Social morality and the primacy of individual perspectives

Gerald Gaus

Abstract: This paper examines themes and concerns about my book, *The Order of Public Reason*, raised in the three essays in this symposium by Peter Boettke & Rosolino Candela, Michael Munger and Kevin Vallier. The three essays present variations on a common theme: I need to embrace deeper commitments than *The Order of Public Reason* acknowledges. In my estimation these proposals lead to places that I do not wish to go — nor should anyone devoted to core Hayekian insights. The goal of the book is show how a diversity of moral views can lead to a cooperative social morality while abjuring as far as possible "external" moral claims — claims that do not derive from the perspectives of cooperating individuals. The diverse individual moral perspectives, and what *they* understand as normative, must be the real engines of social normativity. In this essay I stress the primacy of the individual normative perspectives in generating social morality; this helps show why the urge to embrace deeper commitments should be resisted. Rather than going over the presentation in *The Order of Public Reason* to stress this point, I sketch a modest recasting of the analysis in terms of models of individual moral interaction.

1 THE INDIVIDUAL AND THE PHILOSOPHICAL PERSPECTIVES

I am especially grateful to the *Review of Austrian Economics* for this opportunity to reflect on themes in, and questions about, *The Order of Public Reason* (*OPR*) more than

Review of Austrian Economics
DOI 10.1007/s11138-016-0358-8

five years after it was published. It is gratifying to know that anyone is still interested in the book, and it is especially gratifying to know that those who, like me, appreciate the brilliance of Hayek are still interested. It is also welcomed since, at this point, I am able to stand back and better appreciate how readers view the book. I can now better see what worked well and, well, what did not.

I hope it is not a sign of sheer hardheadedness, but I remain convinced of the core claims of the analysis. I now, however, see different ways to put some of the points — ways that I think help avoid interpretations the text may bear, but which I think lead in wrong directions. *OPR* sought to challenge many of the presuppositions of ethics and social philosophy, yet employed many of the tools ("deliberative models") and concepts ("public justification," "self-legislation," "moral authority") of moral and political philosophy to do so. My audience was primarily philosophers of social ethics and political philosophy, and I self-consciously sought to stress the way my view was continuous with some of the core philosophical traditions — especially Kantian, contractarian, and Rawlsian. I believe this was accurate, necessary and helpful for wider understanding, but it also tended to invite readers to think the view was more Kantian, or more Rawlsian, than it is. It is certainly far more Hayekian than Kantian or contractarian, as Vallier shows. This journal is the perfect forum to stress this point. It is thus especially useful for me to reflect on the interpretations advanced by these three excellent essays, by Peter Boettke & Rosolino Candela, Michael Munger and Kevin Vallier, and see if I can reframe some of these matters in helpful ways, and then see where the conversation might lead.

The three essays present variations on a common theme: I need to embrace more extensive commitments than *OPR* acknowledges. Boettke and Candela suggest a deeper ethical foundation, Munger insists that the analysis supposes a "Kantian Parliamentarian," and Vallier argues that the problem of "global" justification may call for devices such as a civil religion. In my estimation these proposals lead to places that I do not wish to go — nor should anyone devoted to core Hayekian insights. As I stress in my recent *The Tyranny of the Ideal* (2016), the guiding aim is to sketch a framework of moral accountability and cooperation that makes room for the widest possible diversity of moral views, yet which all can see as supporting moral claims on each other in a system of cooperation. Consequently, the goal is to show how a diversity of moral views can lead to a cooperative social morality while abjuring as far as possible "external" moral claims — claims that do not derive from the perspectives of cooperating individuals. The diverse individual moral perspectives, and what *they* understand as normative, must be the real engines of social normativity.

The importance of this was not sufficiently stressed in *OPR,* though I certainly thought it. I distinguished there two perspectives, that of philosophical reflection and that of real moral agents (*OPR*: 266), but this fundamental distinction is underdeveloped. The critical idea was that the normative basis of our shared morality is to be grounded not in the "Archimedean perspective" (Gauthier 1986, 233; Gaus and Thrasher 2016) that reveals correct principles of morality independent of the perspectives of those in a practice of social morality. *OPR* seeks to avoid appeal to any such transcendental source of moral claims and demands. The "normativity" that exists in a system of social morality comes from the normative commitments of the participants.

Yet *OPR* is, after all, a work in moral and social philosophy, and develops a theory of a justified social morality. The claims of *OPR* itself are made from a philosophical or

theoretical perspective, which takes up issues that are, as it were, *about* the nature of social morality but not *in it*. In itself this is not unusual, being characteristic of many theoretical studies in the social sciences and humanities. Economists, for example, seek to better understand the features of systems of consumers and producers — features such as supply and demand and the division of labor — that the participants did not intend to produce and with which they may well be uninterested. This idea is also the core of Hayek's (1988: 6) analysis of morality: "To understand our civilisation one must appreciate that the extended order resulted not from human design or intention but spontaneously: it arose from the unintentional conforming to certain traditional and largely *moral* practices, many of which men tend to dislike, whose significance they usually fail to understand, whose validity they cannot prove, and which have nonetheless fairly rapidly spread by means of an evolutionary selection …." Like Hayek, *OPR* understands social morality as undesigned, arising out of individual intentional moral decisions. Thus it asks, given the normative commitments of individuals holding a certain array of normative views, under what conditions can an entire system of individual interactions reach a moral equilibrium on some rule *R*, such that *R* provides all with shared empirical and normative expectations and grounds a practice of accountability — always from the normative view of each and every participant?

In understanding *OPR*, then, it is critical to keep the two perspectives distinct. The system of interactions based on diverse individual perspectives gives rise to a social morality; the philosophical perspective (the perspective we theorists take up), studies and then tests the system, drawing on concepts such as moral accountability and public justification. The normative task of the philosophical theory is to show that ideas like social cooperation, moral accountability and public justification are attractive and powerful at the philosophical level, so that critical observers will employ these concepts to judge the system. This results in an important, but much more modest role, for the "moral point of view" than in most moral and political philosophy — from Kant to Rawls, from Smith to Sen, from Locke to Nozick. In these traditions almost all normative results derive from the proper specification of the philosophical, "objective," view which in some sense leaves behind, or radically abstracts from, individual perspectives, and then constructs true morality from some favored position. We might say that on this more traditional view the philosophical perspective is the fountainhead of our knowledge of morality. *OPR* certainly does not reject the moral point of view; some notion of objectivity is critical to all moral thinking. According to *OPR* the moral point of view is a perspective you and I as critical observers take up when we ask "Does this moral rule merit our endorsement? Does it pass the test that shows it to be acceptable or is it based on oppression, ignorance and false consciousness?" The system of individual perspectives generates a social morality, but moral agents always have the ability to reflect on it. Only once we have a theory of the unintended system of social morality can we do this; the point of having a philosophical perspective is to allow this sort of evaluation. Here, perhaps, *OPR* departs from Hayek, or at least some of the more extreme remarks in *The Fatal Conceit* (1988: 68 ff.) where he suggests such reflection can get little grip, as no theory can sufficiently understand the normative system. *OPR* rejects both the philosopher who would legislate morality from her philosophical theory, and the quietism that says we must accept the morality that has evolved.

In this essay I wish to stress the primacy of the individual normative perspectives in generating social morality; this will motivate why I think the urge to embrace deeper commitments should be resisted. Rather than going over the presentation in *OPR* to emphasize the individual viewpoint, I shall sketch a modest recasting of the analysis in terms of models of individual moral interaction. The aim is to keep our eye on the primacy of the individual perspective, and see just where the philosophical perspective enters in. As we do that, I believe that many of the queries and concerns of these three insightful essays can be addressed more clearly. I make no attempt to show that all of what I say here can be translated into the language of *OPR*, though I believe that there are no significant alterations in content.

2 MODELING MORAL AGENTS

2.1 The basic idea of moral maximizers

We begin with an assumption of a finite, large system *S* of moral agents {*A, B, C*…}. To say that they are moral agents is to say that: (*i*) each possesses evaluative or normative standards {A_Σ, B_Σ, C_Σ…}; (*ii*) each applies those standards such that she can rank options in terms of them; (*iii*) because they are intelligibly described as moral, everyone is competent at using those standards when choosing among alternative moral rules; and (*iv*) facts (i)-(iii) are known and accepted by all in *S*. They are not under a Rawlsian veil of ignorance (Munger 2016): each knows everything about her situation, all facts and values. And there is no assumption that they care about the same thing, except insofar as they can make the distinction between considerations that are relevant to moral evaluation and those that are not (and on this they may disagree, but each has come to some account of it — more on this anon), and we model all as acting on the former rather than the latter. More formally, the assumption is that they act on a partial utility function, defined in terms of the moral utility given by acting on some moral rule. But the arguments (based on evaluative standards) in their moral utility function can be entirely different.

2.2 Idealizations in the model

As Vallier would point out, there are important idealizations thus far. We have assumed that people know what they normatively care about, are effective at ranking options in terms of this, and that they act on this ranking. These idealizations are justified by the concerns of the philosophical perspective: our question is whether, and under what conditions, people who have deep normative disagreements can live according to freely shared moral rules — to answer that, we must focus on agents who act on their moral views. The question we are interested in answering always shapes the way we model individuals.

There is, though, a deeper idealization in the model: we suppose that each recognizes others as competent moral agents, and thus they implicitly share a common standard of competency. This is a critical point, and I think is at the heart of some of Munger's observations. *OPR* (279–83) advances a criterion of intelligibility; *A* must view B_Σ as *intelligible standards* given the problem at hand, i.e., evaluating whether

other moral agents can embrace the same moral rules as he does. If he cannot see B_Σ as intelligible, then he cannot see *B* as another moral agent with different views, as she is not really acting as a moral agent at all. Maybe she is too incompetent to understand what a moral rule is, or perhaps is someone who simply refuses to play the moral game, with her sole normative criterion being "what's in it for me?" In order to clarify the notion of competency, *OPR* (294–303) endorses a traditional list of necessary conditions for any rule to be moral: it must be public knowledge that it is a rule, and can be taught as such; it must resolve conflicting claims; it must be understood as issuing requirements which are typically decisive considerations for agents; it must be reversible in the sense that, should it identify different roles, a person endorses it regardless of the role she will occupy; and the rule must not be detrimental to anyone's basic interests. The model supposes that at least a large proportion of *S* will not endorse a proposed moral rule that fails to meet these standards.

2.3 What is objective and what is subjective?

Now if the general idea of moral competency qua intelligibility, combined with the articulation of necessary conditions for *R* to be a bona fide moral rule, are understood as dictates from the objective, Archimedean, perspective, then Munger must be correct that "Gaus requires … that the moral agent has *sufficient* reasons [to endorse a rule]. This sufficiency is objective, not subjective. It is not up to the individual to decide if the reasons are sufficient; they are, or they are not. And that is what is meant by the *adequate employment of reason*" (Munger 2016). The objective point of view determines what the reasons for agents are, and what constitutes a competent moral agent.

The approach of *OPR* (244–58) was indeed to employ the philosophical perspective to identify a minimal theory of competent reasoning, but it was supposed that members of *S* share it: it is confirmed by their individual perspectives. The main justification of this was a claim that this standard of competency approximates the actual practice of morality that our theory is trying to explain. As a popular practice it does not set the bar for competency high. To be sure, this is a commitment of the philosophical perspective about the subject matter. The aim is to explain social morality, and that has within it certain standards of competency. Moreover, unless some constraints are placed on what counts as competent reasoning the account will be of little philosophical interest. If our theory seeks to answer the question "can moral agents who disagree live together under freely shared moral rules?" we need to specify some plausible notion of what constitutes a minimally competent moral agent to get the project going. Assume a society of fools, all of whom believe that moral competency involves consulting their Ouija boards and who hold that acting morally caters to the desire to enslave each other, and it will be hard to make much headway. It won't explain the idea of social morality as we understand it.

Any theory that models subjectivity must, in an important sense, "objectively" define the phenomenon. In order to define a wide variety of views about matter *X*, it must delimit *X*-phenomenon, and so will seem as if it is imposing an objective theory of *X*-ness on the participants. The idea motivating *OPR*, then, was to present an individual-centered theory of reasoning, but to do that we must have some account of what is bona fide reasoning, and so we must appeal to the theoretical perspective. On this theory of reasoning, if the individual deliberated to a modest extent on *her own*

beliefs and values given her own norms of reasoning — and these are intelligible as reasoning (doodling doesn't count as reasoning) — she would conclude she had a sufficient reason to endorse a rule. Sufficient reason was thus substantively defined by a person's own system of beliefs, values and norms about good reasoning, but with some weak procedural requirement (she deliberated on her beliefs and values for a reasonable amount of time) so that we could isolate a wide range of *moral reasoning* as opposed to, say, announcing the first thing that came into your head, musing, egoistic planning, etc. We wish to theorize about a wide range of what might constitute moral reasons, but for that we need to classify what counts as moral reasoning. I do not think there is anything very objective about this.

2.4 Inherent evaluative utility

So — to return to our model — we have a finite, large system S of moral agents $\{A, B, C\ldots\}$ who possess evaluative or normative standards $\{A_\Sigma, B_\Sigma, C_\Sigma\ldots\}$, all of whom meet a standard of competent moral agents. Suppose further that each individual identifies a set of possible moral rules that he might act on. There is no need at this stage for a Kantian legislator to winnow down the option set (see further §4). We let each individual identify her own option set of possible moral rules over some areas of social life (say property rights). We can then generate a social option set $\{R_1 \ldots R_n\}$ as simply the union of all individual option sets. Departing a bit from the measurement assumptions of *OPR*, let us suppose that each person $\{A, B, C\ldots\}$ has a cardinal utility function, based solely on her normative standards $\{A_\Sigma, B_\Sigma, C_\Sigma\ldots\}$. For ease of presentation let us suppose that each person $\{A, B, C\ldots\}$ can score every rule in the social option set. [1] Let us call u_A person A's *inherent evaluative utility*, a cardinal ordering of $\{R_1 \ldots R_n\}$ in terms of their normative satisfactoriness given A_Σ. If Alf was the sole moral legislator for all, social morality would be based on this ranking alone. To fix ideas, assume each person's inherent utility assigns a score to each rule in $\{R_1 \ldots R_n\}$ from 0 to 10. A rule R_i scored 0 represents for Alf a rule that fails to adequately satisfy his evaluative standards; $u_A(R_i) = 10$ designates the rule(s) that, in Alf's view, perfectly conforms to them. In terms of the analysis of *OPR,* a rule that scores 0 is, from a person's evaluative perspective, not a *moral* improvement on having no rule at all on this matter.

Assume then we consider the evaluative utility that each person in S gives to each rule in the social option set and identify all rules in $\{R_1 \ldots R_n\}$ that everyone in S scores higher than 0. This *socially eligible set* is, then, a subset of $\{R_1 \ldots R_n\}$. Thus far the analysis looks very much like traditional moral philosophy, which thinks of moral reasoning in terms of each person judging "what I believe we all ought to do." Each thinks of what she considers normatively most sound (her inherent evaluative utility), and then identifies what rule all should follow. Such judgments are a reasonable way to begin thinking about how we should live together, but an unfortunate place to end. *OPR*'s fundamental claim is that moral agents — at least very, very many — do and should care about *sharing* a morality with others, proclaiming not simply "I've concluded that you and I must do such-and-such," but also "You and I have both concluded that you and I must do it."

[1] *OPR* (303 ff.) shows that such completeness is not required.

2.5 The push beyond inherent utility: the practice of accountability

The issue, then, is why, in addition to their inherent utility, agents are concerned with sharing moral rules. The answer I propose in *OPR* is the importance agents attach to participating in a practice of moral accountability. From the philosophical perspective — that from which we think about things such as a practice of accountability — we can identify the conditions for Alf and Betty to secure what we might call:

> MUTUAL ACCOUNTABILITY: Persons *A* and *B* justifiably hold each other accountable for violations of R_i only if each endorses R_i.

The status of MUTUAL ACCOUNTABILITY is an element of the philosophical perspective. It is argued that once we consider the best analysis of the nature of accountability (and responsibility), we shall discover that Alf rationally can hold Betty accountable — and experience the accompanying moral emotions of indignation or resentment at her failure to comply — only if he justifiably believes that she knew better. As a competent moral agent, if she had reasoned as well as the moral practice required, she would have seen the moral force of his demand. And that is because the moral demand is based not just on his reasons, but her reasons too: she also endorses the rule. If he admits that other competent moral agents (reasoning as well as the practice of morality requires) fail to see the moral force of R_i, Alf cannot hold them responsible for not conforming.

And without mutual accountability, *OPR* argues from the philosophical perspective, there simply cannot be an effective social morality. Social morality has a critical function to perform in human life, grounding widespread cooperation by invoking the internal normative guidance of participants. Here I heartily concur with Deirdre Nansen McCloskey's (2016) rejection of the narrow "economistic" view that human institutions — and the rules of social morality are types of institutions — are simply ways that society channels self-interested behavior in socially beneficial directions. Instrumental rationality cannot explain the way humans act *in* rule-based systems: we have evolved strong normative guidance systems such that when we accept a moral rule we have a strong behavioral tendency to comply (Kitcher 2011; Gaus 2015; Boehm 2012). But for this to be rational the normative guidance requires, as a necessary condition, that a person's evaluative utility for this rule is above 0; if it is 0, she cannot draw on any of her own evaluative commitments to see why she should act on it, nor by definition can any others show she has normative reasons to act on it. We thus can reformulate MUTUAL ACCOUNTABILITY as

> MUTUAL ACCOUNTABILITY*: Persons *A* and *B* justifiably hold each other accountable for violations of R_i only if they believe $[u_A(R_i) > 0]$ & $[u_B(R_i) > 0]$.

However, MUTUAL ACCOUNTABILITY* is too weak. As Munger observes, Alf's question is not whether Betty has *some* reason to endorse and act on R_i, but whether she has *sufficient* reason. Again, occupying the philosophical perspective, we reason thus: suppose Alf seeks to hold Betty responsible for failing to act on R_i and he points out that $u_B(R_i) > 0$. She replies: "Yes, but there is an alternative rule R_j in the social option set $\{R_1 \dots R_n\}$ that is inconsistent with R_i, such that $u_B(R_j) > u_B(R_i)$." She cannot accept responsibility for failing to act on R_i when R_j is, from her perspective, a rule that better

satisfies what she sees as the standards of morality. It would be absurd for her to act on R_i in this case, for she thinks it is a morally-inferior rule to R_j: something would be amiss if she held herself accountable for doing what she thinks she has better moral reason to do. This seems to pose an insuperable requirement: to hold Betty responsible Alf must not only hold that $u_B(R_i) > 0$, but that Betty does not rank an alternative as morally better. How can that occur when there is great disagreement about morality? Call this the *Puzzle of Accountability Under Conditions of Diversity.*

It may be questioned whether mutual accountability is really necessary. That a common system of rules is required does not imply that this must be a widely, normatively endorsed system. As McCloskey's *bête noir*, the institutionalist, might argue, a system of incentivized rules can channel self-interested behavior in the socially preferred direction. And as the Hobbesian has long stressed, avoiding punishment is a strong incentive — one of the things that we have learned from the work of Ernst Fehr and others is just how important punishment is to maintaining social cooperation (*OPR*: 103–22). Yet we also know that when punishment fails to correspond to what people believe are legitimate normative expectations, punishment easily evokes "anti-social" counter-punishment. As Bowles and Gintis (2011: 26) stress, effective punishment depends on legitimacy: unless those to be punished and their friends and allies are convinced that the rule being enforced is legitimate, a punishing action taken as a means to protect social cooperation can lead to weakening it. Experimental evidence (e.g., Hopfensitz and Reuben 2009) confirms that attempts at punishment readily evoke counter-punishment when the offender does not experience guilt — that is, given her own perspective, the offender does not see that she violated a moral rule. Accountability really is fundamental for an effective social morality.

3 MODELING AN INTEREST IN ACCOUNTABILITY: DIFFERENTIAL REASONS TO SHARE A RULE

3.1 Modeling individual perspectives on the value of sharing

This gets us to a critical juncture in the analysis. We have described moral agents and how they evaluate feasible moral rules, and we have seen that holding others accountable for violating R_i requires that given their own perspectives, they accept that they have moral reasons to accept it (as the best option in the socially eligible set). However, we have not modeled this second idea into our understanding of moral perspectives. Although from the philosophical perspective we can appreciate the importance of accountability, are we warranted in attributing recognition of this to moral agents from their individual perspectives?

Alf's commitment to a rule to which he can hold others accountable can be modeled in terms of the moral value he places on sharing rules with others. As we have seen (§2.5), if Alf and Betty are to hold each other accountable for violations of R_i they must share it: given their individual moral perspectives each must endorse R_i as the best in the social option set. Thus, we can say that Alf's commitment to a practice of mutual accountability is expressed by how important it is for him to share moral rules with others: the more accountability is important, the more stress he places on sharing rules.

To capture this concern with sharing let us expand each individual's concern to formulate a total moral utility function U, such that each person not only has an inherent moral utility for each rule R_i (expressing his "I conclude we ought" reasoning) but weights this utility by how many competent moral agents act on this rule. Alf's total moral utility function U_A for rule R_i will be his inherent utility $u_A(R_i)$ multiplied by a weighting function $w_A(nR_i)$, where w varies between 1 and 0, and nR_i is his estimate of the number of others who are acting on R_i. So Alf's total moral utility of rule R_i is defined as:

$$U_A(R_i) = u_A(R_i) \times w_A(nR_i) \qquad 1$$

Although from the philosophical perspective we can appreciate the importance of moral accountability (and, so rule sharing), we cannot assume that all competent moral agents place the same importance on it. From the philosophical perspective we are concerned with this question, but it would be highly controversial to impute such a concern to all individual perspectives. (We again see the importance of keeping the perspectives distinct.) Let us, then, allow wide variation in the weighting function among individuals. Alf may be someone who puts great value on accountability; the more people who act on a rule the greater weight he gives its inherent utility. Perhaps for Alf, the w of R_i equals 1 only when the n acting on R_i approaches all in S; perhaps unless some threshold n act on R_i, $w_A(nR_i) = 0$. (This, of course, implies that no matter how high the inherent moral utility of R_i it would have no overall moral utility if the number of S acting on it is below the threshold.) Betty, on the other hand, may place little importance on sharing a rule, so that as long as a very few act on R_i, she weights its inherent utility as 1. Most, we suppose, fall in between; assuming that accountability is important to most, the distribution of weights will be biased toward Alf-like ones and away from Betty's.

We can now ask our core question: under what conditions can S share a social morality, and so ground a practice of moral accountability? *OPR* (395ff) considered a very limited case, where agents disagree on inherent utility but all value sharing a rule with others — indeed an agent received moral utility from acting on a rule only when another acts on it as well. The analysis can be expanded to allow moral utility from unilateral action on a rule (e.g., Betty) in more heterogeneous populations, though of course depending on the distribution of inherent utilities and weightings, at times competent people will achieve a shared morality, and sometimes they will not.

3.2 The basic convergence dynamic

Consider a simple case in which there are only two options, R_1 and R_2. Only R_1 is eligible for everyone in S [i.e., $u(R_1) > 0$)] while another rule R_2, some subgroup g of S (where S-g is significantly greater than g) holds [$u(R_2) > u(R_1) > 0$]. If (*i*) interactions within S are uniform so that each interacts with each at the same rate, (*ii*) there is good knowledge of the actions of others, and (*iii*) the distribution of social weighting functions w among S is varied and does not have deep discontinuities, all total moral utility maximizers in S easily converge on R_1. Here a bandwagon effect is apt to occur because of the greater ability of R_1 to be shared.

To see this, assume we have iterated interaction over multiple periods. In the first period all act in the way we would expect of moral agents characteristic of orthodox

moral philosophy: they commence action on the basis of their "I conclude we ought" judgments — *u*, their inherent utility. They act on their inherent utility, as no one has yet formed expectations of what others are doing. So initially all in *S*-*g* act on R_1 and all in *g* act on R_2. In this very simple model we assume that after the first period everyone reconsiders whether they would have achieved higher total moral utility (*U*) if they had endorsed, and acted upon, the alternative rule. If so, they switch and play the alternative in the next period (we might instead employ more sophisticated Bayesian updating rules, but the basic dynamic would remain). Now examine α-types, members of the subgroup *g* (*i*) whose $u(R_2)$ is only slightly greater than $u(R_1)$, and (*ii*) whose *w* values for both rules is such that they place great value on sharing rules (their *w* is very low when most others do not share a rule). Subgroup α will switch to R_1 because most in *S* acted on it in the first period; so long as the higher weighting of R_1 over R_2 outweighs its lesser evaluative utility, R_1's total moral utility to α-types will be greater [i.e., $U_\alpha(R_1) > U_\alpha(R_2)$]. If so, α-types switch to R_1. Now consider β-types, members of *g* who either have a marginally higher evaluative utility gap between R_2 and R_1 than do αs, or who have marginally lower weighting of sharing than do α-types. Given the defection of αs to R_1 in period 2, β's will switch to R_1 in period 3; given β-types weightings, the defection of α-types to R_1 will make it the case that $U_\beta(R_1) > U_\beta(R_2)$. And so on to types who were increasingly prone at first to act on R_2, so long as the *g* population does not exhibit sharp discontinuities in weightings coupled with evaluative utility gaps. It is important to stress that α's and β's inherent utilities do not change: their weighting of the value of interactions with others tilt them to acting on R_1. Eventually we will come to ω-types: those for whom there is a large gap between $u(R_2)$ and $u(R_1)$, and whose *w* is such that they do not greatly care how many people act on a rule. At this point, however, there are so few people also acting on R_2 that even ω-types are very apt to change to R_1. (Note that in this case the opposite dynamic to R_2 cannot take root, as for all members of *S*-*g*, $u(R_2) = 0$.) To be sure, R_1 need not go to fixation: for some ω-types the gap between $u(R_2)$ and $u(R_1)$ may be so great, and their concern with sharing rules so low, that they continue to act, pretty much alone, on R_2.[2]

It is important that this dynamic does not require an entirely smooth distribution of social weightings or continuous gaps in *g* between $u(R_2)$ and $u(R_1)$. All that is required is that at each stage enough people recalibrate which rule is best given their total moral utility *U* such that, at the next stage, more of those who were still optimizing by R_2 adjust their actions to R_1, until the overwhelming majority does. What the dynamic certainly does depend on, though, is that most of the R_2 advocates in *g* significantly weight the importance of the number of others with whom they interact on shared rules. If a large proportion of *g* essentially only care about their inherent evaluations of R_2 they will not adjust their moral behavior.

3.3 Path-dependence in social morality

Consider another case, in which for the entire group there are only two rules to which all ascribe positive inherent moral utility. The core logic of the first case applies, assuming again that (*i*) interactions within *S* are uniform so that each interacts with everyone at the same rate,

[2] More generally, the basin of attraction of the all-R_1 equilibrium will be large, but not the entire state space.

(*ii*) there is good knowledge of the past actions of others, and (*iii*) the distribution of social weightings and evaluative utility gaps taken together is not deeply discontinuous. Here we have *S* divided into two subgroups, *g* and *g**, where for *g*, $u(R_2) > u(R_1) > 0$ and for *g** $u(R_1) > u(R_2) > 0$. The outcome will significantly depend on the relative sizes of the two groups, the distribution in each of the utility gaps between $u(R_2)$ and $u(R_1)$, and the distributions of the social weightings in each group. The more these are similar in the groups, the more the process will be unpredictable and path-dependent, in which small events (for example, whether *g* or *g** members tend to switch earlier, use different updating rules, or whether the general knowledge about who has switched is biased towards either rule) will determine the outcome (Arthur 1994: chap. 5). While either rule is a possible equilibrium, again there is a strong tendency for the overwhelming majority (essentially everyone except those who put very little weight on sharing a rule with others) to gravitate to one or the other rule if the *w* values are varied and not highly discontinuous. If the typical utility gap between the two rules is wide and the value that the typical person puts on coordination is low, then it is quite possible that *g* and *g** will equilibrate on different rules: here competent moral agents will fail to share a social morality, and so their moral views lead them to accept restricted moral accountability. Thus the "public justification" of a rule of morality cannot be assumed as an a priori requirement, as is the wont in the public reason literature. It is a social and moral achievement of the first order, arising when moral agents put moral value on sharing moral rules with their fellows and so grounding a practice of mutual accountability. Note, however, that in our second case two distinct patterns drive toward convergence: (*i*) when the typical utility gap is modest, as is the weighting of interacting with others, and (*ii*) when typical utility gap is large, but so are the social weightings.

This, then, further explains the result sketched in *OPR* (400), where either rule may be selected in a path dependent process, assuming an equal initial division of the population, as in Fig. 1.

3.4 Solving the Puzzle of Accountability through self-organization

These, of course, are very simple cases, assuming as they do uniform frequency of encounters; other assumptions more easily lead to groups dividing up into different networks of responsibility. But a different point is of interest here. In the two cases just sketched, *S* manages to solve the Puzzle of Accountability Under Conditions of Diversity: each ends up

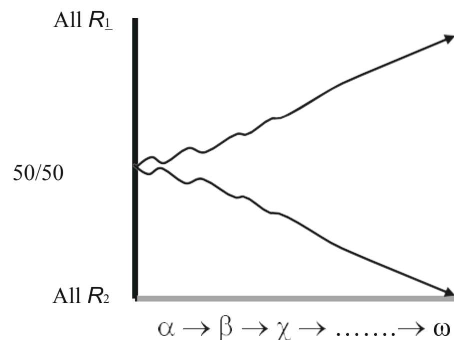

Fig. 1 Cascade dynamic by groups

acting on the rule that gives her the highest total moral utility (U). And thus they secure the strong conditions needed for a practice of mutual accountability. And they do this even if a significant number in S place rather low value on sharing moral rules with others. Thus we can begin to understand when a system of moral accountability arises, based solely on each following her own individual moral perspective, and making the best response to the moral choices of others. There is no central coordinator and the philosophical perspective does not dictate the result. But the philosophical perspective does allow us to grasp how a moral equilibrium on a shared rule of accountability arises, in a variety of circumstances, by the free moral choices of competent individuals. Unlike in social contract theory, an order of public reason is not the result of top-down planning, but is spontaneous and self-organizing. As Boettke and Candela (2016) nicely put it, we see the conditions under which diverse moral agents, "despite their differences with one another must somehow stumble upon rules of social intercourse that enable them to live better together than they would apart."

What I especially wish to stress is that in the analysis sketched above reaching an equilibrium is not secured through a constructivist procedure. The philosophical perspective does not construct an equilibrium; it seeks to show how an equilibrium on a rule can arise. To be sure, *OPR* identifies "devices of public reason" — types of rules that are more favorable to a widespread practice of accountability: basic rights of agency and jurisdictional rights such as private property. Without employing these "devices" it is unlikely that a diverse system of agents will be able to converge on any common rules. As I have put it more recently (Gaus 2016: chap. 4), the rules characteristic of the open society are accommodative to diversity as such; from the theoretical perspective we can see that systems organized around such rules are far more likely to arrive at a widespread equilibrium on social-moral rules. Perhaps Kirzner's "finders-keepers principle" (Boettke and Candela 2016) is one of these. In any event, it is critical to understand that the theoretical perspective does not fully justify these rules as part of a social morality: only the actual convergence of competent moral agents could accomplish that. Rather, the aim of that analysis to demonstrate that these types of rules possess positive evaluative utility for all — are in the socially eligible set (§2.4) — and in that sense fulfill a critical first step in the process of public justification.

Although it is indeed true that *OPR* focuses on the features of moral equilibrium, I cannot agree that in it "public justification of social society is analogous to a market viewed in terms of equilibrium *rather* [than] *a market viewed in terms of an entrepreneurial discovery process*" (Boettke and Candela 2016 emphasis added). Both equilibrium and discovery are critical. In social morality Austrians also must focus on equilibrium conditions. While markets are constantly changing, and at best only move toward or away from theoretical efficiency equilibria, this very dynamism of markets depends, as Hayek stresses, on a relatively settled moral framework that is not constantly in flux (Gaus 2016: 165–75). Dynamic markets are moved by dispersed, individual planning, but this planning requires some relatively fixed parameters, such as basic rights of property. Thus the study of equilibrium states is fundamental to understanding social morality in a way that is not, I believe, true of markets. Here there is a deep disanalogy between markets and morals. However, as I have stressed, finding these equilibrium states is indeed a process of discovery, deriving from the individual moral perspectives and their searching for ways to maximize the satisfaction of their moral standards; the theoretical perspective uncovers the conditions under which this is most likely to result in a widespread practice of mutual accountability.

4 THE ORDER OF JUSTIFICATION

Everything I have said thus far concerns equilibrium on specific rules. Munger points to a fascinating question: just as we can ask "could it be that all markets are simultaneously in equilibrium?" we might ask "could it be that all moral rules are simultaneously in equilibrium?" To translate the problem from economics, suppose that system S has two types of rules, *R-type* and *T-type* rules; with two resulting socially eligible sets, $\{R_1, R_2\}$, $\{T_1, T_2\}$. All four rules, then, are ranked by everyone as having positive inherent utility. Suppose that S has, in fact, equilibrated on R_1: the process of discovery led to a cascade on it, so that everyone endorsed R_1 because it was their total moral utility-maximizing choice. So we have a single-rule equilibrium. But consider some critical subgroup g in S. When they evaluate the inherent utility of T_1 and T_2, it matters to them whether S has already equilibrated on R_1 or R_2. If S equilibrates on R_2, members of g would strongly favor T_2, but if S equilibrates on R_1, they strongly favor T_1. Thus, given the first equilibration on R_1, members of g strongly favor T_1. Unfortunately, while most people in S incline to T_2 because members of g now strongly incline to T_1, they stop any cascade to T_2; but T_1 is widely disfavored and so it could be that S also fails to equilibrate on it. But then S fails to equilibrate on any *T-type* rule, because it first equilibrated on R_1. But if it had first equilibrated on R_2, then it could have achieved a T_2 equilibrium.[3]

I think Munger is correct that one can plausibly read *OPR* in a way that invokes a sort of Kantian Conductor in the form of what I called the "order of justification." The idea is that we take the most basic issues and settle them; take them as fixed, and move on to less basic ones against that background. This would not ensure equilibrium for all rules, but it would at least, as it were, identify the preferred set of partial equilibria for S: equilibrium on basic matters, with perhaps acceptable disequilibrium on less basic questions, or equilibrium on those issues within the parameters set by the most basic, settled ones. So the Kantian Conductor might announce: "Let us first look at T-rules, and find the equilibrium. Then we will consider R-rules, and so on." If we are to achieve the best partial equilibrium — the best "single order of public reason" (Munger 2016), it looks like the Kantian Conductor is necessary in a way that does hint, I think, at the "Walrasian Auctioneer."[4]

On *OPR*'s analysis, such a Kantian Conductor has no place. The theoretical perspective seeks to understand the conditions under which shared social normativity and a practice of accountability can arise, but it is unknown ahead of time what these rules will be. There is no Kantian Conductor, and we do not understand the order of public reason any better by supposing one.

[3] If for each member of S the moral evaluation of all rules is closely tied to the value of other rules in S, then S becomes what we might call a justificatorily complex system (Gaus, 2016; chap. 2); rather than productive moral change we will get endless, wandering, movement in the option space. I do not believe that moral evaluation is that drastically "holistic" — for one thing, competent moral agents simply cannot keep track of such complex evaluations. We tend to partition our evaluations (Gaus 1996: 107–8; *OPR*: 272–5). Thus I do not think questions of overall moral equilibrium are really pressing.

[4] "Hayek believed," writes Vallier (2016), "that the consistent application of the test of Kantian assent will 'amount to a test of compatibility [of a rule] with the whole system of accepted rules.' In other words, while the Kantian contractarian test cannot assess the system of moral rules as a whole, it can bring them into alignment and organize them into a hierarchy if used iteratively by testing each rule in sequence."

The order of justification, then, is not properly understood as imposed by a Kantian Conductor. Rather, the way it entered into *OPR* was as a claim about a feature of any viable system *S*: "All moral orders suppose an order of justification: some things are more or less settled, and that settlement provides a background for further justification" (*OPR*: 275). Thus the idea was that an actual system of interacting individuals will come to some shared understanding of what is generally most basic to their shared social lives and what is not. I still think this claim is plausible. To share a liberal system of morality is to see that, say, avoidance of serious harm is more basic than avoiding insult, while to share an honor code is to deny that. One of the things detailed by McCloseky's study of the rise of bourgeois morality is the rise of a certain order of justification, where some matters are seen as fundamental to moral life, and must be assumed as generally fixed when thinking about others. In this way, without a Kantian Conductor, evolving social moralities work to solve some of the problems that I have discussed in this section.

But of course any actual agreement on such an order of justification is highly imperfect. Once we abjure appeal to a Kantian Conductor we are left with the realization that equilibrium is always imperfect and in flux. Having arrived at an approximate equilibrium on one sort of rule (say, property rules), some moral innovators realize that this equilibrium is inconsistent with their strongly favored agency-related rules, on which society is not presently well coordinated. Thus, for example, a society may find that it has equilibrated on a conception of property (e.g., husbands assume their wives' property at marriage) that is preventing coordinating on agency rights for women. As various innovators stress the importance of women's agency rights, they will evaluate the property equilibrium and perhaps begin to defect from it. Thus the entire system is never in equilibrium. That, though, is a good thing, for while stability is good for planning, too much stability can lead to stasis (Gaus 2016: 223ff). A critical source of moral change is precisely this lack of overall equilibrium, allowing one equilibrium to be challenged as another is established. It is here that we clearly dispense with static equilibrium models, understanding social morality as a series of punctuated equilibria.

5 SOCIAL MORALITY, RULES AND VIRTUE

"Gaus's rejection of neo-Aristotelian virtue ethics as a solution to modernity's problems," Boettke and Candela (2016), write, "while plausibly true in terms of politics, doesn't necessarily mean that neo-Aristotelian virtue ethics doesn't provide a justificatory mechanism, namely self-direction, that informs us of the feasibility of reaching a political solution to generate the demands of social morality." This leads to a fundamental difference between *OPR*'s analysis of social morality and virtue-centered accounts of McCloskey (2007); McCloskey (2016), Rasmussen and Den Uyl (1991) and many others who are committed to an ethics-sensitive analysis of market orders. I have addressed some of these matters in a recent review essay of McCloskey's wonderful *Bourgeois Equality* (Gaus forthcoming). I certainly cannot rehearse even that sketch here. Let me state, in an unfortunately didactic manner, why I do not think the current renaissance in virtue thinking fundamentally helps in explaining extended moral orders, though virtue, like a number of other considerations, certainly enters into

inherent utility functions. In this regard I am reaffirming the Hayekian account of moral rules over the new liberal virtue ethics.

There is, firstly, a great deal of compelling evidence that character traits and attitudes are poor predictors of action (Bicchieri 2006, 2016; Nisbett and Ross 1980; Harman 1999, 2000). This seems counter-intuitive, especially to those committed to straightforward individualist explanations. If Alf acts morally, or well, it must be because of his character or attitudes. I believe that a fair reading of the evidence shows that while a person's personal normative attitudes and character are important in rendering her sensitive to an informal social rule (Bicchieri 2016), an often more important variable is our empirical expectations — i.e., what we expect others will do (Bicchieri and Xiao 2008) — and our beliefs about their normative expectations — what we believe they expect us to do. People care deeply about what they think others expect of them — and this is usually a better predictor of norm-based action than their own moral attitudes (Bicchieri and Chavez 2010). What I did not sufficiently stress in *OPR* is that the importance we place on meeting the normative expectations of others is a critical reason why societies converge on moral rules and maintain them. If I hold a rule to be reasonably consistent with my moral commitments and believe that the rest of *S* normatively expect me to comply, this is usually critical in moving me to endorse it and act on it.

We are much more socially sensitive to what others expect from us, and much more guided by rule-thinking, than most economists appreciate, because their own field is based on a highly individualist, resolutely internalist motivational assumption. McCloskey and her followers reject the narrow maximizing view of such behavior on which much economics is based, but they still tend to view behavior as generated by one's own convictions. Morality is thus understood as a sort of strongly self-directed behavior. Hayek never made this error, and that is why his brand of individualism is so complex and sophisticated. As *OPR* stresses, we are rule-followers, and we are highly responsive to what others expect of us. I think there is very strong evidence that our capacity for rule-based guidance has evolved along with our capacity for social life (Kitcher 2011; Gaus 2015). We are social creatures *because* we are rule- followers. I am delighted to "channel Hayek" (Vallier 2016) on this point: "Our reason did not produce the social order…. Rather, the requirements of social order shaped our reason." All this means that our moral action is often surprisingly distant from our attitudes, character and view of virtues. I realize that people reject this as implausible, but as social scientists we know that our folk and introspective understanding of social phenomena is often unreliable.

None of this is to say that one's moral virtues and attitudes do not enter into one's moral action. They enter into one's understanding of whether a rule is endorsable, and very likely that does affect how sensitive to the rule's requirements — apt to act on it — one is. Rules that don't align with one's attitudes can, it would seem, be more easily undermined. And I agree with Boettke and Candela that when searching for a new moral equilibrium, internally-guided moral innovators are indeed critical in determining where society ends up. But virtues are one of a number of factors, and social rules can lead one to act on, and even enforce, rules that are weakly endorsed — or even run against — these commitments.

6 VALLIER AND THE PHILOSOPHICAL PERSPECTIVE

As I understand them, Vallier's concerns are resolutely from what I have been calling the philosophical perspective, but he brings out the important point that anyone may take up this perspective, stepping back — perhaps at moments of crisis and doubt — and think about the system as a whole. Here, rather than seeing things as a participant in an ongoing way of moral life, one asks whether one can endorse, from this detached viewpoint, the very idea of living in a system of this sort. The question here is not whether the system secures moral accountability among free and equal persons, the focus on *OPR*'s philosophical perspective. Rather the worry here is whether, despite the conclusion that the system secures this, we should participate in it. Should we, as it were, reject the whole enterprise of social morality?

Often when we assume the philosophical perspective we are driven to higher-level, potentially skeptical, questions that press the query: So what? So what if my normative beliefs drive me into participating in this system? So what if this is what I, as a real agent, care about? As I suggest in *OPR* (a thought that utterly failed to move Vallier), in this context questions for global justification often reduce to "So what if what I reflectively care about leads me to endorse this — why should I care about what I reflectively care about?" I can only reiterate: "Because *you do* reflectively care about it." Of course we can try to go further: "You have *reason* to care about what you care about." But now that we are playing this skeptical game, we invite the reply, "So what if I have reason to care about what I care about? Why should I care about reason?" I am not being facetious; in an important work on rationality Hampton (1998) sought to answer this question against the "curmudgeon" who didn't care about reason. Philosophers who take this tack almost always seek to show that our answer ends with a claim to objective truth ("The objective authority of reason demands that you acknowledge its claims!"), as if that reply really resolved any issue, except in the mind of the person who gave it.

So my first, and I think main, response is that often when we ask deep philosophical questions we confuse ourselves — that is why so much philosophy leads to a theoretical skepticism which threatens practical skepticism. To resist this, most contemporary philosophical ethics and political philosophy has retreated into what I have elsewhere (2016: chap. 5) called the "citadel" of inner, apparently certain, "intuitions." As should be clear by now, a main aim of *OPR* is to focus on the normative standards of individuals and how these lead moral agents to form themselves into orders of public reason. Moral philosophy has, in my view, been far too enamored with reasoning from the detached perspective, as if the philosopher was not herself part of a network of social relations. The result has been an abstract, overly-individualistic, understanding of social morality.

But, as I have stressed, *OPR* does not do away with the philosophical perspective; it rather seeks to curtail its ambitions. Reflective people sometimes do take a detached point of view, and question their involvement in practices, and one task of the philosophical perspective is to address some of these worries.

In answering these questions and worries we must keep firmly in mind that we are thinking about the practice of social morality from an external point of view and not legislating within it. This philosophical perspective is not part of the normative system of social morality — unlike the Smithian impartial spectators, in *OPR* this detached

perspective is not the ultimate source of adjudication within social morality, but the locus of a variety of reflective questions that we often seek to answer about social morality. Now as I read his comments, Vallier (2016) may not see things this way.

> In Chapter II of *OPR*, Gaus argues that while the "instrumentalists" are right that social morality is necessary for human cooperation and social life, the authority and justification of social morality cannot depend on an appeal to those benefits alone. This is why Gaus claims that the Kantians are right to "insist that our moral relations can never be reduced to their instrumental benefit." If that is so, how can members of the public assure themselves that they have not landed in a confused practice by reflecting on the benefits of the social-moral order as a whole? After all, Gaus wants us to comply with moral rules for non-instrumental reasons. Recognizing the instrumental benefit of our moral rules cannot morally require us to comply with them. Similarly, recognizing the instrumental benefit of the system of moral rules cannot morally require us to comply with or trust it.

In this passage Vallier supposes that when employing the philosophical perspective we are seeking to show that we have moral reasons to comply with the rules of social morality: the philosophical perspective is to explain the "authority and justification of social morality." This, though, is to embrace precisely the role of the philosophical perspective that *OPR* rejects: that, at the end of the day, it is the detached perspective that provides us with the ultimate justification of the authority of social morality.

The response of *OPR* to the "worry that we are stuck in a 'confused' moral practice" is not to provide a deeper moral justification of the practice but to point out that, given the very analysis of social morality that the theory has advanced, such social morality is the bedrock of human life. This does not show that it has moral authority, for, as I have stressed, the critical aim of the theory was to ground the source of normative authority in the participating agents. And it certainly does not show that "our moral practice is justified on purely instrumental grounds" (Vallier 2016). At the end of the day to retreat to the philosophical perspective and from it advance the ultimate account of the normative nature of morality or its justification would unravel the core thesis. But as any overall theory of a human phenomenon can do, once we take a systemic view of it we may see that it is not a "confused" practice that we stumbled into. The questions of whether it is a confused, or stupid, practice can be most effectively answered by showing its broad and fundamental instrumental value to human social life. There is no inconsistency here with the claim that, from the individual participant perspective, morality cannot not simply be seen as instrumental to securing our ends. As always, we must try to keep the two perspectives distinct.

Take an analogous case, the practice of baseball. Suppose we appeal to *OPR*'s philosophical perspective to analyze the practice of our local baseball league, and find that its rules have authority for all participants. They all endorse the rules of baseball, see them as authoritative, including the calls of umpires. But a participant asks, "Is this is a stupid game? Would I be better off switching to a better game — maybe basketball?" (Recall Michael Jordan's attempt to switch the other way.) Our worried baseball player may then seek to step back from baseball and look at it from a detached

perspective. And, as usual, she finds that Rawls has come to the rescue, providing a proof that baseball is the best game, alone satisfying six conditions for excellence.[5]

Like the philosophical perspective advanced by *OPR*, Rawls's proof assuages the worry of the baseball player by providing new, systemic, considerations of which she may have been unaware. That is what a good theory does — it provides a new understanding of the practice, one that was unavailable to participants about which it theorizes. But while Rawls's proof may assuage the worries of our player, it would not add authority to the umpire's calls, or to the infield fly rule. Their normativity has already been ensured by endorsing the practice of baseball as a player. In *OPR* the philosophical perspective assures players that social morality is a game that humans receive great benefits from playing. That may help assuage the concerns of those who worry that they should switch to a different moral game.

But it may not. *OPR*'s theory of social morality, with its modest understanding of the power of the philosophical perspective, can only do so much. Some may need more, and social morality allows them to look for it. Perhaps our baseball player needs be assured that Jesus is on the side of baseball (something that strikes me as plausible).[6] Not even Rawls's theory of baseball can assure her of that, but her religion can.

However, what cannot be admitted into an account of a free society is a civil religion that is a part of social morality — even, I must admit, baseball[7] — unless it is integrated into individual evaluative standards, and then it has no more authority than any other individual standards. My doctoral supervisor, John W. Chapman, was both a classical liberal and a Rousseau scholar; unlike many he insisted that Rousseau described the conditions for a society of free and equal citizens. However, the advocacy of civil religion is where, as he used to tell us, Rousseau lost his nerve that a free society could maintain itself. "The purpose of the civil religion may be to preserve man's political freedom, but it is a means which destroys his moral freedom and dignity" (Chapman 1956: 86). A civil religion must end up imposing moral standards on some who do not endorse them. An open society, I have argued, draws on an extraordinary range of individual moral perspectives that converge upon, and endorse, a social morality of equality and freedom, which is friendly to the exploration of new ideas and ways of living (Gaus 2016: chap. 4). From the philosophical perspective, a theory of the open society can show how its basic moral framework allows the maximum number of individuals to live in ways they find morally acceptable while searching for better ways of living, and to this extent may increase their confidence that the open society is a great

[5] "The Best of All Games," www.bostonreview.net/rawls-the-best-of-all-games

[6] Or perhaps, like Annie, she joins the Church of Baseball: "I believe in the church of baseball. I've tried all the major religions and most of the minor ones. I've worshipped Buddha, Allah, Brahma, Vishnu, Shiva, trees, mushrooms, and Isadora Duncan. I know things. For instance, there's 108 beads in a Catholic rosary and there's 108 stitches in a baseball. When I learned that, I gave Jesus a chance. But it just didn't work out between us. The Lord laid too much guilt on me. I prefer metaphysics to theology. … I've tried them all, I really have. And, the only church that feeds the soul, day in, day out, is the church of baseball." *Bull Durham*, written by Ron Shelton.

[7] Which has a plausible claim to be the American civil religion. Again, Annie in *Bull Durham:* "Walt Whitman once said [this turns out to be a paraphrase], 'I see great things in baseball. It's our game, the American game. It will repair our losses and be a blessing to us.'" Or, alternatively, in the words of a fictional J. D. Salinger, "I don't have to tell you that the one constant through all the years has been baseball. America has been erased like a blackboard, only to be rebuilt and then erased again. But baseball has marked time while America has rolled by like a procession of steamrollers" (Kinsella 2002: 252).

achievement. But more than this cannot be said by a theory that takes moral diversity seriously. Each must find his own deepest answers to his gnawing worries, but these are not part of the social morality of a free and diverse society.

7 ON THE THEORY AND PRACTICE OF SOCIAL MORALITY

I am deeply grateful to the authors of these three essays for spurring me to think harder about one of the most perplexing problems in moral theorizing: the relation between a philosophical theory of morality and the normative convictions of moral agents as they work out their moral lives. As the reader has seen, it is anything but easy to keep these perspectives distinct. But unless we do our best, we will fall into the fundamental error of most contemporary moral and political philosophy, of being blind to the individual perspective, falling prey to the theorist's vice of supplanting the way real humans organize their moral relations with her philosophical construction.

Perhaps more than anything else, it was this conception of moral theorizing that *OPR* rejects as asocial and dangerous to a free society. Hayek's social and rule-based analysis of morality was its great inspiration. Even many who are friendly to Austrian economics fail to appreciate the deep insights of the Hayekian account of a free society as a *moral* rule-based order — a spontaneous order that rises through intentional moral choices. The primacy of the individual perspective is the fundamental commitment of a Hayekian-inspired account of social morality. Yet, *OPR* agrees with Rawls (1958 [1999]): 53) that individuals sensibly advance justified complaints against the rules of social morality, and so a fundamental task of a theory of social morality is to investigate when individuals have such complaints — when they are being subjected to a morality that they would not reflectively endorse. Thus the philosophical perspective is always necessary to give us a critical leverage on our social morality, but it should never presume to dictate what our shared morality is.

Acknowledgments: My thanks to Kelly Gaus for her comments and suggestions. My thanks also to Kevin Vallier for organizing this symposium and, of course, to the authors of the other papers in it.

References

Arthur, W. B. (1994). *Increasing Returns and Path Dependence in the Economy*. Ann Arbor, MI: University of Michigan Press.

Bicchieri, C. (2006). *The Grammar of Society*. Cambridge: Cambridge University Press.

Bicchieri, C. (2016). *Norms in the Wild*. New York: Oxford University Press.

Bicchieri, C., & Chavez, A. (2010). Behaving as Expected: Public Information and Fairness Norms. *Journal of Behavioral Decision Making, Vol., 23*, 161–178.

Bicchieri, C., & Xiao, E. (2008). Do the Right Thing: But Only if Others Do So. *Journal of Behavioral Decision Making, Vol., 22*, 191–208.

Boehm, C. (2012). *Moral Origins: The Evolution of Virtue, Altruism and Shame*. New York: Basic Books.

Boettke, Petery J. and Rosolino A. Candela. (2016). "A Social Morality for Mortals: A Review Essay of *The Order of Public Reason: A Theory of Freedom and Morality in a Diverse and Bounded World*." *Review of Austrian Economics*, this issue.

Bowles, S., & Gintis, H. (2011). *A Cooperative Species*. Princeton: Princeton University Press.

Chapman, J. W. (1956). *Rousseau — Totalitarian or Liberal?* New York: Columbia University Press.

Gaus, G. (1996). *Justificatory Liberalism*. New York: Oxford University Press.
Gaus, G. (2011). *The Order of Public Reason*. Cambridge: Cambridge University Press.
Gaus, G. (2015). The Egalitarian Species. *Social Philosophy and Policy, 31*(Spring), 1–27.
Gaus, G. (2016). *The Tyranny of the Ideal: Justice in a Diverse Society*. Princeton: Princeton University Press.
Gaus, G. (forthcoming). "The Open Society as a Rule-Based Order." *Erasmus Journal for Philosophy and Economics*.
Gaus, G., & Thrasher, J. (2016). Rational Choice in the Original Position: The (Many) Models of Rawls and Harsanyi. In *The Original Position,* edited by Timothy Hinton (pp. 39–58). Cambridge: Cambridge University Press.
Gauthier, D. (1986). *Morals by Agreement*. Oxford: Clarendon Press.
Hampton, J. E. (1998). *The Authority of Reason*. Cambridge: Cambridge University Press.
Harman, G. (1999). Moral Philosophy Meets Social Psychology: Virtue Ethics and the Fundamental Attribution Error. *Proceedings of the Aristotelian Society, New Series, vol., 99*, 315–331.
Harman, G. (2000). The Nonexistence of Character Traits. *Proceedings of the Aristotelian Society, New Series, vol., 100*, 223–226.
Hayek, F. A. (1988). *The Fatal Conceit: The Errors of Socialism*, edited by W. W. Bartley. Chicago, IL: University of Chicago Press.
Hopfensitz, A., & Reuben, E. (2009). The Importance of Emotions for the Effectiveness of Social Punishment. *The Economic Journal, 119*(October), 1534–1559.
Kinsella, W. P. (2002). *Shoeless Joe*. New York: Rosetta Books.
Kitcher, P. (2011). *The Ethical Project*. Cambridge, MA: Harvard University Press.
McCloskey, D. N. (2007). *The Bourgeois Virtues: Ethics for an Age of Commerce*. Chicago, IL: University of Chicago Press.
McCloskey, D. N. (2016). *Bourgeois Equality: How Ideas, Not Capital or Institutions, Enriched the World*. Chicago: University of Chicago Press.
Munger, M. C. (2016). "Human Agency and Convergence: Gaus's Kantian Parliamentarian." *Review of Austrian Economics*, this issue.
Nisbett, R., & Ross, L. (1980). *Human Inference: Strategies and Shortcomings of Social Judgments*. Englewood Cliffs, NJ: Prentice-Hall.
Rasmussen, D., & Den Uyl, D. (1991). *Liberty and Nature: an Aristotelian Defense of Liberal Order*. Chicago, IL: Open Court.
Rawls, John. (1958 [1999]). "Justice as Fairness." In *John Rawls: Collected Papers*, edited by Samuel Freeman. Cambridge, MA: Belknap Press of Harvard University Press: 47–72.
Vallier, Kevin. (2016). "Gaus, Hayek and the Place of Civil Religion in a Free Society." *Review of Austrian Economics*, this issue.