some inequalities produced by institutions that satisfy the Difference Principle would be unjust, since justice requires that citizens be motivated by an egalitarian ethos to make productivity-promoting occupational choices without compromising socio-economic equality. More specifically, Cohen denies that in a just society the inequality-creating incentive payments demanded by productive individuals are just, because in his view, two possible interpretations of premises 1 and 2 in the argument sketched just above must be disambiguated, and, once these are disambiguated, it can be argued that the defensible interpretation of premises 1 and 2 are not ones which can warrant the conclusion expressed in 3. Let me explain.

Cohen notes that incentive payments can be “necessary” to induce the talented to work productively in two importantly different senses: in an *intention-dependent* and an *intention-independent* sense. An incentive payment is necessary in an *intention-dependent* sense if it is needed to induce the talented to work productively, where working productively is in the talented person's *power* to do (i.e. it is something they are capable of doing if they so choose), but she is *unwilling* to do it, without the incentive payments in question. If the talented could, by contrast, not *bring themselves* to work productively even if they so chose, then the incentives that they receive in order to *enable* them to work are necessary in an *intention-independent* sense.⁸

Now, according to Cohen, when they are correctly understood, the demands of the Difference Principle (and of the related principles of fraternity and of com-

⁸ Incentives that are necessary in this sense, Cohen points out, are not, strictly speaking inequality-engendering, in that they are merely *compensatory*, i.e. they offset a disadvantage. An illustration is a monetary reward that is necessary to buy oneself a holiday one *needs* to be able to rest enough to remain productive.

munity) are such that only inequalities that are necessary in an intention-independent sense can be just. But inequality-generating incentives are necessary only in an intention-dependent sense. Conclusion 3, then, does not follow from a duly revised formulation of premises 1 and 2.⁹ According to Cohen, in a just society, one whose institutions are regulated by the Difference Principle, and whose citizens accept that principle, the talented members of society would not hold out for incentive payments only because they can and in order to maximise their income within the limits permitted by just institutions. Instead, just citizens would be guided by the Difference Principle in their personal choices, and more specifically, in their choices of which occupation to pursue and how hard to work, and talented citizens would be motivated by an “egalitarian ethos” to work productively without demanding inequality-generating incentive payments – by accepting, for example, higher redistributive taxation taxes while still working at their productive jobs. The worse off would thereby benefit *more* relative to the inequality-producing incentives scenario, while at the same time, fewer or no inequalities would obtain.¹⁰

Given my interest in this paper, it bears mentioning at this point that Cohen’s critique of the incentives argument assumes that the inequalities produced by the Difference Principle as it applies to institutions would be compatible with the other principles of justice, including the fair equality of opportunity principle: Cohen does not take issue with the claim that such inequalities can arise

9 Cohen’s argument could then be summed as follows (the underlined words indicate the changes to the original incentives argument):

1*. Inequalities that are necessary, only in an intention-independent sense, to benefit the worse-off are just;

2*. The inequality-creating incentive payments demanded by the talented members of society are necessary to benefit the worse off only in an intention-dependent sense;

Therefore,

3*. The inequality-creating incentive payments demanded by the talented are not just.

10 The following table illustrates, in a simplified form, the payoffs structures that would be just according to the incentives argument (B) and according to Cohen’s egalitarian ethos argument (C). A represents the egalitarian baseline relative to which B constitutes an improvement such that the inequality in B is (allegedly) just.

	Worse-off	Talented
A – Egalitarian baseline	10	10
B – Incentives inequality	12	20
C – Egalitarian ethos payoffs	16	16

compatibly with the principle of fair equality of opportunity's being respected. That claim, by contrast, is part of the argument I develop.

As already announced, in what follows I participate in the discussion over the merits of the justification of socio-economic inequalities provided by the incentives argument from a particular angle, that is, given a concern with assessing the rights and obligations of a particular group in a just society, namely, *parents* – by which I mean adults who are the primary carers of children. Thus, while my argument is, like Cohen's, an argument *for equality* (or at least, against inequality), it is different from Cohen's insofar as it grounds the egalitarian case in the special obligations of parents to their children, rather than in general obligations of justice which citizens owe to one another. My argument is also different from Cohen's in other respects, some of which see my view as better aligned with the Rawls-Scanlon liberal egalitarian view than with Cohen's view, and others which see my argument as opposing some convictions which the views of Cohen, Rawls and Scanlon all share despite their differences. I will return to these points in section 4, after presenting the argument from parental partiality in favour of socio-economic equality in the next section.

3 Talented parents and the limited scope of the incentives argument for inequality

Whatever the merits of Cohen's critique of the incentives argument, the latter, I now suggest, fails as a *general* justification of incentive-generated inequalities, that is, as an argument that is supposed to apply to a class of inequalities – incentive inequalities – created by *all productive members of society alike, parents as well as non-parents*. More specifically, I argue that parents may not demand inequality-creating incentive payments while also *both* upholding equality of opportunity *and* heeding their parental obligations, that is, a set of special obligations which they have, as parents, to their children, in a just society.¹¹ Two key sets of premises are needed to support this conclusion. They concern, respectively, certain assumptions about social facts which all just citizens must take as valid given the operation, in a just society, of the equal opportunity principle;

¹¹ In his discussion of equality of opportunity, Joseph Fishkin helpfully brings out that, under certain conditions (e.g. when access to positions of advantage is very competitive), families will have greater incentives to act in ways that exacerbate inequalities (Fishkin 2014, pp. 127–8). Note that my claim is different from Fishkin's, in that it is primarily a claim about *the obligations* parents have, not about what they are motivated to do.

and a view about what obligations parents have towards their children in a just society.

Consider, first, what all citizens of a just society must assume regarding the implications of a principle of equal opportunity's being in operation. That that principle is in fact in operation is, recall, a condition that must be satisfied for the inequalities justified by the difference principle to be just. The incentives argument sketched earlier is one that holds against the assumption that the equality of opportunity principle is satisfied, at least to a substantial degree. This much is uncontroversial; it is an implication of Rawls' according lexical priority to fair equality of opportunity over the Difference Principle, and is something that, as was mentioned earlier, Scanlon's treatment of equal opportunity's role in justifying socio-economic inequalities makes perspicuous. What should also be uncontroversial, and, I now suggest, has implications that are typically not noticed, is the following social fact which citizens in a just society should assume holds true in their society: in a society whose institutions are regulated by the equality of opportunity principle, *there is a considerable degree of social mobility, so that the social class one is born into is not a reliable predictor of what social class one will belong to as an adult*. Unless this social fact obtains, we cannot say that equality of opportunity in *any sense that is meaningful for liberal egalitarians* is satisfied. If this social fact does not obtain, at best only a formal principle of equality of opportunity, which requires that people who hold the qualifications for competing for positions of advantage not be excluded from competing for them on irrelevant grounds, is realised. Liberal egalitarians like Scanlon and Rawls, however, clearly believe that equality of opportunity demands more than this. As they note, it requires that those who have equal motivation and equal ability (or could have equal ability if provided with the right conditions) have roughly "the same prospect of success". It seems safe to say that, unless a substantial degree of social mobility obtains, nothing close to this ideal is approximated, let alone fully satisfied.¹²

Since much has been written on the thorny questions of exactly *what constraints* on individual parental action the principle of equal opportunity imposes (Is reading bedtime stories compatible with respecting equality of opportunity?

12 There is a scenario in which at least Rawls' understanding of fair equality of opportunity could in principle be realised compatibly with there being little social mobility, i.e. one in which natural talents are closely correlated with social class. (Whether Scanlon's version of the ideal would allow for this possibility is less clear, given that Scanlon denies an assumption Rawls makes, that is, that there is a pre-institutional understanding of "natural talents".) I assume this is not the case. It is an interesting question what sort of value equal opportunity would have if this were, indeed, the case, but I do not address that question here.

Is sending one's children to after-school activities?), and what *public policy measures* must be put in place in order to implement this ideal (Is high quality, equal publicly funded primary education for all required? Is that all that is required?), it is important to note that here I am not attempting to give an answer to those questions. Instead, I am suggesting that a substantial degree of social mobility is a social fact that we must assume would be the *outcome* of (more than formal) equality of opportunity's being in place, whatever constraints on individual actions and whichever policy measures we believe equality of opportunity might require.

By way of analogy, consider the claim that, whatever gender equality might require in terms of the division in heterosexual households of unpaid domestic and paid non-domestic work between male and female partners, we can confidently claim that unless *statistical equal split* obtains (such that taking the aggregate of men in such households and women in such households, we find that overall they share equally in domestic and in non-domestic labour), a society is not gender egalitarian.¹³ We can say that a substantial degree of social mobility and statistical equal split are the macro-level, society-wide outcomes that must obtain if certain phenomena condemned by (a range of defensible interpretations of) certain ideals do not occur. Such phenomena would be, respectively, systemic failures of a more than formal principle of equal opportunity; and widespread gender injustice, such that it would generally be possible to predict, from an individual's sex, where he or she falls in the distribution of paid and unpaid work.

These social facts, then, can be taken as indicators of certain principles' being respected, and insofar as citizens in a just society make choices whose justice is conditional upon certain principles' being realised in their society, and must be informed by their knowledge that those principles are so realised, these social facts are ones which just citizens must assume obtain when making those choices. In a society in which equality of opportunity is in operation, and in which people's prospects in the allocation of positions of advantage are not (at least, not largely) affected by their social class, there will be a substantial degree of social mobility, both upwards and downwards. Parents in a just society, then, would be wrong to assume that their and other people's children will grow up to be in their own social class. For parents who are talented individuals and could thus, in principle, command incentive payments, this means that they

¹³ Richard Arneson defends this claim (1997), as an alternative to Susan Moller Okin's view that gender justice requires, specifically, that there be an equal split between men and women *within individual households*.

would be wrong to assume that their children will be among the better-off, rather than the worse-off. This, I now suggest, has important consequences for what socio-economic inequalities these parents may demand, *compatibly with* their fulfilling their obligations to their children. In a just society, talented parents could only render inequality-engendering incentives necessary (thus making premise 2 of the incentives argument true) if they failed to respect special obligations which they have as parents, namely, the obligation to their children to not gain at their expense, and the obligation to help ensure for them the circumstances in which they will be comfortably well-off as fully cooperating members of their society.

The first of these two obligations bears resemblance to Rawls' principle of fraternity: parents' having an obligation not to gain benefits at their children's expense involves their refraining from rendering social positions which their children could come to occupy less good than they could be if the parents worked as productively without seeking to benefit by demanding incentive payments. Whether or not we agree with Rawls that this obligation binds citizens to one another, it seems plausible that it is one that parents have to their children.¹⁴ Parents' obligations to their children extend beyond that. It is important, here, that the obligation to help ensure for one's children the conditions for them to be comfortably well-off is an obligation to help secure for one's children the option of being well-off *by being fully cooperating members* of their society – that is, by being self-supporting and by participating in the social role which they choose to participate in and which, compatibly with background constraints (e.g. regarding the supply of jobs) and with the principle of equality of opportunity's being respected, they will be able to participate in.¹⁵ It is in virtue of having this obligation (as opposed to merely an obligation to supply for their children a certain level of material well-being) that parents would be failing their children if they intentionally undermined their children's employment chances, or worsened their children's prospective employment conditions.

14 My discussion here makes two important assumptions I do not examine. The first is that parents have special obligations to their children, as opposed to having only general obligations towards them, obligations which other citizens also have. There are several different accounts of what grounds such responsibilities, which I do not take a stance on here. (See O'Neill 1979, Brake 2010, Archard 2010, Porter 2014, Olsaretti 2017 for some discussions.) The second assumption I make is that some of these obligations are owed to children as future adults, or once they are adults. Some special obligations parents have to their children do not cease once their children are of age.

15 The obligation at hand may share its basis with parent's obligation – which more readily springs to mind – to help raise his children so they acquire the capacities to be free and equal members of society. I do not explore this point here.

They would fail their children – it is important to note this – even if they were both willing and able to secure their children’s material well-being by lavishing them with generous gifts, and even if they undermined their children’s employment chances, or worsened their children’s prospective employment conditions, *in order to* secure for themselves material resources with which they will improve their children’s material well-being *to a greater extent* than if they did not undermine their children’s employment opportunities. This would be so whether parents’ undermining of their children’s employment opportunities occurs through individual “private” acts, or *via* institutions. So, for example, a parent who, knowing he is better qualified than his child at a job and would command a higher salary than her (on grounds of seniority, say), applies for that job in order to outcompete his child and get the higher salary, violates the obligation in question, *even if* he were willing to share his higher salary with his child. Similarly, to take another example, parents who vote for labour market policies that are reliably predicted to have, as an effect, a reduction or worsening of the conditions of youth employment, may also be failing in this obligation – again, even if they are able and willing to more than compensate their children monetarily.

If talented parents are unwilling to work productively unless they receive inequality-generating incentives (for example, by not working productively under a higher income tax regime), thereby not improving the worse off social positions as much as they could if they accepted higher redistributive taxation, then, given that their children may be among the worse-off and come to occupy those social positions, they would violate an obligation to their children to help ensure for them the circumstances in which they can be comfortably well-off as fully cooperating members of their society. So, in a society regulated by fair equality of opportunity, parents may not justify inequality-generating incentive payments by appeal to the liberal egalitarian argument, compatibly with respecting their parental obligations.

If parents who are just citizens in a just society can only demand inequality-engendering incentive payments at the price of failing in their parental obligations, then, since I assume that parents who are just citizens are also good parents (that is, parents who comply with their parental obligations), the socio-economic inequalities that will be demanded in a just society many of whose talented members are parents will be quite limited. Talented parents who recognise both their obligations as just citizens and as parents would, in just circumstances, be motivated to both abstain from demanding inequality-engendering incentive payments, *and* to nonetheless work (nearly) as productively as they would if they received incentive payments. They would make economic choices that benefit the worse-off members of their society, *without* demanding incentive payments (or demanding only limited incentive payments) thereby contributing

to there being only limited or no inequalities and benefiting the worse-off *more* than would be possible if they, the parents, did demand those payments as a condition for working productively.¹⁶ Parents would, in other words, accept greater sacrifices for the sake of helping the worse-off than if they did not have, or failed to act on, their obligation to benefit their children.¹⁷ They would do this because they acknowledge, as they must if they are just citizens who accept that their institutions are duly regulated by the fair equality of opportunity principle, and who therefore assume the social fact of social mobility, that the worse-off members of society could include their own children, and that they have an obligation to their children to uphold just institutions which render all social positions, including those of the worse-off, ones which they could want their children to occupy.

4 The public-private distinction, and the role of parental partiality in egalitarian theory and practice

Scanlon's treatment of equality of opportunity sees this ideal as part of a justification of socio-economic inequalities: socio-economic inequalities are only just if equality of opportunity is satisfied. Taking this claim as a starting point for my discussion, I have suggested that it can be deployed as part of an unfamiliar argument for socio-economic equality, an argument that appeals to the obligations of parents to their children in a just society. If equal opportunity real-

16 I believe that there are other concerns that parents (reasonably) have for their children, and which have implications for the kind of society's employment structure which parents would want to help establish and maintain in a just society. In particular, I have in mind parents' concern with ensuring that their children achieve some self-realization in the jobs and positions they occupy; parents would, accordingly, support institutions that restructure jobs in order to render possible the pursuit of a substantial degree of self-realisation even if one occupies less "prestigious" occupations.

17 We can use the simple table used earlier to illustrate Cohen's egalitarian ethos view: just parents would opt for C if they recognised their parental obligations.

	Worse-off	Talented
A – Egalitarian baseline	10	10
B – Incentives inequality	12	20
C – Just parents	16	16

ly must be satisfied (at least to a substantial degree) for socio-economic inequalities to be just, then well-off parents have reason to reduce socio-economic inequalities and ensure that all social positions, including the worse-off ones, are reasonably good, since their children could come to occupy those positions. So, viewing the fair equality of opportunity principle as Scanlon does, as part of a justification of inequalities, can have some generally neglected inequality-constraining implications. The fact that the fair equality of opportunity principle must be in operation for any socio-economic inequalities sanctioned by the Difference Principle to be just has implications, then, not only for how *access* to positions of advantage should be regulated, but also – and this is what I have focused on in this paper – for what *the payoffs* of those positions of advantage should be. This point has been missed in discussions of the incentives argument, which have focused only on the Difference Principle and neglected to ask what implications the fact that lexically prior principles of justice are satisfied has on which socio-economic inequalities can be just. By way of conclusion, I would like to highlight two other ways in which the argument I have sketched relates and contributes to existing positions on the liberal egalitarian justification of socio-economic inequalities and on the family.

First, as I already remarked, while my argument supports the same implications as those drawn by Cohen, it does so on different grounds, and unlike Cohen's view, it does not rest on rejecting the moral division of labour between institutions and personal behaviour championed by defenders of the liberal egalitarian argument for inequality. According to the liberal egalitarian view that Cohen attacks, there is, in a just society, a division of labour between institutions, which enact justice, and personal behaviour, which does not itself aim at realizing justice.¹⁸ Cohen believes, by contrast, that the very same principle that regulates institutions – the Difference Principle – should also constrain people's personal choices. To be sure, the demands of the egalitarian ethos would only extend as far as is compatible with the independent requirement to respect an agent-centred prerogative, which Cohen believes exists, and which protects an area in which each individual is at liberty to pursue his self-interest, broadly understood (thus understood, self-interest includes parents' interest in their child-

18 In other words, just citizens uphold just institutions, and accept that their behaviour, in order to be legitimate, must not violate the constraints imposed on it by those just institutions, but, when making choices in their personal life which are permitted by the institutions they live under, they need not ask themselves which of those choices best promote justice. I have argued against this understanding of the moral division of labour, which both Cohen and his targets display, in "The Inseparability of the Personal and the Political", *Analysis*, but leave this point aside here.

ren's well-being).¹⁹ But on Cohen's view, personal economic choices, as well as a society's institutions, would have to be regulated by the Difference Principle in order for justice to be realised.

The argument I have sketched for a society of limited socio-economic inequalities, by contrast, is compatible with endorsing the moral division of labour endorsed by liberal egalitarians, since it supports conclusions that overlap with Cohen's, for equality, for reasons that are very different from Cohen's. These reasons appeal to the demands of parental obligations in a just society, rather than the need to extend the reach of demands of justice into personal life. We could thus agree with Cohen that we should aim to have more socio-economic equality than some liberal egalitarians have allowed for, but also agree with Rawls and Scanlon, against Cohen, that we should not demand that people's economic choices be guided by an egalitarian ethos.

A second, related point highlights a respect in which the argument I have sketched differs from both Cohen's and liberal egalitarians' views. On Cohen's view as much as on the view of the liberal egalitarians he criticises, the family is seen as primarily an equality-disrupting force. Mention of parents and the family in the context of discussing both equality of opportunity and socio-economic inequalities typically points to the various respects in which the family and parental partiality pull away from equality, both equality of opportunity and the maintenance of socio-economic equality. There is, to be sure, an important difference between Cohen's view and the liberal egalitarian accounts he criticises when it comes to the family's disruption of equality. According to the Rawls-Scanlon argument for socio-economic inequalities, parents' disposition to favour their children leads to inequalities that are sanctioned as *just*, whereas for Cohen, it would be protected by the agent-centred prerogative as *legitimate departures from justice*. But the possibility I have canvassed in this paper is altogether different, and has brought to view one way in which parental concern can be aligned with justice, and push towards greater socio-economic equality, rather than being viewed as necessarily pulling away from justice and equality.

¹⁹ So, Cohen does, in fact, permit *some* inequalities that are not necessary in an intention-independent sense to benefit the worse-off. He believes these would not be very extensive (for an argument to the contrary, see Estlund 1998, to which Cohen responds in his General Appendix of *Rescuing Justice and Equality*). In any event, the inequalities in question would not, as I understand Cohen, be ones that are either *required* or even just *permitted by justice*. Instead, they are unjust, but, on balance, justified inequalities.

References

- Archard, D., 2010. The Obligations and Responsibilities of Parenthood. In D. Archard and D. Benatar (eds.). *Parenthood and Procreation*. Oxford: Oxford University Press
- Brake, E., 2010. Willing Parents: A Voluntarist Account of Parental Role Obligations. In D. Archard and D. Benatar (eds.). 2010. *Parenthood and Procreation*. Oxford: Oxford University Press
- Cohen, G. A., 2008. *Rescuing Justice and Equality*. Cambridge, Mass.: Harvard University Press
- Daniels, N., 1978. “Merit and Meritocracy”, *Philosophy and Public Affairs* 7: 206–223
- Estlund, D., 1998. “Liberalism, Equality and Fraternity in Cohen’s Critique of Rawls”, *Journal of Political Philosophy* 6, 1: 99–112
- Olsaretti, S., 2017. “Liberal Equality and the Moral Status of the Parent-Child Relationship”. *Oxford Studies in Political Philosophy*, vol. 3. Oxford: Oxford University Press
- Olsaretti, S., 2011. “The Inseparability of the Personal and the Political: A Review of G. A. Cohen’s *Rescuing Justice and Equality*.” *Analysis* 72: 147–156
- Rawls, J., 1999. *A Theory of Justice*, Oxford: Oxford University Press (revised edition)
- O’Neill, O., 1979. Begetting, Bearing, and Rearing. In O. O’Neill, and W. Ruddick (eds.). *Having Children: Philosophical and Legal Reflections on Parenthood*. New York: Oxford University Press
- Porter, L., 2014. Why and How to Prefer a Causal Account of Parenthood. *Journal of Social Philosophy* 45, 2: 182–202
- Scanlon, T., 2018. *Why Does Inequality Matter?* Oxford: Oxford University Press
- Williams, B., 1962. “The Idea of Equality”. In P. Laslett and W. G. Runciman (eds.). *Philosophy, Politics and Society, Second Series*. Oxford: Blackwell.

Derek Parfit

Improving Scanlon's Contractualism

Scanlon's *What We Owe to Each Other* is, I believe, the best book on ethics published in the 20th Century. Scanlon proposes and defends one of the two best versions of contractualism, the other being a revised Kantian theory. Of Scanlon's several other achievements, the greatest seem to me the ways in which he has shown that the most fundamental normative truths are not about what is right or wrong but about normative reasons. Things go worse, for example, if they go in ways in which we all have stronger impartial reasons to want things to go, and morality matters because we have reasons not to act wrongly.

I shall now discuss some ways in which Scanlon could revise and extend his version of contractualism.

In *What We Owe to Each Other*, Scanlon claimed that, rather than describing the facts that can *make* acts wrong, his theory gives an account of wrongness itself, or of *what it is* for some act to be wrong. This claim, I believe, was a mistake.¹ According to one statement of

Scanlon's Formula: An act is wrong just when such acts are disallowed by some principle that no one could reasonably reject.

If Scanlon was here using 'wrong' in a Contractualist sense, to mean 'disallowed by such an unrejectable principle', he could claim that his formula gives an account of *what it is* for acts to be wrong in this Contractualist sense. But *Scanlon's Formula* would then be a concealed tautology, which told us only that acts are wrong in this Contractualist sense just when they are wrong in this sense. We could all accept this trivial claim, whatever our moral beliefs. Scanlon's claim should instead be that, if some act is disallowed by a principle that no one could reasonably reject, this fact makes this act wrong in one or more other, non-contractualist senses. Scanlon might for example claim

Editors's note: Since Derek Parfit sadly died on January 1st, 2017, only a few months after the symposium, he was not able to revise his symposium talk for publication. We are very grateful to Tim Scanlon for supplying a title for Parfit's paper and making some edits of a non-substantive nature, including a few footnotes and a list of cited works at the end. We would also like to thank Derek Parfit's literary executor, Jeff McMahan, for granting us permission to publish this version of Parfit's talk from the 2016 Lauener Symposium in Bern.

¹ See Parfit (2011) 213–215.

SF*: When some act is wrong in this Contractualist sense, that makes this act wrong in the justifiabilist, blameworthiness, and reactive-attitude senses.

These four senses of ‘wrong’ are all abbreviations of longer phrases. So this version of *Scanlon’s Formula* could be more fully stated as

SF***: When some act is disallowed by some principle that no one could reasonably reject, this fact makes this act unjustifiable to others, blameworthy, and an act that gives its agent reasons for remorse and gives others reasons for indignation.

Scanlon could not claim that when some act is wrong in his Contractualist sense, that is *the same* as being, or is *what it is* for this act to be, wrong in these other senses. Being disallowed by a principle that could not be reasonably rejected couldn’t be *what it is* for an act to be blameworthy, or what it is for an act to give its agent reasons for remorse and others reasons for indignation. These are different properties. But being disallowed by such a principle might be one of the highest level wrong-making properties under which other wrong-making properties can be subsumed. Scanlon has shown that, as well as having intuitive beliefs about which acts would be wrong, we have intuitive beliefs about what would be reasonable grounds for rejecting some moral principle. This is one way in which Scanlon has developed and improved some of our moral thinking. Scanlon, I believe, may now accept that his Contractualist theory should take this second, substantive, non-reductive form.²

I turn next to some of Scanlon’s claims about what would be such reasonable grounds for rejecting moral principles. According to what we can call Scanlon’s

Impersonalist Restriction: In rejecting some moral principle, we cannot appeal to claims about the impersonal goodness or badness of outcomes.

In Scanlon’s words, “impersonal values are not themselves grounds for reasonable rejection.”³ All reasons for rejecting principles, Scanlon also claims, must

² Scanlon writes: “I should have avoided describing contractualism as an account of the property of moral wrongness. [...] The claim ... can be dropped from my account without affecting the other claims I make for contractualism.” Scanlon (2004) 137. He also writes: “The fact that an action would cause harm may make it reasonable to reject a principle that would permit that action, and thus make that action wrong in the contractualist sense I am describing. It is also true that an action’s being wrong in this sense makes it morally wrong in the ... general sense of that term ...” (ibid. 136). For a longer discussion, see Scanlon (2007).

³ Scanlon (1998) 222.

be *personal*. Of those who appeal only to personal reasons, some believe that there is no sense in which outcomes can be impersonally good or bad. That is not Scanlon's view. Scanlon believes both that outcomes can be impersonally good or bad, and that we can have strong reasons to try to produce or prevent such outcomes.⁴

Scanlon gives, as one example, reasons provided by the suffering of animals. He writes: "... like the pain of humans, the pain of non-human animals is something we have reason to prevent and relieve, and failing to respond to this reason is a moral fault."⁵ Scanlon then imagines someone saying: If there are impersonal reasons of this kind, why should they not count as possible grounds for reasonably rejecting principles? He replies: "The contractualist formula is meant to describe one category of moral ideas: the requirements of 'what we owe to each other'. Reasons for rejecting a principle thus correspond to particular forms of concern that we owe to other individuals. By definition, impersonal reasons do not represent forms of such concern."⁶

Since Scanlon himself defines his Contractualist sense of 'wrong', he is entitled to claim that, when we ask which acts are in this sense wrong, we should not appeal to impersonal reasons, since by definition such reasons are irrelevant. But on the substantive reading of *Scanlon's Formula*, when certain acts are wrong in Scanlon's Contractualist sense, that makes them wrong in other senses. Scanlon could not say that, when we ask which acts are wrong in these other senses, claims about what is impersonally worse in reason-implying senses are *by definition* irrelevant.

This version of Scanlon's view could take either of two forms. If Scanlon keeps his *Impersonalist Restriction*, he might claim that, when certain acts are wrong in his Contractualist sense, that would often make these acts wrong in these other senses, but such acts might be justified if they would prevent things from going much worse in the reason-implying sense.

It would be better, I believe, if Scanlon drops his *Impersonalist Restriction*. On this version of Scanlon's view, when we ask whether we could reasonably reject some principle, we are allowed to appeal to our beliefs about the goodness or badness of outcomes. Scanlon could then make the wider claim that acts are wrong in the other senses just when, and in part because, these acts are wrong in Scanlon's Contractualist sense. If that were true, Scanlon's Contractualism

⁴ Scanlon (1998) 182.

⁵ Scanlon (1998) 181.

⁶ Scanlon (1998) 219.

would unify, and help to explain all of the more particular facts that can make acts wrong. That gives Scanlon a strong reason to make this wider claim.

If Scanlon allowed us to reject some principles by appealing to our beliefs about the badness of outcomes, that would not make Scanlon's theory Consequentialist. Of the facts that can make outcomes impersonally worse, many are facts about various effects on people's well-being. Scanlon could claim that, in most cases, we would have no need to appeal to claims about the badness of these outcomes, since we can appeal directly to these effects on people's well-being.

Suppose that some act would give small benefits to many other people, in a way that would impose a great burden on me. On Scanlon's view, I could reasonably reject any principle that would allow this great burden to be imposed on me. None of the other people could reasonably reject any principle which condemned such acts, since none of these people could claim that his being denied the small benefit gave him a stronger ground for reasonable rejection than my great burden would give me.

In such cases, the great burden imposed on me might make the outcome worse. But I would not need to appeal to this fact. I might also reasonably reject any such principle even if, because the small benefits would come to very many people, my being burdened would make the outcome better. As these remarks suggest, if Scanlon allowed us to appeal to claims about impersonal reasons, and the badness of outcomes, that would make little difference to much of the moral reasoning that his Contractualism describes.

There is, however, at least one important exception. When Scanlon asks what we owe to others, he intends these *others* to include all future people. In his words: "... contractualism provides no reason for saying that people who do not now exist but will exist in the future have no moral claims on us."⁷ He also writes: "... a restriction to presently existing human beings seems obviously too narrow."⁸ When we consider how our acts may affect future people, we face some new questions. Suppose that we and others in our community are choosing between two energy policies. What I shall call the *selfish* policy would be somewhat better for us, but would greatly lower the quality of people's lives more than a century from now. This fact, I believe, would give us strong reasons not to choose this policy, and may make this policy wrong. We may assume that this great lowering in the quality of life would be worse for the people who would later live. But that may not be so. These people's lives would be worth liv-

7 Scanlon (1998) 187.

8 Scanlon (1998) 186.

ing, and given the facts about human reproduction, if we had chosen the other, better policy, these particular future people would never have existed. It would have been other, different people who would have later lived and had a higher quality of life. The selfish policy would then be worse for no one.

This policy would be wrong, *Scanlon's Formula* implies, if this policy would be disallowed by some principle that no one could reasonably reject. Could these future people reasonably reject any principle which permitted us to choose this selfish policy? If all grounds for reasonable rejection should appeal to personal reasons, the answer seems to be No. These future people could not reject any such permissive principle by appealing to the burdens that this policy imposed on them. There are no such burdens, since these people would have lives worth living, and they would owe their existence to our choice of this policy. Nor could we claim that this principle could be reasonably rejected by the people who would have existed, and had a higher quality of life, if we had chosen the better policy. If we choose the selfish policy, these people would never be actual. When we apply *Scanlon's Formula*, we cannot defensibly appeal to claims about what could be reasonably rejected by people who are merely possible, and never actual. Since we cannot appeal to the *personal* reasons that would be had by people who *never* exist, we should appeal to the *impartial* reasons that are had by people who *do* exist. On this version of Scanlon's view, we could reasonably reject any principle that would permit the selfish policy, and other such policies, because these policies would make things go much worse. That is how Scanlonian Contractualism could, as he intends, apply to the acts with which we affect future people.

I shall end by discussing one other way in which Scanlon might revise his view. According to what we can call Scanlon's

Individualist Restriction: In rejecting some moral principle, we must appeal to this principle's implications only for ourselves and for other single people.

In Scanlon's words: "... the justifiability of a moral principle depends only on individuals' reasons for objecting to that principle and alternatives to it."⁹ We can also call such reasons *personal grounds* for rejecting some principle. The strength of these grounds depends in part on how great the burdens are that this principle's acceptance would or might impose on us. This strength may also depend on certain other facts, such as how badly off we are, and whether we are responsible for the fact that either we or others will have to bear certain burdens. Some reasonable personal grounds for rejecting principles, Scanlon

9 Scanlon (1998) 229.

adds, may have nothing to do with our well-being. I shall not discuss such grounds here.

Scanlon's *Individualist Restriction* is given some support by one of Scanlon's most appealing ideas, that of justifiability to *each* person. Since we are asking which are the principles that *no one* could reasonably reject, we must consider each person's grounds for rejecting some principle, and we can plausibly claim that these grounds are provided by this principle's implications for *this* person.

Scanlon also defends this claim in another way. Like Rawls, Scanlon intends his Contractualism to provide 'a clear account of the foundations of non-Utilitarian moral reasoning'.¹⁰ Act Utilitarians believe that it would always be right to impose great burdens on a few people, if we could thereby give small benefits to enough other people. In one of Scanlon's imagined cases, *Jones* has suffered an accident in the transmitter room of a television station. To save Jones from one hour of severe pain, we would have to cancel part of the broadcast of a football game, which is giving pleasure to very many people.¹¹

If this broadcast was giving pleasure to enough people, Utilitarians would believe that we ought to let Jones have his hour of severe pain. Utilitarians reach such unacceptable conclusions, Scanlon suggests, because they mistakenly *add together* different people's benefits and burdens. By appealing to the *Individualist Restriction*, Scanlon writes, we can avoid such conclusions "in what seems, intuitively, to be the right way."¹² In his words: "A contractualist theory, in which all objections to a principle must be raised by individuals, blocks such justifications in an intuitively appealing way. It allows the intuitively compelling complaints of those who are severely burdened to be heard, while, on the other side, the sum of the smaller benefits to others has no justificatory weight, since there is no individual who enjoys these benefits ..."¹³

On the simplest form of Scanlon's *Individualist Restriction*, benefits to different people cannot ever be *morally summed*. In applying *Scanlon's Formula* to any two conflicting principles, we should compare only the strongest personal objection that any one person would have to one of these principles, and the strongest objection that anyone else would have to the other principle. We can ignore significantly weaker objections to some principle, whatever the number of people who would have these objections. In Scanlon's phrase, *the numbers do not count*.

¹⁰ Scanlon (1998) 267. He also writes that he is one of those "... who look to views such as contractualism specifically as ways of avoiding utilitarianism." (1998) 215.

¹¹ Scanlon (1998) 235.

¹² Scanlon (1998) 241.

¹³ Scanlon (1998) 230.

Scanlon qualifies this view in two ways. He suggests that, when different possible acts would impose equal burdens on different people, numbers can break ties, since we ought to impose such burdens on as few people as we can.¹⁴ Scanlon also suggests that, when one burden is not much smaller than another, the numbers count. To avoid these complications, I shall discuss cases in which we could either save one person from some great burden, or save many other people from *much* smaller burdens.

Scanlon's *Individualist Restriction* is not, I believe, the right way to avoid unacceptable Utilitarian conclusions. Scanlon misdiagnoses how Utilitarians reach these conclusions. Their mistake is not their belief that the numbers count, but their belief that it makes no moral difference how benefits and burdens are distributed between different people.

To illustrate this distinction, we can suppose that Jack and several other people are aged 25, and have life-shortening medical conditions. We are doctors, who must choose whom to treat. In *Case One*, the relevant facts are these:

- If we did nothing, Jack would live to the age of 30, and the other people would live to 70.
- If we treat Jack, he, like the others, will live to 70.
- If we treat the others, Jack will die at 30 but the others will live to 75.

Scanlon's view implies that we ought to give Jack his 40 more years of life, whatever the number of other people to whom we could instead give 5 more years. If the number of the other people would be very large, this view would, I believe, be too extreme. But it would be fairly plausible to claim that we ought to give Jack his 40 more years of life rather than giving 5 more years to each of eight, twelve, twenty, or even more of these other people.

Though *Scanlon's Formula* gives a fairly plausible answer here, it does not, I believe, support this answer in the right way. If we ought to treat Jack rather than these other people, that is not because we would be denying Jack the eight times greater benefit of 40 more years of life. It is because, if we don't give Jack this benefit, Jack would be much worse off than these other people, since Jack would die at 30 and the others would live to 70. To show this fact to be what matters, we can change this feature of this case. Suppose that, in *Case Two*, the relevant facts are these:

- If we did nothing, Jack and all the other people would live to only 30.
- If we treat Jack in one way, he will live to 70 but the others will live to only 30.
- If we treat Jack and the others in a different way, Jack and the others will all live to 35.

14 Scanlon (1998) 240.

On Scanlon's view, we ought to give Jack his 40 more years of life rather than giving 5 more years to Jack and to as many as a million of these other people. That is clearly false. And what makes it false is not merely that, compared with 40 more years, 5 *million* more years of life would be a vastly greater total sum of benefits. These benefits would also be more fairly distributed between different people. It would be clearly better if, rather than Jack's living to the age of 70 rather than 30, Jack and a million other people each live to 35 rather than 30. This second outcome would be better, I believe, even if these 5 extra years came to as few as seven, or six, or perhaps even fewer of these other people.

These cases show, I believe, that Scanlon ought to drop his *Individualist Restriction*.¹⁵ Rather than giving Jack 40 more years of life, we ought to give 5 more years to Jack and at least seven other people who, without these years, would die as young as Jack. For *Scanlon's Formula* to give the right answer in such cases, Scanlon must allow that these many other people could reasonably reject any principle that did not require us to give these benefits to them. Since the benefits to *each* of these other people of five more years would be much smaller than the benefit of forty years that we could give Jack, these other people must be allowed to appeal to the fact that, as well as being as badly off as Jack, if they die at 30, *they together* would receive a greater total sum of benefits, in significant amounts of five years per person. Each of these people must be allowed to appeal to this fact, speaking on behalf of this group.

As such cases also show, it is not only Utilitarianism that gives weight to the numbers of people who might receive benefits or burdens. So do all plausible distributive principles. Rather than greatly benefitting one person, we often ought to give much smaller benefits to each of many people who are just as badly off.

Scanlon claims that his *Individualist Restriction* is central to the guiding idea of Contractualism, and is also what enables it to provide a clear alternative to Utilitarianism.¹⁶ This claim implies that, if Scanlon dropped this restriction, Scanlon's view would cease to provide a clear alternative to Utilitarianism. But

¹⁵ These claims apply only to those cases in which both (1) the baseline is equal and (2) we can give much greater benefits to some people than to others. If we could give equal benefits to each person, as is often true, no one could reasonably reject a principle requiring us to give everyone such benefits. But cases in which (1) and (2) are true, though they are much less common, help us to see more clearly what is distinctive in the version of Scanlon's view that includes his *Individualist Restriction*.

¹⁶ Scanlon (1998) 229.

that is not so. Even without the *Individualist Restriction*, Scanlonian Contractualism would provide such an alternative.

Here is one of the many ways in which that is true. According to what we can call

The Contractualist Priority View: People have stronger moral claims, and stronger grounds to reject some moral principle, the worse off these people are,

Scanlon supposes we could either do something to relieve the pain of someone who has much pain, or do more to relieve the pain of someone who has little pain. In such a case, he writes, the fact that the first has more pain strengthens her claim to be given the lesser of these two benefits.¹⁷ As well as dropping his *Individualist Restriction*, Scanlon ought, I believe, to give more weight to this *Contractualist Priority View*.

I have now described four ways in which Scanlonian Contractualism could be developed further. Scanlon, I believe, should claim that, if acts are wrong in his contractualist sense, that makes them wrong in other senses. He should reject his *Impersonalist* and *Individualist Restrictions*, and he should return to a stronger version of his *Priority View*. These changes would all, I believe, strengthen what is already one of the best moral theories.

References

- Parfit, Derek (2011) *On What Matters*, Volume II (Oxford: Oxford University Press)
- Scanlon, T. M. (1998) *What We Owe to Each Other* (Cambridge, MA: Harvard University Press)
- Scanlon, T. M. (2004) "Replies" in Stratton-Lake, Philip *On What We Owe to Each Other* (Oxford: Blackwell Publishing)
- Scanlon, T. M. (2007) "Wrongness and Reasons," in Russ Shafer-Landau, ed., *Oxford Studies in Metaethics* (Oxford: Oxford University Press)

¹⁷ He then writes: "... the way in which A's situation is worse strengthens her claim to have *something* done about her pain, even if it is less than could be done for someone else." Scanlon (1998) 227.

Zofia Stemplowska

Substantive Responsibility and the Causal Thesis

1 Introduction

Scanlon has given us an account of substantive responsibility. Why is such an account difficult to construct? Holding people substantively responsible involves an imposition of tangible burdens: those substantively responsible for a given outcome cannot appeal for assistance with their situation on the grounds that others should have done more to protect them from whatever befell them – insolvency, medical trouble, loss of one’s home, and worse. This may mean having no claim to assistance at all.

One way to justify the imposition of such burdens is to appeal to the idea that people chose freely when they made the choices that triggered the burdens. But this strategy is blocked if we assume the truth of the causal thesis. According to that thesis, one’s actions and attitudes are caused by factors external to the agent (and in this specific sense, are not under one’s ultimate control). On a popular version of the causal thesis that I will assume throughout, a person’s actions and attitudes are made inevitable by the state of the world prior to that person’s birth. If so, the imposition of burdens on the person seems misplaced – akin to the medieval practice of putting pigs on trial and punishing them for stealing food.

Famously, Scanlon defends a version of substantive responsibility that is meant to be immune to the threat of the causal thesis. His strategy is to appeal to the idea that those who have no complaint against others that an outcome arose had an ‘adequate’ or ‘good enough’ opportunity to avoid the outcome.¹ In what follows, I question Scanlon’s account and gesture at an alternative that is Scanlonian in spirit and also meant to be compatible with the causal the-

Acknowledgment: I am grateful to Simon Caney, Ben Jackson, Jonathan Quong, and Victor Tadros for written comments delivered at short notice and to the audiences at the 7th International Lauener Symposium on Analytical Philosophy on Themes from T. M. Scanlon, in Bern, and at the Centre for the Study of Social Justice, in Oxford, especially to Cécile Fabre and Thomas Sinclair for their follow-up comments.

1 T. M. Scanlon 1998, chapter 6 and Scanlon 2008, pp. 198–204. See also Scanlon 2013.

sis. Any account of substantive responsibility worth having, in my view, does need to be compatible with the causal thesis.

2 Substantive responsibility according to Scanlon

To be substantively responsible for an outcome is to have no complaint against others that the outcome arose. Scanlon memorably asks us to imagine the case of toxic waste removal in our neighbourhood. The relevant bodies have issued warnings (leaflets, radio announcements and similar) to alert people that they need to stay indoors at the time of the removal and erected a fence round the excavation area to keep people out. The warnings, however, only awoke the curiosity of Curious who climbed the fence and went to look at the site. Assuming the protection that was offered Curious is deemed ‘good enough’, she is substantively responsible for the medical condition she develops as a result of exposure to the toxins.

It is a testament to Scanlon’s example that the fictional character of Curious does not seem too far-fetched. We can be less sure of the motivations of real people, of course, but Curious-like behaviour seems on display in the real world:

Since the 2004 Indian Ocean tsunami, the Australian government has spent [tens of millions of dollars] on the development and deployment of the Australian Tsunami Warning System... The Chilean tsunami was forecast to arrive at Sydney at about 8.30 in the morning on 28 February. Despite widespread television and radio broadcasts of the tsunami warning, live television coverage showed hundreds of people standing on beaches waiting for the tsunami to arrive. Worse, some of them were deliberately swimming into the incoming tsunami, despite the efforts made by volunteers from Surf Life Saving Australia to prevent this.²

As the authors of the text recognise, we cannot be sure that all those involved heard the warnings but it seems reasonable to assume that many of them did. For a different example of conduct that seems designed to court self-inflicted disaster, consider the popular British practice of binge-drinking on the weekend.³

² Dominey-Howes and Goff 2010, 350. I am grateful to Jacob J. Krich for drawing my attention to the letter in *Nature*. I mention it also in Stemplowska 2013, 488–507.

³ Many binge-drinkers would not qualify as addicts, rather they engage in a cultural phenomenon (raising further questions about cultural and social determinants of human behaviour that I put aside here since, in any case, I assume the truth of the causal thesis).

According to a 2015 report, binge drinking costs the UK taxpayer 4.9 billion GBP a year. This figure includes the costs of increased ‘A&E admissions, road accidents and police officers on duty’.⁴

Conduct that leads to disadvantage and even disaster despite ‘good enough’ protection is widespread in our world. Clearly, then, we need a philosophical account of what can be said to those who, like Curious, took steps that led to their own disadvantage. When are they substantively responsible for their fate? Such accounts are needed all the more in countries, such as most EU countries, where there are spiralling medical costs but medical assistance continues to be offered free at the point of use.⁵

A natural interpretation of what it means to be substantively responsible invites the conclusion that people in the situation of Curious are not entitled to assistance here, but this is not Scanlon’s vision. Rather, being substantively responsible means that one has no complaint that the outcome arose, but one might still have a claim for assistance on other grounds (and, if so, a complaint should assistance be withheld). So Curious might have no complaint that she is suffering pain, but may have a complaint if her medical costs are not covered. That said, to give teeth to the conclusion of substantive responsibility here, it might be best to imagine situations in which one has no further claims against others for assistance once the outcome in question arises. This may be, for example, because resources are in short supply and any claim to assistance one might have had are trumped by the claims of those in even greater need or those who are not themselves substantively responsible. From now on, I will assume that the relevant harm is that of the medical condition and the costs of managing it.

As reported, Scanlon thinks that to be substantively responsible, a person must have had an adequate (or ‘good enough’ – I will use the terms interchangeably) opportunity to avoid the relevant outcome. Good enough does not mean sufficient to actually avoid it. If a deterministic version of the causal thesis is true then no one who ends up harmed has had a genuine (ultimate) opportunity to avoid the outcome. An opportunity can be good enough, argues Scanlon, even if it fails to actually protect, provided we have ‘done enough’ to protect the person. So a weaker interpretation of ‘good enough’ is needed for Scanlon to claim his account is immune to the threat posed by the causal thesis. Of course, we might have ‘done enough’ in one sense because there is nothing else left to

4 BBC News, ‘Binge drinking ‘costing UK taxpayers £4.9bn’’, 31 March 2015, available online: <http://www.bbc.co.uk/news/uk-32126518> [accessed 25 July 2017].

5 I do not mean to imply here that the medical costs are spiralling due to reckless human conduct; they are primarily growing even in well-established public health systems due to medical advances and aging populations.

do. But we can dispense with that interpretation right away: Scanlon does not mean for his idea to rest on this thought. After all, in the toxic removal cases where Curious ends up substantively responsible for her contamination, we could have installed a 24 h guard at the excavation site and erected a screen there to show a movie on a loop about the dangers of toxic exposure.

I tackle what might count as a good enough opportunity below. First, let me discard the suggestion that such an adequate opportunity is a necessary condition for substantive responsibility. I do not think that Scanlon would disagree, but this is worth clarifying. Consider a situation – unlike the toxic waste one – in which the cost to the person who ends up harmed is lower than the cost to at least one other person if more is done to protect the former. Suppose, for example, that a pregnant woman faces a high risk of piles in her calves unless £1 m is spent on her pregnancy medication (and this is the only measure available to mitigate the risk to her). Since we are considering policies that would cover many, the costs would mount quickly if all were to be offered such protection. This would doubtless affect the availability of other medical provision. If so, then it is not clear that the pregnant woman needs to be offered any protection against the condition at all and yet, it would seem, she would still lack a complaint against others when she develops piles. Similarly, we can imagine the need for toxic waste removal in our war-torn country where resources are so scarce that Curious could not even ask for a warning and a fence – we just have to get on with the waste removal as soon as possible. And still Curious would presumably have no complaint against others if she ended up exposed. My point here, of course, is not that people have no claim to be warned/protected from disadvantageous outcomes, but that they may lack the relevant complaint despite the absence of such measures and a corresponding adequate opportunity to avoid the outcome. Still, giving people an adequate opportunity to avoid a given outcome can be thought of as a way through which people can be permissibly exposed to a greater risk of harm than would otherwise be possible. Perhaps if the pregnant woman is informed of increased risk of piles from some procedure, the procedure can be offered to her. It is easy to think of further cases where being able to expose people to a greater risk of harm could let us adopt beneficial policies: for example, we can undertake hazardous waste removal.

3 An opportunity to avoid harm

But if we must assume the truth of the causal thesis, what would it mean to say that people had a good enough opportunity to avoid getting harmed? Can such a good enough opportunity ever be offered? Scanlon thinks that it can. Indeed, he

thinks Curious was offered such a good enough opportunity. This can be contrasted with another person, Walker who – in an example developed from Scanlon’s writings by Alex Voorhoeve⁶ – goes out for a stroll on the day of the toxic waste removal because the warnings about it happening did not reach him. In fact, following Voorhoeve, suppose that we are facing a choice of policies: we can adopt either Inform Everyone – a policy that manages to reach everyone with a warning about the toxic removal but that also sparks the curiosity of Curious, or Vivid Warning – a policy that manages to persuade anyone whom the warnings reach to avoid toxin exposure but whose coverage is more limited and the warnings fail to reach those who frequently go on strolls, e.g. Walker. Suppose we conclude that the opportunity offered to Walker was not ‘good enough’ such that, although steps were taken to protect him, he retains a complaint and is not substantively responsible for the exposure to the toxins. Should we not say the same thing about Curious?

But Scanlon sees Curious as substantively responsible in the stylised scenario. To bolster this point, and building on the work of Johann Frick, Scanlon offers an analogy with vaccinations.⁷ We can say that people have been given a good enough opportunity to avoid a disease if they have been vaccinated against it, even if the vaccination did not in fact protect them in the sense that they did not develop immunity (I will discount throughout the herd immunity impact of vaccinations). Specifically, adapting Frick’s examples, we can imagine three schemes which, let us assume, end up leaving different people to catch the disease (and suppose that due to budget constraints, we can only develop one of the schemes). For simplicity assume also that those for whom a vaccination scheme is ineffective all catch the disease and die:

- (1) The first scheme works by providing everyone with a dose of vaccination that is known ex ante to only work for a specific 95% of the population (the remaining 5% are offered it as a gesture in case they don’t want to know which group they belong to).
- (2) The second scheme works by providing everyone with a dose that is known ex ante to be 95% effective against the disease. We are under a ‘natural veil of ignorance’⁸ in that we have no ex ante knowledge – and no reasonable prospect of acquiring it – of who will end up in the unprotected 5% (though, given that the vaccination analogy is meant to work as an analogy for choice if even the deterministic version of the causal

⁶ Scanlon 1999; Voorhoeve 2008. I also discuss the example in Stemplowska 2013. This article offers a further defence of the position I develop in that paper.

⁷ Scanlon 2008, p. 205 and Scanlon 2013, *passim*. Frick 2013. My discussion that follows was developed before Covid-19 pandemic and is not intended as a commentary on the current vaccination regimes.

⁸ Frick 2013, p. 190.

thesis is true, we can assume that the vaccination was always bound to fail those it did fail; we can imagine, for example, they had antibodies that blocked the development of the new antibodies).

- (3) The third scheme works by providing a randomly selected 95% of the population with a dose of vaccination that is 100% effective against the disease, while the remaining 5% get a placebo.

Clearly, adopting the first scheme means that there are some people to whom we cannot say that we did anything to protect them. Similarly, Scanlon argues that the 3rd scheme – sufficiently analogous to the case of Walker under the Vivid Warning policy – also fails to offer protection to the 5%: instead of reducing their risk of the disease we offered them a placebo.

However, on Scanlon's account, the 2nd scheme – despite not being 100% effective – is seen as offering protection to everyone. Why? We can say to each person: 'our actions reduced your risk of catching the disease'. We can look each person in the eye and say this because we are under a 'natural veil of ignorance' here so although we know some people will end up catching the disease, we do not and cannot reasonably know in advance who they are. If so, we can say that the 2nd scheme offers a good enough protection, despite the fact that some are bound to lose out. By analogy, at least if we did not know in advance who would act like Curious, we can be said to have protected people like Curious when we adopted Inform Everyone, even if, given that the causal thesis is true, they were bound to come to harm.

The analogy with vaccinations is illuminating. But, to my mind, we are still left with three reasons to reject the conclusion that Curious is substantively responsible for her fate. First, it seems to me that those who end up disadvantaged by their choices, like Curious, are often in a situation analogous to that of the 5% from the 1st scheme rather than the 2nd. We may have offered our warnings against toxic excavation sites, but of many people we can reasonably know, or can come to know, in advance that they will not heed the warning. For example, we each know some people who – we can reliably predict – if given a phone and a car will text and drive or, if alcohol is affordable, will drink too much and otherwise act self-destructively. If so, we cannot always claim that our policies designed to protect people do offer them an opportunity to avoid the outcomes any more than we can claim that a medicine that works only with chromosome Y works for women. We can think of Curious here like we would of a 2 year old to whom we show a biscuit in a cupboard and warn her not to eat it as it is poisoned with norovirus. Pointing out a poisoned biscuit to a two year old cannot be said to protect her from the poison. If so, Scanlon cannot argue that people like Curious, about whom we can reasonably know in advance that they will court

harm, can be held substantively responsible on the grounds of having been offered a good enough opportunity to avoid the harm.

Second, even when we do not know who such people are in advance, we may know that this type of person exists; offering some people warnings, or at least some types of warnings, will trigger destructive behaviour. If so, the vaccination scenarios presented so far are disanalogous to the case of toxic removal as specified above. Recall that Curious only went to the excavation site because she was being offered the warning: the warning triggered her curiosity. If so, a more analogous vaccination scheme would look more along the lines of:

- (4) We can offer everyone a vaccination that will be 100% effective against contracting the disease, but in some people, it will cause suicidal thoughts that will lead to death. 5% will be so affected.

Even if we are behind a natural veil of ignorance here it seems to me that, should this 4th scheme be adopted, the 5% will have a complaint that their situation was not improved by the vaccination (and relative to scheme (3), might have been worsened). If we would not say in this case that the people were offered a good enough protection by us then we should not say that Informed Everyone in a world with people like Curious in it offered a good enough protection either.

Third, suppose, again, that we are indeed simply comparing schemes (2) and (3). It is not clear to me that even in this constrained case the 2nd scheme is really any better at offering protection than the third, so if Scanlon is right that the 3rd scheme fails to protect the 5% then so does the 2nd scheme. Consider the two schemes. Notice, first, that even if we adopt the 3rd scheme, we can still say to those who lose out: ‘since you were included in the lottery, we did something to protect you relative to had no vaccination been on offer’. Second, if we adopt the 2nd scheme, as Scanlon suggests, those who lose out can complain that had the 3rd scheme been adopted they would have been protected from their bodies’ physiological response, but as things stand they were not. This, on its own, does not mean that we should have adopted the 3rd scheme – that scheme also would have failed to protect some – but it does, I think, undermine the claim that one of the schemes is better at protecting everyone. After all, we cannot say that if we adopt the 2nd scheme we will have done more to protect everyone than if we adopt the 3rd scheme: there is a point at which all people enjoy a 95% chance of developing an immunity (and the randomised procedure for offering placebos we devise can keep us ignorant of who is getting what as we administer the doses). We may have other reasons to pick scheme 2 over scheme 3 but offering better protection does not appear to be such a reason.

If we cannot pick scheme 2 over 3, we cannot pick Inform Everyone over Vivid Warning. Thus if we would balk at seeing Walker as substantively responsible when he comes to harm, we should also not see Curious as substantively responsible for her fate on the grounds that she was offered ‘good enough’ protection from the harmful outcome she suffers.

4 An opportunity to choose one’s fate

Scanlon might appeal to a separate strategy for grounding substantive responsibility in the case of Curious. This strategy is to suggest that if we offer someone an opportunity to avoid an outcome then her fate depends on who she is.⁹ I call this a separate strategy since, following Scanlon, what matters here is not whether the opportunity increases or decreases a risk of something and is for that reason good enough but whether it makes what happens depend on one’s preferences. Curious might not have been protected by the choice she was offered but she was put in a situation in which what happens to her reflects her preferences (to go to the excavation site despite warnings). The thought here is that it matters when our fate is determined by who we are.

I do not want to dismiss this as an important consideration when it comes to policy selection, but it is not obvious to me that on its own it is powerful enough to block one from having a complaint against the outcome (and indeed I take it that Scanlon does not seem to think it either, which is why he emphasises the need for offering good enough opportunities). Suppose some harm has to befall me no matter which policy is adopted. For example, imagine that I am in Pompei the day before the volcano eruption and I can be either offered a horse to ride away (which, if I try to ride it, I will fall off and kill myself) or I can be offered shelter in a public bath house in Pompei. It is not even clear to me that it is preferable for my fate to reflect that I ineffectively tried to ride a horse than that I sat in the bathhouse when the volcano erupted, let alone that either option would have protected me. Without some further argument – which I hope to offer below for a restricted set of cases – it is not clear to me why I would have less of a complaint if a policy was adopted that would mean that any harm that comes to me reflects my *choice* under conditions in which offering it was bound to lead me to choose badly.

⁹ Scanlon 2008, 192–198 and Scanlon 2013.

5 Substantive responsibility: A narrower account

Can we ever hold people substantively responsible, then, for self-inflicted harm, even if the causal thesis is true? Let's call choosing to act (from a menu of options at least nominally open to one) in a way that reasonably foreseeably and noticeably increases one's risk of harm (or just straightforwardly leads to harm) as 'courting harm', and let's call one's demands that one be protected from courting harm or protected from bearing the resulting harm as demands to 'share the costs of harm'. Finally, let's call costs that people would not need to bear for one another if they each acted of their own free will when courting harm 'undue costs'. Not all shared costs of harm are undue costs. For example, even if the causal thesis is false and there is proper free will, there are costs that I can permissibly impose on others – for example, as I skillfully and attentively drive to work in my hybrid car I can permissibly impose a greater risk of accident on others and ask for medical assistance when I suffer harm. But there are also costs that would count as undue – for example, the costs that arise due to an accident I suffered because I texted as I drove.

Clearly, people may court harm out of a range of attitudes (whether they know it or not). What I suggest below is that sometimes we can ground attributions of substantive responsibility in the attitudes that gave rise to people's conduct. Suppose, then, that the actions that lead to self-imposed harm can be attributed to the following attitudes (though, of course, the list is not exhaustive):

1. I court harm even though I sincerely try not to; I just cannot act on my sincere preferences.
2. I court harm only if, as far as I can reasonably know, this will not impose undue costs on others.
3. I court harm without any regard if (or because) this will impose undue costs on others and ask for others to share the costs of the harm that I court this way, even if this amounts to undue costs to them.

I will put the 1st case to one side as it is clinical.¹⁰ Attitudes captured by (2) are innocent and I do not have an argument for grounding substantive responsibility for conduct exhibiting them. And, in any case, I think that in a deterministic universe, people who come to harm acting on such innocent attitudes are not substantively responsible for their fate for the reasons adduced in sections 3 and 4.

¹⁰ In my view such a person should not bear substantive responsibility for any harm once we assume the causal thesis is true.

The fact that they ‘chose’ to court harm is irrelevant to their substantive responsibility given the truth of the causal thesis.

But the unreasonable attitudes captured by (3) allow us, in my view, to ground substantive responsibility. An iconic case of conduct of type (3) would be, say, that exhibited by 19th century literary gamblers or, say, Lydia in *Pride and Prejudice* (at least if the harm that might befall her and her sisters was not so preposterous and unjust). But for the purposes of my argument here it is best to imagine someone, call her Self-centred, who upon hearing about the waste excavation goes to the site to satisfy her curiosity thinking that should harm befall her the free-at-the-point-of-use rescue and health systems will be mobilized to help her out. The more familiar character would be that of a texting driver or binge drinker in a country such as the UK. Those who, like Self-centred, act on attitudes of type (3) can manifest disregard for the interests of others in two ways: first, when they court harm expecting others to share the undue costs or, second, when they demand expensive policies that would stop them courting harm the costs of which they would otherwise demand to be shared. They manifest disregard for the interests of others here because they are acting in ways designed to impose undue costs on them.

How should we respond? It is tempting to say that although the person in question disregards the interests of others through her conduct, her attitudes are not ultimately due to her, if the causal thesis is true, and so she is entitled to no less (and has no less of a complaint) than anyone who comes to harm while acting on attitudes of type (2). If this is correct then, in my view, the Scanlonian project of holding people substantively responsible in a world accurately described by the causal thesis must fail.

But I am not convinced that we should be satisfied with this response. Why? Such a response would amount to allowing people to impose costs on others out of disregard for their interests. After all, by assumption, we only face the need to share the costs of the harm because those acting on the unreasonable attitudes have no regard for the interests of others. But people should not end up with (non-trivially) less than they would otherwise have simply because others disregard them.

How often are we in the presence of cases where the agent acts out of attitudes of the 3rd type as she courts danger? It is hard to know, of course, but it is not unheard of. For example, people who live with public healthcare provision know that if they miss appointments, text and drive, or bungee jump drunk, the system bears the costs of depriving others of the resources. This raises the hard question of which actions should count as imposing undue rather than acceptable costs on others and when people can be reasonably expected to be aware of which costs are undue. We can expect to face much controversy on

this issue including over conduct such as smoking¹¹ or voting for policies and persons who diminish a country, though perhaps also some degree of consensus on drunk driving. As Samuel Scheffler once pointed out, we need a new conception of responsibility ‘in a world of complex interdependencies, massive institutional structures, and breathtaking new technologies’, a world in which we are not on a sure footing regarding individual agency.¹² I think we now not only inhabit such a world but many of us are also in a position to realise that we do.

Admittedly, the appeal to objectionable attitudes offered here does not allow us to hold people substantively responsible in all the same cases that Scanlon thinks people should be held substantively responsible. But then, if I am correct, Scanlon’s argument also does not allow us to hold people substantively responsible if the causal thesis is assumed to be true. Of course, it may turn out that the causal thesis is false, and thus attributing substantive responsibility is easier than we had thought. Or, even if the causal thesis is true, it may turn out that the objectionable attitudes are rarely, if ever, displayed or that we cannot discern them when they are. If so, my argument will not allow us to attribute substantive responsibility to people either. However, if we cannot do so because people do not seem to act on objectionable attitudes, it will be less urgent to offer accounts of substantive responsibility.

References

- Dominey-Howes, Dale and Goff, James, “Tsunami: Unexpected Blow Foils Flawless Warning System”, *Nature* 464 (18 March 2010): 350.
- Frick, Johann, “Contractualism and Social Risk”, *Philosophy & Public Affairs* 43 (2015): 175–223.
- Scanlon, Thomas M., *What We Owe to Each Other*, Cambridge, MA: Harvard University Press, 1998.
- Scanlon, Thomas M., *Moral Dimensions: Permissibility, Meaning, Blame*, Cambridge, MA: Harvard University Press, 2008.
- Scanlon, Thomas M. “Response to Stemplowska”, *Journal of Moral Philosophy* 10 (2013): 508–514.
- Scanlon, Thomas, M., *Being Realistic about Reasons*, Oxford: Oxford University Press, 2014.
- Scheffler, Samuel, *Boundaries and Allegiances*, Oxford: Oxford University Press, 2001.
- Stemplowska, Zofia, “Responsibility and Respect” in Carl Knight and Zofia Stemplowska (eds.), *Responsibility and Distributive Justice*, Oxford: Oxford University Press, 2011.

¹¹ I briefly defend the view that smoking, despite the shared healthcare costs, may not be an imposition of undue costs on others in Stemplowska 2011.

¹² Scheffler 2001, chapters 2 and 3, quote p. 65.

Stemplowska, Zofia, “Harmful Choices: Scanlon and Voorhoeve on Substantive Responsibility”, *Journal of Moral Philosophy* 10 (2013): 488–507.

Voorhoeve, Alex, ‘Scanlon on Substantive Responsibility’, *Journal of Political Philosophy* 16 (2008): 184–200.

T. M. Scanlon

Responses to Forst, Mantel, Nagel, Olsaretti, Parfit, and Stemplowska

The very interesting and stimulating papers by Rainer Forst, Susanne Mantel, Thomas Nagel, Serena Olsaretti, Derek Parfit, and Zofia Stemplowska give me a great deal to think about. What follows are some initial thoughts in response.

Response to Rainer Forst

Rainer Forst offers a sympathetic reconstruction of my views about practical reason and morality, and goes on to raise an important question about the relation between the role of justification in determining moral truth and the kind of justification that must underlie actual political institutions.

Forst suggests that I should be a justification fundamentalist rather than a reasons fundamentalist. What is basic according to the justification fundamentalist is a certain process of justification determined by principles of reasoning. "Reasoning guided by principles of reasoning determines good reasons, not the other way around." Reasons fundamentalism holds, by contrast, that there are basic facts about which considerations are reasons for certain actions and attitudes and which are not. Reasoning is valid when it gets these facts right.

Justification fundamentalism is an appealing position. I wish that there were principles of reasoning that determined truths in the normative domain. As I say in Chapter 4 of *Being Realistic about Reasons*, however, I do not see that there are any such principles, at least not ones that provide a basis for substantive truths about the reasons we have. There are formal constraints, such as requirements of consistency and the principle of instrumental reasoning, but these formal constraints alone do not tell us what reasons we have. We can discover substantive truths about reasons by engaging in a process of seeking reflective equilibrium, but this is only a way of discovering these truths, not a basis for them.

Morality, or at least that part of it that I call the morality of what we owe to each other, is a different matter. According to my contractualist view, facts about (this kind of) moral right and wrong are determined by facts about reasons, namely by facts about whether there are sufficient reasons for rejecting certain principles of action. A justification fundamentalist about morality could, as Forst suggests, be agnostic about whether the facts about reasons that the contractualist process of justification appeals to are themselves dependent on a fur-

<https://doi.org/10.1515/9783110733679-012>

ther form of justification or are fundamental normative facts in their own right. I myself, however, am not agnostic about this, but take the latter, realist view. So I am a reasons fundamentalist about normativity in general and a justification fundamentalist about (at least this part of) morality.

The kind of justification that I take to be fundamental in the case of morality is a matter of abstract normative truth. Whether a moral principle allowing certain actions is justified in the relevant sense is simply a matter of what the normative facts are: whether certain considerations are or are not sufficient reasons for rejecting such a principle. We can engage in reasoning about morality—a process of thinking about what the normative facts are. And we can engage in a process of justifying our actions to others—a process of pointing out to them what we take these facts to be. But these activities of actual reasoning and justifying are not fundamental to morality, according to my view. What is fundamental is whether certain principles are in fact justified.

Things are, however, more complicated in the case of political justification. Speaking of the political realm, Forst writes, “Here the ideal of justifiability towards others has a concrete, institutional meaning calling, among other things, for a legally, politically and socially secure status of non-domination, i. e., of not being subject to a normative order without proper justification and, especially, of not being subjected to a normative order without proper procedures and institutions of justification in place.” He also writes, “human beings have a basic moral right to justification which is so to speak a veto right against reciprocally and generally non-justifiable normative claims.”

This is something I could agree with, although I would not put it in terms of a right. The idea of a “veto” can be understood in two different ways, however. One way of understanding it is in terms of the abstract idea of justifiability that I mentioned above. To say that a person could veto a principle in this sense is just to state the normative fact that a person in that position has sufficient reason, of the relevant kind, to reject such a principle. According to contractualism, it would be wrong to treat a person in ways that would be allowed only by principles that he “could veto” in this sense, or to subject him to political institutions that would be allowed only by such principles.

But Forst’s reference to “procedures and institutions of justification” suggests a different “more concrete, institutional meaning” for the idea of a veto. In the case of political institutions, the idea would be that it is wrong to subject people to institutions that do not include “proper procedures and institutions of justification” giving them a voice in determining the laws and policies that apply to them. This idea itself might be understood as claiming that principles permitting institutions that do not include such procedures “could be vetoed” in the abstract sense I have described. On this reading, we would have two layers of jus-

tification: laws and policies would be justified through a participatory political process that is required by principles that are justified in the more abstract sense of being allowed by principles that are supported by sufficient reasons.

But Forst can also be read as saying that the principles by which political institutions are judged must be subject to justification through an actual process in which individuals take part, not merely justified in this abstract sense. He writes, “Basic rights are not just rights to be protected in one’s status as a legally, politically and socially non-dominated person; they are, in a reflexive sense, also rights to determine the rights and duties that define this status. Thus, the authority to determine your rights must reside in a discursive procedure of reciprocal and general justification in which all participants are justificatory equals.” That we are “justificatory equals” means that what is to be taken into account in determining the principles by which our conduct and our institutions are to be judged is not only the reasons that we have, in virtue of how we would be affected, but also our judgment about what principles are justified in the light of these reasons.

This way of looking at the matter draws support from the fact that the principles by which institutions are assessed in an actual society do not exist merely in the abstract sense of being morally correct. They also play a role in actual politics by being generally accepted. For example, a well-ordered society in Rawls’ sense is one in which there are shared principles of justice that provide a public standard by which complaints against the basic structure are assessed.¹ The existence of such “shared” principles is a social fact about the actual attitudes of the members of the society. Forst can thus be read as saying that such a consensus must be maintained through “a discursive procedure of reciprocal and general justification.” That is to say, it must be open to contestation and possible revision.

This view, which I find very plausible, has definite implications. Consider, for example, Jeremy Waldron’s argument that laws restricting hate speech are justified by the need to protect what he calls the dignity of all members of society. Individuals have dignity in this sense if there is a general consensus that recognizes them as full members of the society, and if they have reasonable confidence that this consensus will be maintained.² Waldron argues that hate speech needs to be restricted insofar as it undermines this general consensus and gives individuals reason to doubt the security of their status as full and equal members of society.

1 See Rawls 1971, section 69.

2 Waldron 2012, Chapter 4.

But if the basic norms of the society, which specify what the basic rights of citizenship are and who has these rights, must be open to contestation by all, then it would seem that speech questioning these norms must be allowed, even if it threatens the dignity of some members in Waldron's sense. This conclusion is one that I accept, although I do not welcome it. I mention it here as a possible implication of Forst's requirement that the most general norms of a society must be subject to a "discursive procedure of reciprocal and general justification" in which all members are entitled to participate as justificatory equals.

This leads me to the conclusion that in Forst's view there are not two but in fact three layers of justification. First, there are political institutions through which particular laws and policies are enacted and justified. Second, there is, as we have just seen, a more informal "procedure of reciprocal and general justification" through which the accepted norms governing these political institutions are justified and maintained. But if, as Forst says, there is a *right* to participate in the latter process, this seems to presuppose a more abstract form of justification supporting the conclusion that it would be wrong to exclude some people from participating in this process, a form of justification that is based on the reasons that individuals have, rather than on their judgment about these reasons.

In fact, it seems to me (although this may be a stretch) that Forst's discursive procedure of reciprocal and general justification presupposes a deeper level of facts about justice. This process is, by his account, one in which the participants are recognized not merely as having interests that count for the purposes of moral and political justification, but also as "equal justificatory authorities when it comes to evaluating good moral reasons." The discursive process of justification is, as I have put it, one in which their *judgments* are taken into account. But what are these judgments about? They cannot, it seems to me, be judgments about what the results of this discursive procedure will in fact be. They must instead be judgments about what the outcome *ought* to be, thus implying some idea of normative truth beyond the process itself.

Response to Susanne Mantel

Susanne Mantel's insightful paper deals with the motivational challenge to a cognitivist account of normative judgments. It seems obvious that the normative judgments that an agent accepts can explain his or her subsequent action. The motivational challenge claims that this would not be possible if normative judgments were simply a kind of belief. My response appeals to the idea that acting

in accordance with one's beliefs about the reasons one has is part of being a rational agent.³

Mantel is quite correct that this response would be too narrow if it appealed only to the idea that a rational agent (insofar as she is not irrational) will do what he or she believes herself to have *conclusive* reason to do. An account of the practical significance of an agent's normative judgments must explain a much wider range of phenomena. The fact that a person sees *some* reason to act in a certain way can explain her so acting, even if she does not believe this reason to be conclusive. Indeed, we often do things that we judge ourselves not even to have sufficient reason to do, or even to have conclusive reason not to do. Even when we act irrationally, we act *for reasons*. These reasons "motivate" our action and (at least partially) explain our acting in these ways.

As Mantel says, a plausible account of these matters should be based on dispositions rather than just on strict requirements of rationality. It is part of being a rational agent to be disposed to act in ways that are responsive to—that "manifest" as she says—the judgments about reasons that we at least implicitly accept. When a rational agent acts for a reason, acting in this way manifests this disposition, and the fact that he sees this consideration as a reason (at least partially) explains his so acting. I say "at least partially" because the disposition in question must, as I indicated above, be understood very broadly, to include even cases in which the person does not judge the reason in question to be sufficient. In the example Mantel gives of the woman deciding to stay home from the party, whether she stays home because she judges herself to have conclusive reason to work, or stays home because she is angry at the host of the party, she is acting for *some* reason. Being angry often involves, among other things, seeing oneself as having reason to do things that will hurt, offend, or just express disrespect for, the person toward whom one is angry. Whichever reason the woman is acting for in staying home, and even if doing so is irrational, her staying home (for a reason) manifests a disposition that is part of being a rational agent.

If I understand Mantel correctly, she also suggests that I should rely simply on the idea of a disposition, without tying this to the idea of rational agency. This would, she suggests, broaden the appeal of the account. I am not inclined to accept this suggestion, however. Appealing simply to the claim that the connection between an agent's normative beliefs and her actions is explained by the fact that she has the disposition to act in accord with these beliefs, without saying more, seems *ad hoc*. Adding that having this disposition is part of being a ration-

³ Set out in Chapter 3 of Scanlon 2014.

al agent anchors this claim in a widely accepted view of what human agents are like.

Rather than narrowing the appeal of response to the motivational challenge, this seems to me to ground it in something that even those who favor a desire-based view ought to accept (although there will of course be some who disagree.) Holders of a desire-based view also appeal to an idea of rationality. Desires are not mere urges, but are supposed to “rationalize” an action. As Davidson says in the passage Mantel quotes, “a rationalizing explanation lets us see something the agent saw, or thought he saw, in his action – some feature, consequence, or aspect of the action the agent wanted, desired, prized, held dear, thought dutiful, beneficial, obligatory, or agreeable.” I would say that these are all ways of seeing something as “counting in favor of” the action, that is to say, of seeing something as a reason for it, and making it rational. This is something that a proponent of the motivational challenge should not deny. What they deny is that any cognitive state can have this role. This is a claim, as it were, within an idea of rational agency: a claim that any combination of mere beliefs has no implications about what it would be rational for an agent to do.

What makes my account of normative judgments cognitivist is just that it claims that these judgments are capable of being true or false. Why should this feature of an attitude rule out its having any implications for what it would be rational for the person to do? I can see why this would be so if the only attitudes capable of truth and falsity were beliefs about the natural world. But, in my view, judgments about normative questions can also be true or false, and accepting them is thus a matter of belief. So being capable of truth or falsity does not debar an attitude from also having normative content, and thus from having implications about what it would be rational for a person who holds that attitude to do. In this way, it seems, my response to the motivational challenge to normative cognitivism depends on the success of my response to the metaphysical objection to this view.⁴

Response to Thomas Nagel

Thomas Nagel’s paper raises challenging and perplexing questions about moral truth and moral progress. It seems to me that his argument can be understood in several different ways. The degree to which it is perplexing, and the reasons why it is perplexing, are different depending on how it is understood.

⁴ Set out in Chapter 2 of Scanlon 2014.

Nagel frames his argument in terms of a contrast between scientific progress and moral progress. Scientific progress can be understood in terms of an idea of truth that is independent of what people have at various times had reason to believe. The ancient Greeks had no reason to believe that salt was sodium chloride. But this was just as true in 400 B.C. as it is now. They did not have a reason to believe this because their lack of the relevant scientific understanding rendered this fact inaccessible to them. Later, when chemistry developed, and people came to understand the molecular structure of salt, they came to believe something that had been true all along.

“By contrast,” he writes, “the truth of a moral proposition cannot be distinguished from there being a reason for people to conduct themselves in the way it prescribes. And I would claim that there cannot be such a reason if there is no reason for anyone to believe there is.” In some cases, Nagel believes, the reasons we have now for believing that a certain practice is wrong and should be abandoned imply that there was never any justification for this practice, and people were simply mistaken in failing to see that it was wrong. He suggests that this is true in the case of slavery. But in other cases he believes that this is not true. “For example,” he writes, “I believe the moves to popular sovereignty and equality of opportunity between the sexes required certain levels of political, economic, and educational development before they were even imaginable as options.”

One way of understanding Nagel’s thesis that “the truth of a moral proposition cannot be distinguished from there being a reason for people to conduct themselves in the way it prescribes” and that this is inseparable from their having reason to believe that this is a reason, can be illustrated by cases of non-moral practical reasoning. If I have strong reason to escape through a door that is locked with a combination lock, then the fact that it would open if it were turned left to 6, right to 23 and then left to 17 is, in an abstract sense, a reason for me to do that. But if I have no reason to believe that this is the combination, then this reason is inaccessible to me, and I have no more reason to turn the dial in that way than in any other way.

So, on one way of understanding it, Nagel’s thesis is that the claim that a certain practice is morally wrong is true only if people *have reason* in this sense to condemn that practice and to change it if they can. By contrast, the truth of scientific claims does not depend on anyone having a reason, in this sense, to accept it. A weaker version of the thesis would rely on the distinction between the question of whether a certain practice is morally wrong and whether it is *wrong of* certain individuals to go along with this practice and not condemn it. We might, for example, say that denying equal economic opportunity to the sexes was always wrong, but that it was not wrong of individuals at certain

times to accept this practice, because the reasons for believing it to be wrong were not accessible to them.

This weaker interpretation would make Nagel's thesis less perplexing, since it would preserve a sense in which moral claims are always true (as long as the non-moral facts that make them true obtain.) But he seems to have a stronger claim in mind. His thesis that "the truth of a moral proposition cannot be distinguished from there being a reason for people to conduct themselves in the way it prescribes" is naturally understood as denying the significance of the distinction on which this weaker claim depends.

Even the weaker version of his thesis remains perplexing in one way, insofar as it relies on the possibility that moral facts can be inaccessible. In the example of the locked room, what is inaccessible to me is any reason to believe the non-normative fact that acting in a certain way would be a means to my end of escaping from the room. So in such a case there is nothing puzzling about divergence between what there is reason for one to do and what one has reason to do. It arises simply from our limited access to facts about the natural world.

But Nagel suggests that in some cases what is inaccessible to us is the normative facts themselves, not just non-normative facts that have normative significance. This possibility is more puzzling. Normative facts, and in particular moral facts about right and wrong, seem to be things that we should be able to discover simply by thinking about the matter in the right way. It thus may seem that if such things are accessible to us at all, they should always be accessible, unlike the combination to the lock, or the chemical composition of salt, which no amount of mere careful thinking could be expected to disclose.

Nagel says that "if there were a single governing moral principle, such as utilitarianism, from which all moral reasons were derived ... all moral progress would be the discovery of moral truths that in a sense had been true all along, like the chemical composition of salt—because they followed from a premise that any rational being could have recognized." What seems most important about utilitarianism is not that it is a single principle but that it is a principle "that any rational being could have recognized." Once we have come to believe in the truth of any moral proposition it may seem to us to be like this—something that any rational being could have recognized and that we, and others, were mistaken not to have recognized sooner. Nagel's thesis is that in some cases we are mistaken to think this. So one interesting question is how the cases in which we would be mistaken differ from the case of utilitarianism, in his view.

One possible answer is that some moral concepts are applicable only under certain empirical conditions. This is suggested by Nagel's observation that "moves to popular sovereignty and equality of opportunity between the sexes required certain levels of political, economic, and educational development before

they were even imaginable as options.” To take a slightly different example, it would have been foolish for an early nineteenth century Russian nobleman simply to free his serfs. Given the institutions prevailing at the time, there might have been no economic role for them, and no life for them to live, outside the system of serfdom. The act utilitarian principle would not be subject to this problem, since what it prescribes depends on the actual consequences of individual actions, not on moral judgments such as that the institution of serfdom was unjust. This line of thinking would not take us all the way to Nagel’s conclusion, however, since on this account it would still be true to say, even at that time, that serfdom was unjust. There would just be context-dependent limits on what that fact would give people reason to do.

A stronger interpretation of Nagel’s thesis would take the factors that determine accessibility to be more conceptual, and less empirical and institutional. This interpretation is suggested by what he says about the argument for freedom of speech in my 1972 article. His point there seem to be that before a certain time people lacked access to the idea that political institutions needed to be justified in the way I described. This interpretation makes the thesis more perplexing, since it runs against the idea that in the normative realm if an idea is accessible to a person at one time it was accessible to others at other times. So one wants to know why this is not so.

Nagel’s answer seems to be that the normative idea that certain institutional arrangements are desirable depends on the idea that such institutions are feasible. Access to the latter idea can depend on a certain understanding of how social institutions work that is, like scientific conclusions, available only to people with certain kinds of experience. But, in contrast to the case of science, this lack of access undermines the truth of the claims of desirability rather than just making it the case that people were not open to criticism for failing to see these truths.

I am not certain what to think about this stronger claim. Plato was able to imagine, and see the desirability of, selecting individuals for positions of power on the basis of their ability, without regard for their sex. But it might be said in response that it is one thing to imagine this for positions in Plato’s imaginary republic and quite a different thing to think of it as applying to actual positions in a real society, which may seem to people at a given time to belong “by their very nature” to individuals of a certain sex, or class. I am not certain that the difficulty, for these people, of imagining a system in which this was not so undermines the truth of the claim that positions conceived of in this way are unjust. But I do feel the force of Nagel’s challenging argument.

Response to Serena Olsaretti

It is generally recognized that economic inequality is a serious threat to equality of opportunity. Richer parents can help their children develop abilities that give them advantages over poorer children in the competition for desirable positions, and given the desire of parents to help their children they are likely to do this. This provides an argument for limiting economic inequalities in order to preserve equality of opportunity. In Rawls's theory, for example, the idea that there is such a limit follows from the fact that the requirement of fair equality of opportunity is a limitation on inequalities that might be permitted by his Difference Principle.

Olsaretti provides a different argument against economic inequalities, based on their incompatibility with parent's obligations to do as well as they can for their children. The structure of her argument is similar to that of G. A. Cohen's argument against justifications of economic inequality based on the supposed effectiveness of incentives.⁵ I believe, however, that her argument can also be understood in a different way that can have wider implications.

The account of equality of opportunity that I offer in *Why Does Inequality Matter?* focuses on unequal positions that are justified by the benefits that result if these positions are filled by individuals with the right qualifications.⁶ The aspect of equality of opportunity that I call procedural fairness (selection by merit) is just a corollary of this justification: if individuals are not selected on the basis of the abilities required to produce the relevant benefits, then the positions are not being administered in a way that fits with the justification for having them. Inequalities that are allowed by Rawls' Difference Principle because they benefit the worst off are examples of inequalities justified in this way, but they are not the only examples. Inequalities might be justified by other beneficial consequences. The general rationale for procedural fairness that I have described thus fits with Rawls' Difference Principle but also applies more generally.

Inequalities justified in this way may consist in the greater income and wealth attached to certain positions, but this is not the only possibility, or even, I think, the most likely one. Positions may be unequal because some positions involve better working conditions, or provide opportunities for the development and exercise of capabilities that individuals have reason to value. If the inequalities claimed to be justified consist of economic benefits, it may well be asked how giving some these benefits can raise the expectations of the worse

⁵ Cohen 1997.

⁶ Scanlon 2018, Chapters 4, 5. See esp. pp. 41–44.

off, or produce other beneficial consequences. One familiar answer is that economic inequalities do this by providing incentives, which attract more qualified individuals to the positions in question or induce the individuals who hold these positions to work more effectively. Again, this is not the only possibility. Unequal economic benefits might arise simply from the exercise of powers or liberties that are assigned for some other reason. I am rather skeptical about the magnitude of inequalities that could be justified by the need for incentives, and I believe that Rawls was as well. But this is the case that Cohen and Olsaretti focus on.

Cohen's argument turns on the fact that economic incentives are effective in promoting the well-being of the worse off only if those who receive them refuse to do the same work without these greater rewards. But, he argues, they cannot refuse to do this insofar as they themselves accept the Difference Principle, and therefore believe that inequalities are just only if they lead to the greatest benefit for the worst off. The worse off would benefit more if the better off were willing to work without special incentive payments. So the better off who accept the Difference Principle must be willing to do this, thereby making incentives unnecessary. Insofar as incentives do benefit the worse off, Cohen says, this is not a case of justice, but, at best, a case of injustice justified by the fact that it is made necessary by the unjust attitudes of the better off.

Olsaretti's argument has a similar structure. As Nagel and others have noted, rich parents in an economically unequal society face a dilemma.⁷ They have strong reason to use the resources at their disposal to provide their children with good schooling and other developmental conditions that will enable them to do well in a competitive economy. But in doing this they will be contributing to injustice, by enabling their children to prevail over poorer children who are equally or even better qualified.

Olsaretti turns this into an argument against (one source of) economic inequality by combining it with Cohen's argument, and adding two further elements to Nagel's description of the predicament of richer parents. The first element is the plausible claim that parents not only desire to give their children better lives. As parents they also have an obligation to do so. The second element is the assumption that the rich parents are better off because their higher pay is justified on the ground that it is needed as an incentive to get them to take certain jobs and work hard in them, with resulting benefits to the worse off.

It follows that richer parents who are concerned with justice have two reasons not to demand these incentives. One is the reason that Cohen cites, that forgoing these incentives will further benefit the worse off. But in addition, Olsaretti

⁷ Nagel 1991, pp. 110–111.

argues, there is the fact that insofar as they accept these payments it will become unjust for them to fulfill their obligation to do the best they can to help their children develop valuable talents. Insofar as they take their parental obligation seriously, they will want to be able to fulfill it without acting unjustly.

Understood in one way, Olsaretti's argument, like Cohen's, operates within Rawls' framework, showing that conscientious parents cannot be motivated in the way they would need to be in order for incentive payments to be justified by the Difference Principle. But her argument can also be understood in a broader way. Like Nagel, Olsaretti calls attention to the fact that economic inequality gives richer parents the ability to give their children special care and education that will give them unjust advantages over poorer children. It thus presents these parents with a conflict between their sense of justice and their obligation to do the best they can for their children.⁸ This conflict can be seen as a direct *pro tanto* objection to economic inequality: namely that it puts richer parents in the position that fulfilling their obligations as parents involves behaving justly. This objection is more general than the one Olsaretti raises since it applies to inequalities that enable parents to give their children better educations whether or not these inequalities are economic and whether or not they are justified by the supposed need for incentives. But her argument is distinctive in that it does not merely call attention to an objection to economic inequality, based on its consequences, but attacks the legitimacy of economic inequalities by undermining one possible justification for them.

Response to Derek Parfit

I have responded in my own contribution to the symposium to a number of the points that Derek Parfit raises. I will add here replies just to two further issues.

The first of these is the problem for contractualism presented by Parfit's non-identity problem. I see two lines of possible response to this challenge. One, the pluralist response, is to concede that contractualism cannot explain the wrongfulness of a policy that makes people who live at some future time much worse off than the people who would have lived at that time if some alternative policy had been followed. According to the pluralist response, the wrongfulness of such a policy is to be explained simply by the fact that it brings about a situation that

⁸ As Nagel pointed out (*Ibid.*) this conflict is not merely a problem for the parents. It is also a conflict between the needs of justice and the motivation to help their children that society must rely on parents to have.

is much worse than the situation produced by such an alternative. Although this account delivers the correct verdict, that adopting the harmful policy would be wrong, it does not explain the sense we have that people who, as a result of this policy, exist in the future and have bad lives have been *wronged*.

The alternative line of response is to show that this wrong can be accounted for within contractualism, properly understood. The first step in this direction is to address the question of which individuals it is whose reasons for objecting to a principle can make the actions that principle would allow wrongful. The answer is that the objections that are relevant in the process of contractualist justification are not objections of particular individuals. Rather, they are reasons that any individual would have in virtue of being in a certain position, namely the position of being affected in a certain way by a principle—affected by being required to comply with the principle or being affected by other people's acting in ways that that principle would permit. But a person cannot “be in a certain position” unless that person exists. So the question is whether people in a certain position, if there were any, would have reason to object to a certain principle. The position in question is that of living under conditions that are very bad in some way, due to policies that this principle would allow (for example, bad because of pollution, increased temperatures, or high sea levels) but who would not have existed if some alternative policy had been followed. It is not obvious to me that people in this position do not have such an objection, although I do not have a worked out view of the matter. If they do not have an objection, then the former, pluralist view would seem to me the most plausible account. In that case, the main possible deficiency of this account would also be undermined, since the fact that these future people would have no complaint would strongly suggest that they are not wronged.

I turn now to Parfit's question about whether contractualism is best understood as an account of what it is for an act to be morally wrong, or as an account of what makes actions morally wrong. I have claimed the former, and he has argued that this is a mistake. To make clear why I do not agree, it will help to begin with a few general points about the objectives of a moral theory.

A moral theory offers an interpretation of the subject matter that our ordinary moral beliefs are beliefs about. We may believe firmly that certain actions are morally wrong, and that if an action is morally wrong then we have strong, normally conclusive reasons not to do it. But it is not clear exactly which actions are morally wrong or what reasons we have not to behave wrongly. So, as I say in my paper in this volume, a moral theory seeks to answer two questions about morality: the question of *content* and the question of *acceptance*. It seeks to explain which actions are wrong, and what makes them wrong, and to explain why moral wrongness is important and something that we have strong reason to

avoid. Since our pre-theoretical beliefs about morality are inchoate and sometimes conflicting, a moral theory is bound to be in some ways reformist. Plausible moral theories will vary in the degree to which this is so. But in order to be an interpretation of the subject matter of our pre-theoretical beliefs, a theory needs at least to account for the things that seem most clearly to be morally wrong and to provide a plausible account of the distinctive importance that moral rightness and wrongness seem to have.

With these general points as background, consider Parfit's question about whether contractualism is an account of what makes acts wrong or an account of what wrongness is. Answers to the question of what makes a certain action wrong are often quite specific, referring, for example, to the consequences of that particular action. One might say, for example, that what made an action wrong was that it caused serious, easily avoidable harm, or that it led to serious foreseeable harm without sufficient justification. Other actions will be morally wrong for different reasons. Parfit believes I should also say that what makes the action wrong is that any principle that permitted it could, for that reason, be reasonably rejected. This property is much more general, applying to many actions that, I would say, are wrong for different reasons. This generality makes it suitable to be an account of what wrongness is that serves the purposes of a moral theory.

In order to answer the question of content a theory needs to offer a general characterization of the subject matter of morality in terms of which we can explain which particular actions are morally wrong and why. Different moral theories, such as forms of utilitarianism, contractualism, and Ross's intuitionism offer different characterizations, but they are all *general* characterizations of moral right and wrong. That is, they characterize a property of wrongness that different actions have for different reasons. Some actions are wrong because they cause serious avoidable harm, others because they involve violating a promise, others because they are invasions of a person's privacy. But they are all wrong in the same sense. This is why a general characterization of this kind is plausibly called an account of what wrongness is, rather than an account of what makes actions wrong. This generality is also important for an answer to the question of acceptance, which explains the reason we have for caring about whether our actions are wrong or not. These reasons, as revealed by what I called the remorse test, are common to acts that are wrong, in the same sense, for different reasons. I take contractualism to be an account of what wrongness is because it aims to be a general account of moral wrongness that provides answers to these two questions.

Parfit offers several arguments against taking contractualism to be an account of what moral wrongness is and for instead taking it to be an account

of a property that makes actions wrong in some other sense. The first of these begins with his statement of what he calls *Scanlon's Formula*:

Scanlon's Formula: An act is wrong just when such acts are disallowed by some principle that no one could reasonably reject.⁹

Parfit asks how the term 'wrong' in this formula should be understood. He says that if it means "wrong in the contractualist sense" then this formula is a thinly disguised tautology. To avoid being trivial in this way, he suggests that the formula should be understood as claiming that when actions are, as he puts it, disallowed by some principle that no one could reasonably reject this makes them wrong in some other sense. The candidates he mentions for this other sense of wrongness include being unjustifiable to others, being blameworthy, and being an act that gives its agent reasons for remorse and gives others reasons for indignation.

But this is not the only way that *Scanlon's Formula* can be understood as making a claim that is not trivial. As a moral theory, contractualism is an interpretive thesis about the subject matter of our ordinary moral beliefs. *Scanlon's Formula* is plausibly understood as an interpretive claim of this kind. It claims that what we ordinarily think of as moral wrongness is best understood in the way that contractualism proposes: that the contractualist account of wrongness captures an important subclass of the cases that are commonly regarded as morally wrong and that this characterization provides a plausible account of the significance for us of the fact that these acts are wrong. Understood as an interpretive claim of this kind, the formula escapes Parfit's charge of being a tautology. But a related charge that it involves a concealed tautology is still lurking.

When I first put forward contractualism as a moral theory, I proposed it as an account of "morality" in a general sense. I later recognized that this was a mistake. It is plausibly said to be morally wrong to destroy great works of nature for trivial reasons, or to be indifferent to the value of developing one's talents, or to fail to take care of one's children. But none of these things is plausibly understood as morally wrong because it would be reasonable to reject principles that would permit it.¹⁰ Recognizing this, I shifted to a narrower claim, that contractualism captures only a subpart of what is commonly called morality, a part that

⁹ I have stated the formula as Parfit does. See Parfit, "Improving Scanlon's Contractualism," this volume. I would, however, prefer the formulation, "An act is wrong just when any principle that permitted such acts could, for that reason, be reasonably rejected." I will interpret it in this way in what follows.

¹⁰ See Scanlon 1998, Chapter 4, section 7, pp. 171 ff.

I named “what we owe to each other.”¹¹ This might be seen as shifting my claim to a tautology—that contractualism captures the class of actions that are “wrong in the contractualist sense”—while attempting to conceal this by inventing a new name for this subpart of morality. My defense against this charge is that on reflection, applying the remorse test, it is clear that the moral faults I have recognized as not explained by contractualism represent distinct forms of moral wrongness, which we have different reasons to care about and feel remorse about violating.

Parfit also argues that if contractualism captures only one sense of moral wrongness among others this undermines the possibility of genuine disagreement between contractualists and those who hold other views. If my claim about an action is just that it is wrong in the sense contractualism describes (even if I give this sense another name, such as “what we owe to each other”) and a person who holds a different view claims that it is not wrong, we will not be disagreeing but only talking past one another. In order to keep alive the possibility of genuine disagreement, he believes, we must both be making a claim about wrongness in the same sense of “wrong,” and my claim must be that being allowed only by principles that it would be reasonable to reject *makes* the action wrong in this shared sense.

I believe, however, that genuine moral disagreement can be preserved without appeal to such a shared idea of wrongness, and that in many cases disagreement is best understood by recognizing that different values are being appealed to. As I have said, moral theories such as contractualism and utilitarianism can be understood as making interpretive claims about the best way of making sense of our ordinary moral beliefs. Insofar as this is so, utilitarians who claim that an action is wrong and contractualists who claim that it is not wrong can be disagreeing (among other things) about what we should recognize as being wrong in this everyday sense. If this is the case, being contrary to what would maximize utility and being permitted only by rejectable principles are not claimed to be properties that *make* an action wrong in this “ordinary” sense. Rather, each theorist is making a claim about how this ordinary sense of moral wrongness should be understood.

Contractualists and utilitarians can also be disagreeing in a different way insofar as each claims to have identified the form of moral wrongness that we have most reason to accept as an authoritative standard of conduct—as the moral

¹¹ It is an open and disputed question what idea of reasonable rejectability, if any, best captures this intuitive idea, or whether, as Johann Frick suggests, contractualism can only capture a part of it. See Frick 2015, pp. 219 ff.

theory for which we can give the best answer, or perhaps the only satisfactory answer, to the question of acceptance. What each of these parties claims for their view is not that it correctly describes “morality” in some independent sense, but rather just that it is the moral view that we have most reason to accept (or perhaps is the only one we have good reason to accept).¹²

A slightly different analysis applies to disagreements of the kind I have mentioned above involving forms of moral wrongness that are not violations of “what we owe to each other.” When someone claims, for example, that it is morally wrong to harm the environment even if this does not have harmful effects on human life, and someone, perhaps a contractualist, disagrees, their disagreement is obviously not about the content of “what we owe to each other.” Nor, I think, is it best understood as a disagreement about morality in some single more inclusive sense. This seems to me to be revealed, again, by the remorse test. The kind of distress that one party claims properly to feel for despoiling the environment in a way that has no ill effects on human life (or, we may suppose, even on the lives of other animals) is quite different from what we properly feel when we have injured a person, or failed to keep a promise, or refused to help someone in need. This seems to me to show that the claim being made is about the importance of a different form of value. The disagreement, if there is one, is not about contractualism, or about moral wrongness in a more inclusive unified sense, but rather about whether there is a distinct value of this kind, and if so how it is to be understood.¹³

Therefore, despite Parfit’s challenging objections, I remain convinced that contractualism is best understood as an account of a distinct sense of moral wrongness, rather than as an account of a property that makes actions wrong in some further sense.

Response to Zofia Stemplowska

Stemplowska’s challenging paper raises questions about my account of what I call substantive responsibility. The aim of this account is to explain how it can make a difference to the moral justifiability of a policy that that policy makes what happens to a person depend on how that person responds when presented with a choice that will determine what happens to him or her. The

¹² This appears to be what Bentham is claiming when he says, in his *Introduction to the Principles of Morals and Legislation*, that the terms ‘ought’ and ‘ought not’ make sense only when understood in the manner that utilitarianism offers. See Ryan 1987, p. 67.

¹³ The same is true, I believe, of some arguments about sexual morality. See Scanlon 2011.

guiding idea is that one reason people have for wanting outcomes to depend on their responses is that generally, although not invariably, this makes it more likely that the outcome will be one that they have reason to prefer. In particular, warning a person that a certain course of action will lead to an outcome she has reason not to want is one way of making it less likely that that she will suffer that result. The effectiveness of such warning for a given person will vary, depending on the conditions under which the choice in question would be made, and on that person's psychological makeup. Moreover, warnings are not the only way that people can be and should be protected against unwanted outcomes. As I indicate in my example of hazardous waste, there are often many other things that should be done—fences, careful removal procedures, etc. But a warning that gives people the choice of avoiding this outcome is one kind of protection that people generally have reasons to want.

Even if we had free will, our reasons of this kind to want what happens to us to depend on our choices would still vary, depending on the conditions under which these choices were made. So even with free will we would have reason to prefer that what happens to us should be made to depend on choices we make under better conditions. My main claims are that we have these same reasons even if the causal thesis is correct, and all our choices are ultimately caused by factors outside of us, and that these reasons fully explain the significance of choice in cases of what I call substantive responsibility.¹⁴ Stemplowska questions the latter claim.

What does it mean for a person to be “substantively responsible?” Stemplowska says at one point that “To be substantively responsible for an outcome is to have no complaint against others that the outcome arose.” I may have said things that suggest this, but as she points out it is much too broad. So some clarification is in order.

To begin with, there are cases in which a person “has no complaint” about some unwanted outcome the occurrence of which has nothing to do with any choice she made or could have made. This can be so simply because no one had any duty to do more to protect her against suffering that outcome. This may be the case in Stemplowska's example of the pregnant woman who suffers piles.¹⁵ The occurrence of this condition may have had nothing to do with any-

¹⁴ As I have argued elsewhere, most recently in Scanlon 2015, the significance of choice for the appropriateness of moral reactive attitudes is a different matter.

¹⁵ Stemplowska writes that she will discard the suggestion that having a good enough opportunity “is a necessary condition for substantive responsibility.” She goes on to offer the case of the woman with piles as supporting this. Since the woman in this example is pretty clearly not substantively responsible for suffering piles, I think that what Stemplowska may have meant

thing the woman did, so it would be odd to say that she is responsible for this outcome, even though it is true that she has no complaint against anyone about suffering it.

In order for a person to be substantively responsible for what happens to him, it must be the case, first, that this outcome depends on how the person responded or failed to respond when presented with a choice. Second, it must be true that enough was done to protect the person against suffering this outcome, including (but not limited to) making what happens depend on the person's response in this way. This "response" need not involve a conscious decision. I could be substantively responsible for failing to pick up my theater tickets half an hour before curtain time (as I was told when I bought the tickets), and the theater owner is therefore justified in selling them to someone else, even if I simply forgot to pick them up, and never made a decision not to do so.

As this example also indicates, we can have "done enough" to put a person in a good position to make a choice even if what we have done falls short of what would be required to ensure that the person chooses wisely. Whether we have done enough can depend, as in this example, on what is at stake in the choice, and on the costs to others of providing greater protection. In the case of the tickets, if the policy of selling tickets that are not picked up half an hour before curtain time is justified, this is because theater owners should not be required to forego these sales so that the absent-minded will not lose their claim to tickets. The stakes are higher in my hazardous waste case, but even so, as Stemplowska says, we do not need to put a 24 hour guard around each person's house to make sure that they do not go out.

Stemplowska's examples also bring out the need to distinguish clearly between the question of whether we have, overall, done enough to protect a person and the question of whether he was actually put in sufficiently good conditions to avoid harm by choosing appropriately. Just because a person chose badly under the conditions in which he was placed does not mean that others did not do enough to protect him by providing good conditions. It may be difficult to put everyone in a good position to make a choice, and there is a limit to how far we are required to go in order to do this. Moreover, the question of whether a person has, *overall*, any complaint about the various forms of protection he was provided with can come apart from the question of whether he is

(and what she goes on to say in that paragraph) is that being put in good enough conditions to avoid an unwanted outcome by choosing appropriately is not necessary in order for a person to have no complaint about suffering an outcome.

substantively responsible for the outcome. Both of these possibilities are brought out by the examples of Curious and Walker.

Curious was aware that the decision to go out involved a high risk of contamination. And she was put in a position to decide whether to go out under conditions that would have been, for most people, sufficient to make it very likely that they would not go out. (Although, as in the example of the tickets, people who have been warned do sometimes forget.) In Curious' case, however, the warning did not make this less likely. As the example of theater tickets indicates, a warning need not be successful in order for a person to be substantively responsible for an outcome. Conceivably, more could have, and should have, been done (short of perfect protection) to guard against foolish choices in this case. But unless that is so, it follows that Curious is substantively responsible for what happened to her, on my view.

Walker, on the other hand, never heard that going out would involve a risk of contamination. So his decision to go out was not made in response to being presented with this choice. It follows that he is not substantively responsible on my view. If, however, others did as much as they could be required to do in order to make him aware of the risk, then it may be that he has no complaint, even though he is not himself substantively responsible.

With regard to Curious, Stemplowska asks not only how we can say that "enough" was done for her, given that it did not prevent her from going out and becoming contaminated, but also how we can say that we have done anything at all to protect her against this result, given that the warning she received was ineffectual and even counter productive. I have already responded to the first of these questions, but the second is more challenging.

What was done for Curious that could be called a benefit to her? It was, I would say, being made aware of the choice she had and of relevant facts about the alternatives she could choose between. She was, as a result, better placed to decide whether to stay home or to satisfy her curiosity by going out than she would have been if she had been told nothing. It is true that if she had been told nothing she would, in the end, have been better off, but this is not because she would have been put in better conditions to make the decision. But why should one want to be, in this sense, in better conditions for deciding (more aware of the alternatives, etc.)? The answer on my view is that in general this will make it more likely that the outcome will be one that one has reason to prefer. The problem is that in Curious's case this was not true: being put in a position to decide whether to go out made her more likely to do go. So in what way was it a good thing, from her point of view?

We need to distinguish here between two kinds of cases. In the first kind of case, more common when we are choosing among large scale policies, we may

know that among those affected by a policy, there are a few who, like Curious, will choose badly if the outcome is made to depend on their responses. But we have no idea which individuals in particular are like this. In cases of this kind, looking at things *ex ante*, from behind what might be called a natural veil of ignorance, being put in a good position to decide what to do—in this case being warned of the danger of going outside—is what each person has reason to want, because it reduces the probability of suffering contamination.¹⁶

The problem that Stemplowska describes arises only in a different kind of case, in which we know, of particular individuals like Curious, that they will respond foolishly to a warning. Given this knowledge, we can no longer say that by putting them in a better position to decide what to do we are reducing the likelihood that they will suffer a bad outcome. This raises the question of whether more should be done to protect these people. As Stemplowska says, it would be more than we could be required to do to put a 24 hour guard outside every house. But if there are only a few people like Curious perhaps we should provide guards for them or, more realistically, give them a more effective kind of warning. Stemplowska's argument as I understand it, uses the question of whether this should be done, and at what cost, as a way of raising the question whether the fact that Curious has already been "given a choice" has greater significance than my view can explain.

Simplifying her argument slightly, suppose that a proposed way of warning people about the dangers of hazardous waste removal will fail to reach a certain number of people like Walker, and fail to dissuade an equal number of people like Curious, who will hear of the project but go outdoors anyway to see what is happening. Suppose that, given available resources, we could either expand the current warnings to reach those in Walker's position or make the current, more restricted warnings more effective, thus deterring people like Curious. It may seem that we should do the former. Why might this be so?

One answer is that choosing the latter option would involve spending more resources on protecting Curious and those like her than on protecting those in the position of Walker, thus giving greater weight to the interests of those in the former group. This would be in line with the point made earlier, that whether we have "done enough" to protect people against some unwanted outcome depends not just on how much good we have already done them but also on the costs to others of providing greater protection.

The alternative explanation, which Stemplowska presses, is that the fact that Curious has been given the choice of avoiding contamination has moral signifi-

¹⁶ I take this idea, and the phrase, 'natural veil of ignorance,' from Frick 2015, pp. 175–223.

cance that is independent of any benefit to her of having this choice and of the cost to us of providing it, and is therefore not captured by the account I have offered. I do not agree with this conclusion, although I do feel its force. I would make two observations about it.

The first is that although Stemplowska presents the problem as arising from my aim of explaining the significance of choice in a way that is compatible with the causal thesis, this seems to me not quite right. The fundamental question is whether the fact that a person was given the choice (under certain conditions) of avoiding harm has intrinsic moral significance that is independent of the benefits to a person of being given such a choice. It is a further question whether having a choice could have this kind of significance if the way we respond to the choices we are presented with is, ultimately, caused by factors outside of us. I myself doubt that this is true. But even if it is not true, examples seeming to show that having a choice in this sense (that is, having what happens to a person depend on how she reacted when presented with the choice under “good” conditions) has intrinsic moral significance of this kind would present a problem for my account of substantive responsibility.

My second observation is that what would need to have this intrinsic significance is not choices that individuals make (the engagement of their will, so to speak) but rather the circumstances in which they are placed—the fact that they “*had* a choice.” This seems to me to lend support to my view that substantive responsibility is more a matter of how much others have done rather than the degree to which this has, in the end, benefitted a particular recipient.

Insofar as the alternative view sketched by Stemplowska in section 5 of her paper puts the emphasis on fair sharing of costs, it is in line with what I have just said. But insofar as it makes conclusions about a person’s substantive responsibility depend on his or her motives in “courting harm” I would resist it. First, it does not seem to come to grips with the problem presented by Curious and Walker, which seems to me to be independent of any assumptions about Curious’s attitudes. Second, making a moral assessment of a person’s motives relevant to judgments of substantive responsibility invites a kind of moralism in discussions of public policy that I think it is important to avoid.¹⁷ That said, however, I would repeat that Stemplowska’s examples present a serious challenge to my view of substantive responsibility, and I am not confident that I have responded to it adequately.

17 As I argue in Scanlon 2018, pp. 60–63, 72–73.

References

- Frick, Johann, "Contractualism and Social Risk," *Philosophy & Public Affairs* 43 (2015).
- Rawls, John, *A Theory of Justice* (Cambridge, MA: Harvard University Press, 1971).
- Ryan, Alan, ed., *John Stuart Mill and Jeremy Bentham. Utilitarianism and Other Essays* (London: Penguin Books, 1987).
- Cohen, Gerald A., "Where the Action Is: On the Site of Distributive Justice," *Philosophy & Public Affairs* 26 (1997), pp. 3–30.
- Nagel, Thomas, *Equality and Partiality* (Oxford: Oxford University Press, 1991).
- Waldron, Jeremy, *The Harm in Hate Speech* (Cambridge, MA: Harvard University Press, 2012).
- Scanlon, Thomas M., *What We Owe to Each Other* (Cambridge, MA: Harvard University Press, 1998).
- Scanlon, Thomas M., "What is Morality?" in *A Harvard Sampler: Liberal Education for the Twenty-first Century*, Jennifer M. Shepard, Stephen M. Koslyn, and Evelyn M. Hammonds, eds. (Cambridge, MA: Harvard University Press, 2011), pp. 243–266.
- Scanlon, Thomas M., *Being Realistic about Reasons* (Oxford: Oxford University Press, 2014).
- Scanlon, Thomas M., "Forms and Conditions of Responsibility," in *The Nature of Moral Responsibility: New Essays*, edited by Randolph Clarke, Michael McKenna, and Angela M. Smith (Oxford: Oxford University Press, 2015), pp. 89–111.
- Scanlon, Thomas M., *Why Does Inequality Matter?* (Oxford: Oxford University Press, 2018).

Susanne Mantel

Words of Thanks

I am deeply grateful, and feel greatly honored, to receive the Lauener Award for young philosophers. I am particularly thrilled to be honored together with Tim Scanlon and to participate in this remarkable symposium on his lifetime achievements, since he has deeply inspired and influenced me ever since I started to do philosophy. To let you see just how true this is, I would like to tell you a little story.

When I was a student at Tübingen University, I started to take an interest in analytic philosophy and in particular in reasons for actions. I was very lucky to have some teachers who shared these interests and I quite eagerly attended the analytic courses they gave and the conferences they organized. When one day we were told that even students could organize little workshops, some other students and me founded a little group with the aim of inviting, every year, one of our most favorite analytic philosophers. Of course we were not sure whether we would succeed, but with a lot of support we raised the necessary money. Each year we invited one of the authors whom we particularly admired – and, obviously, Tim Scanlon was one of them. It is hard to describe what it felt like to write an invitation to him, being nothing more than students. Compare it to writing to the author of your favorite novel. It felt like inviting the world famous author of our favorite novel to fly over to our place and meet us, just because we loved his books and had a lot of questions about the plot.

You will be able to imagine how enthusiastic and excited we were when Tim accepted our invitation to come to the – quite tiny – town of Tübingen. To be more precise, we almost burst with joy and pride to be his hosts, to attend his lecture, and to give our little student presentations to him, to which he responded in an incredibly patient and thoughtful way. His deep and selfless commitment to us became especially apparent when Tim joined us on a punt boat trip on the river, which was remarkable since it was a very shaky boat and he suffered from a heavy pain in his back from sleeping in an uncomfortable position on the airplane.

To sum up: He flinched from nothing and was the friendliest and most encouraging teacher we could have imagined. The graduate conference with him was one of the greatest sources of inspiration and encouragement for me as a student, and, as it turns out, the themes from the conference with Tim have never left me: Their deep influence on me is evident in almost any philosophical text that I write today.

<https://doi.org/10.1515/9783110733679-013>

Let me end this story by saying that because of these experiences I think I very well understand what Prof. Lauener's engagement with analytic philosophy must have meant to his students in Switzerland when he enabled them to study analytic philosophy and, for instance, when he organized his famous international congresses. He thereby gave his students the chance to meet the authors they were working on and to have inspiring discussions with them, and I know how important such chances are. Notably, he did so already at a time when analytic events were much rarer in Europe than at the time when I was a student.

Now that my little story is over, I would like to cordially thank all the members of the Lauener Foundation for this great honor, especially the president Mr. Føllesdal and the managing member, Mr. Frauchiger, who took so much care in preparing this event.

Heartfelt words of thanks go also to my supervisors, Prof. Fehige and Prof. Wessels from Saarland University. They always believed in my work – at times much more than I did myself – and I can hardly express how much they supported me and how deeply they engaged with my thoughts. I am very grateful to my colleagues from Saarbrücken, who are truly wonderful, and to countless philosophers from other universities, as well, for invaluable discussions and feedback. Special thanks go to my husband, Frank Hofmann, also a philosopher, who must often endure to be the first person I test my philosophical ideas on, and to my two daughters, who, along with my husband, give me so much love and joy.

Contributors

Rainer Forst is Professor of Political Theory and Philosophy at Goethe University Frankfurt. He has been Co-Founder and Speaker of the Research Center “Normative Orders” and of the Center for Advanced Studies “Justitia Amplificata.” His research focuses on questions of justice, democracy and toleration as well as critical theory and practical reason. In 2012 he was awarded the Gottfried Wilhelm Leibniz Prize of the German Research Foundation. He is a Member of the Berlin-Brandenburg Academy of Sciences and Humanities and Corresponding Fellow of the British Academy. His most important publications are *Contexts of Justice* (2002), *Toleration in Conflict* (2013), *The Right to Justification* (2012), *Justification and Critique* (2014) and *Normativity and Power* (2017); forthcoming is *Die noumenale Republik*.

Michael Frauchiger, Dr. phil., has been lecturing in the fields of philosophy and methodology of science and the humanities with the University of Applied Sciences of Zurich, the Open University and the University of Bern, and additionally has been an Advanced Research Fellow of the Swiss National Science Foundation. Since its establishment in 2003, he has been the Managing member of the board of trustees of the Lauener Foundation for Analytical Philosophy. His research cuts across epistemology, philosophy of language and logics, metaphysics, philosophy of religion, philosophy of action, ethics and political philosophy.

Susanne Mantel, Dr. phil., studied philosophy, German literature, and psychology at the University of Tübingen. She was a visiting scholar at the University of Berkeley and attained her doctoral degree at the Saarland University. For her thesis “Determined by Reasons: A Normative Competence Account of Acting for a Normative Reason”, which was published with Routledge, she received the Wolfgang-Stegmüller Prize of the German Society for Analytic Philosophy. Her research focuses on the theory of action, rationality and reasons, and normative theory in general, as well as on political and applied philosophy.

Thomas Nagel is University Professor Emeritus at New York University. He was trained in philosophy, first at Cornell (B.A. 1958), then at Oxford, where he went on a Fulbright Scholarship (B. Phil. 1960), and then at Harvard (Ph.D. 1963). He is the author of *Mortal Questions* (1979), *The View From Nowhere* (1986), *Mind and Cosmos* (2012), and other books.

Serena Olsaretti is ICREA Research Professor, Department of Law, Universitat Pompeu Fabra. She has published articles in political philosophy in various journals, including *Analysis*, *Economics & Philosophy*, the *Journal of Applied Philosophy*, the *Journal of Political Philosophy*, *Philosophy & Public Affairs* and *Utilitas*. She is the author of *Liberty, Desert and the Market* (Cambridge University Press, 2004) and the editor of *Desert and Justice* (Oxford University Press, 2003), *Preferences and Wellbeing* (Cambridge University Press, 2006) and *The Oxford Handbook of Distributive Justice* (Oxford University Press, 2018). Her recent work focuses on questions concerning justice and the family.

Derek Parfit was a fellow of All Souls College, Oxford, where he taught throughout his career. He is the author of *Reasons and Persons* (Oxford: Oxford University Press 1984) and the

three-volume work *On What Matters* (Oxford: Oxford University Press 2011–2017). The last volume appeared after his death on January 1st, 2017.

Thomas M. Scanlon is Alford Professor of Natural Religion, Moral Philosophy, and Civil Polity, *emeritus* at Harvard University. He received his B.A. from Princeton in 1962 and a PhD from Harvard in 1968. He has written widely on topics in moral and political philosophy. His books include *What We Owe to Each Other* (Harvard University Press, 1998) and, most recently, *Why Does Inequality Matter?* (Oxford University Press, 2018).

Zofia Stemplowska is Professor of Political Theory at the University of Oxford and Asa Briggs Fellow of Worcester College, Oxford. Her research interests include domestic, global and historical justice. She is the co-editor (with Carl Knight) of *Responsibility and Distributive Justice*, Oxford University Press (2011).

Markus Stepanians is Professor of Political Philosophy at the University of Bern, Switzerland. His early interests were in the Philosophy of Language, Logic, and Husserlian Phenomenology. He is the author of *Frege and Husserl on Thinking and Judging*, Mentis 1998 (in German) and *Frege—an Introduction*, Junius 2001 (orig. in German; French and Spanish translations 2007). He is the editor of *Individual Rights*, Mentis 2007 (in German) and co-editor (with B. Kahmen) of *Critical Essays on 'Causation and Responsibility'*, De Gruyter 2013.

Index

- Anderson, Elizabeth 12
Ansccombe, Gertrude Elizabeth Margret 30
Archard, David 102
Aristotle 2, 19, 89
Arneson, Richard 101
Ashford, Elizabeth 23
Ashwell, Lauren 68
- Benacerraf, Paul 12
Bentham, Jeremy 147
Bird, Alexander 68
Björklund, Fredrik 62
Björnsson, Gunnar 62
Brake, Elizabeth 102
Bridges, Jason 75
Brink, David 62
Broome, John 62, 64, 68, 75
- Clarke, Randolph 68
Cohen, Gerald A. 53, 93, 97–99, 104–106, 140–142
Cohen, Marshall. 12
- Daniels, Norman 91
Davidson, Donald 63, 136
Dominey-Howes, Dale 120
Dreben, Burton 12
Dreier, James 73, 75
- Enoch, David 49
Eriksson, John 62
Estlund, David 106
- Fehige, Christoph 156
Fishkin, Joseph 99
Føllesdal, Dagfinn 11, 156
Forst, Rainer 11, 45, 47, 49f., 52f., 55, 131–134
Frankfurt, Harry G. 63
Frauchiger, Michael 11, 156
Frick, Johann 27, 30, 34, 36f., 123, 146, 151
- Gaut, Berys 64
Goff, James 120
Gregory, Alex 62
- Hempel, Carl Gustav 63, 65
Henry (Tudor) the Eighth 88
Hofmann, Frank 156
- Johnston, Mark 68
- Kamm, Frances 27f., 34
Kant, Immanuel 11, 14f., 38, 45–47, 50f., 87, 89, 109
Kolodny, Niko 64
Korsgaard, Christine 46
Krich, Jacob J. 120
- Lauener, Henri 156
Lewis, David 68
- Mackie, John Leslie 60
Mantel, Susanne 11, 59f., 67f., 70, 131, 134–136, 155
Martin, Charles Burton 68
Mill, John Stuart 13f., 87–89
Millgram, Elijah 71
Moller Okin, Susan 101
More, Sir Thomas 88
Munoz-Dardé, Veronique 30
Murphy, Jeffrie 18
Murphy, Liam 39
- Nagel, Thomas 1f., 12, 17, 22, 62, 83, 131, 136–139, 141f.
- Olinder, Ragnar F. 62
Olsaretti, Serena 91, 102, 131, 140–142
Olson, Jonas 60
O’Neill, Onora 102
- Parfit, Derek 12, 17, 23, 25f., 29–32, 34f., 37–41, 61f., 109, 131, 142–147
Porter, Lindsey 102

<https://doi.org/10.1515/9783110733679-015>

- Rawls, John 1f., 11–13, 17, 20, 50, 52f.,
55, 91, 94–97, 99f., 102, 106, 114, 133,
140–142
- Rousseau, Jean Jacques 1
- Ryan, Alan 147
- Scanlon, Thomas M. 1–9, 11–18, 20, 22,
25, 30, 45–55, 59, 61–67, 70–73, 75f.,
78f., 87f., 92–97, 99f., 104–106, 109–
117, 119–126, 128f., 131, 135f., 140,
145, 147f., 152, 155
- Scheffler, Samuel 12, 19, 129
- Sen, Amartya 17
- Singer, Peter 7, 27, 38
- Smith, Michael 61, 64, 66, 68, 71, 74
- Stemplowska, Zofia 119f., 123, 129, 131,
147–152
- Stepanians, Markus 1
- Stocker, Michael 63
- Tadros, Victor 27
- Thomson, Judith Jarvis 62
- Tresan, Jon 73
- van Roojen, Mark 60, 74
- Voorhoeve, Alex 15, 45, 123
- Waldron, Jeremy 133f.
- Wedgwood, Ralph 68
- Wigner, Eugene 4
- Williams, Bernard 17, 38, 68, 89, 91
- Zangwill, Nick 73, 75
- Zhou Enlai 9