



The Social Philosophy of Gerald Gaus: Moral Relations Amid Control, Contestation, and Complexity

ABSTRACT: *Gerald Gaus was one of the leading liberal theorists of the early twenty-first century. He defended liberal order based on its unique capacity to handle deep disagreement and pressed liberals toward a principled openness to pluralism and diversity. Yet, almost everything written about Gaus's work is evaluative: determining whether his arguments succeed or fail. This essay breaks from the pack by outlining underlying themes in his work. I argue that Gaus explored how to sustain moral relations between persons in light of the institutional threats of social control, evaluative pluralism, and institutional complexity, and the psychological threat of acting solely from what I shall call the mere first-personal point of view. The idea of public justification is the key to sustaining moral relations in the face of such challenges. When a society's moral and political rules are justified to each person, moral relations survive the threats they face.*

KEYWORDS: Gerald Gaus, liberalism, public reason, public justification, diversity

Gerald Gaus (1952–2020) was a social philosopher with the unusual trait of discouraging others from writing about his work. Nearly everything written about Gaus is evaluative: determining whether his arguments succeed or fail. This essay breaks from the pack by identifying a unifying theme in Gaus's work.

I argue that Gaus explored how to sustain *moral relations* between persons in light of the institutional threats of social control, evaluative pluralism, and institutional complexity, and the psychological threat of acting solely from what I shall call the *mere first-personal point of view*. The idea of public justification is the key to sustaining moral relations under these conditions. When a society's moral and political principles are justified to each person, moral relations survive the threats they face.

To expand on the above, the modern world challenges our ability to maintain valuable relationships with each other. Large and powerful states threaten us with coercion, harm, and death. Members of free societies seem to disagree ever more frequently. And our institutions have become so complex that we barely understand them. How can we sustain our relationships with others when we live

I am deeply grateful to the Gaus family for providing drafts of Gaus's forthcoming articles. I also thank everyone who gave me feedback on this article in writing or in conversation. These include two anonymous referees for the journal, Jacob Barrett, Fred D'Agostino, Adam Gjesdal, Keith Hankins, Brian Kogelmann, Liz Levinson, Alex Motchoulski, Ryan Muldoon, Alex Schaefer, John Thrasher, Chad Van Schoelandt, Michael Vincent, and Paul Weithman, among others.



under threat, have fewer common values, and are governed in ways that are inscrutable to us? Call these the *institutional* threats to moral relations.

What is more, human beings have a built-in temptation to reason only from their own point of view and to impose their values and commitments on others. When we give in to this temptation, we suffuse our lives with stunted relationships (Gaus 1990: 362), brow-beating, coercion, and moral dogmatism (Gaus 1996: 123–29, 165, 182–84), authoritarianism (Gaus 2011: xvi, 16, 30), ignorance and tyranny (Gaus 2016: 96, 139–44, 169, 196, 250), and failed reconciliation and self-governance (Gaus 2021a: 56–60, 74, 101, 278–83). Call these the *psychological* threats to moral relations.

Societies only resist these threats by installing moral and legal rules that all can accept. They govern themselves with rules that can achieve *public justification*. Consider the institutional threats: If states observe mutually acceptable rules, they must protect a wide range of liberties for all. In doing so, citizens can form and sustain moral relations with others without having to worry that their governments will manipulate or harm them. If the rules of social life are mutually acceptable to multiple perspectives, then pluralism need not undermine social cooperation. And when all can accept social rules, they will not resist governance based on them. Officials can then manage complex institutions.

Consider the psychological threats: if we constrain our actions by following rules that all can accept, we both express and reach moral maturity. Our actions take the perspective of others into account. Directing one another to follow rules they accept avoids browbeating and authoritarianism. And because each group can accept the rules, each can engage in free moral inquiry about how to govern its part of the social world.

I will unpack these ideas in the following nine sections. I explain the nature of moral relations (section 1) and their value (section 2). I then identify the institutional (section 3) and psychological (section 4) threats to moral relations. The solution to neutralizing these threats lies in a principle of evaluation implicit in our moral relations and moral emotions (section 5), specifically a principle of public justification (section 6). After reviewing the evolution of Gaussian public justification, I will explain how it helps to neutralize institutional and psychological threats to moral relations (section 7). I conclude by exploring the radical implications of Gaus's social philosophy for philosophical methodology (section 8) and then summarize the spirit of Gaus's project (section 9).

1. The Nature of Moral Relations

The term 'moral relations' and the related ideas 'moralized relations' and 'relations of accountability' figure centrally in Gaus's five major treatises: *Value and Justification* (Gaus 1990: 281, 287–315, 366), *Justificatory Liberalism* (Gaus 1996: 183, 214, 288–91), *The Order of Public Reason* (Gaus 2011: 8, 13, 174, 183–84, 193, 199–200, 223, 282–83, 426, 431, 463, 475), *The Tyranny of the Ideal* (Gaus 2016: 180–83, 211–12, 217–20, 230, 233–34), and *The Open Society and Its Complexities* (Gaus 2021a: 74, 99, 108, 190), and also in many articles.

Despite typically defining terms and ideas in detail, Gaus never carefully defined moral relations. We must reconstruct a definition from his work. Moral relations are human relationships enabled and structured by the reactive attitudes, especially guilt, resentment, and indignation (Gaus 1990: 260–305; 2011: 53–100; Strawson 1974, 1–49).

To define moral relations, one must understand Gaus's conception of moral agency (Gaus 1990: 292–300; 2011: 146, 192, 208–9n57). Gaus characterized moral agency by contrasting it with Hobbesian instrumental reasoners (Gaus 1986) and psychopathy (Gaus 2014). Hobbesian political philosophy claims that we can establish social cooperation through instrumental rationality alone. Reflecting solely on our goals and the best means to reach them, we will agree on the terms of social life. Gaus dissented. One can find Hobbesian moments in his *Value and Justification* (1990: 439–75). Yet, Gaus here still claims that we cannot explain social cooperation if humans pursue only their own ends (362–67). In his later work, Gaus (2011: 53–100) abandons Hobbes with a fifty-page argument against any 'instrumentalist' approach to the social contract. Fortunately, evolution equipped us with the capacity to follow moral rules that frustrate our ends—nature made us deontologists, albeit ones who can sometimes rank goals above their preference to follow rules (Gaus 2011: 513). Thankfully, we are as much rule-following creatures as goal-seeking ones (Gaus 2011: 101; Hayek 1973: 11). We reciprocate (Gaus 2008; 2011: 158–60, 193–204). Our moral emotions, specifically the negative reactive attitudes, drive us to direct others to follow moral rules, create resentment and indignation in response to violations, and push us to punish violators (Gaus 1996: 123–29; 2021a: 46–56).

Gaus thought that social cooperation requires moral agency. Because human beings can choose to follow rules that frustrate their goals, they can escape the relentless drive to maximize personal value, undermining the temptation to defect from cooperative arrangements (2011: 101–81). Once we establish cooperation, we will continue so long as others reciprocate (Gaus 2021a: 46–56). The reactive attitudes can drive us to abide by social rules when we internalize them. These rules structure our moral emotions so that we follow rules and punish violators. These emotions override our temptation to act noncooperatively (Gaus 2011: 232–58). Our moral emotions successfully override temptation because they implicitly evaluate moral responsibility for rule violations. We then want to cooperate with rule-followers and defect from culpable rule-violators.

Gaus (1996: vii) also contrasted moral agency with psychopathy. While psychopaths can follow rules and cooperate with others, they lack functional moral emotions. We therefore cannot relate to them morally (Gaus 1990: 287–92; 2011: 192n12). Psychopathic cognition sets the outer bound on who can stand in moral relations.

I see no further outbounds. Gaus sometimes indicates that moral relations include trust, friendship, and love. Yet, he never claims that these exhaust the class of moral relations (1990: 287; 2011: 192, 315). Indeed, we can have moral relations with strangers whom we barely trust. Moral relations are not always *good* ones. They exist wherever mutual accountability arises. And yet, moral relations contain the

basis for ensuring the moral relations improve. Good moral relations can return when immoral behavior ends.

Based on the foregoing, I define moral relations as follows:

Moral Relations: A and B stand in moral relations with one another if and only if they respectably believe that the other is an apt target of the reactive attitudes.

What has respectable belief got to do with anything? Gaus (1996: 63–84; Gaus and Thrasher 2021: 6–34; Gaus 2011: 63–69) thought that we should understand rational choice in terms of beliefs and inferences. We cannot stand in moral relations with others unless we can recognize that we are in moral relations, and thus moral relations partly consist in our beliefs about others.

Believing that our reactive attitudes aptly target others requires good evidence though not all relevant evidence. Beliefs about the accountability of others can be justified solely from evidence we uncover through a ‘respectable amount of reasoning’. Once we expend respectable cognitive effort, others cannot blame us for failing to grasp all the moral reasons that apply to our actions (Gaus 1996: 138–44; 2011: 249–51).

2. The Value of Moral Relations

Moral relations have value because we rationally value them, both in themselves and for further reasons. In *Value and Justification*, Gaus (1990: 87–92, 429–31) argues that value is subjective: value is explained by valuing; if this is so, then moral relations have value because we value them. Yet, Gaus also argues that we do not value moral relations arbitrarily. This argument pervades his future work.

We should value moral relations for two reasons. First, we should value them for cooperation. Moral relations enable the formation of large-scale cooperative social orders with all their benefits. Social orders with mere instrumental reasoners will break down (Gaus 2011: 53–101; 2015). One need not delve deeply into Gaus’s theory of practical rationality to see why: pure instrumental reasoners will often pursue short-term gains by defecting from cooperative relations because defection has greater value in the moment. Our interactions become purely strategic: we defect from cooperative arrangements when we can get away with it. If Gaus is right, moral relations in large-scale societies yield gigantic benefits. We undermine those benefits by acting purely instrumentally.

Fortunately, evolution equipped us with the moral emotions and the reactive attitudes in particular. This is likely due to group selection, where clans of primates with the reactive attitudes outcompete clans that reason only instrumentally (if the latter sort of clan ever existed). In *The Open Society and Its Complexities*, Gaus argues that *Homo sapiens* outpaced chimpanzees and bonobos precisely because the former overlaid instrumental reasoning with moral reasoning. We acquired strong motives to avoid and punish moral violations even at a cost to ourselves, which Gaus (2021a: 45–46, 63) calls ‘human

ultra-sociality'. The group selection process created the 'Modern Egalitarian Moral Package' (Gaus 2021a: 63).

If we evolved, therefore, to be natural-born deontologists, we can set aside our personal valuations to comply with deontic moral requirements, specifically publicly recognized moral rules. Human beings often prefer rule-following to goal pursuit. When we act on this preference, we capture the value of moral relations without endangering them, and this preference is not arbitrary; we evolved as reciprocators (Gaus 2021a: 43). Since we can follow moral rules because we find them fair or just, we will not violate them to maximize personal value.

We rationally value moral relations for instrumental reasons—they help us enjoy the benefits of large-scale cooperation. Our moral psychology allows us to do so by undermining our temptation to damage the social order.

The second reason to value moral relations is that they include thicker moral relationships of love and friendship. As Gaus (2011: 192) notes, 'A practice such as social morality is deeply embedded in our view of the world; it affects our understanding of interpersonal relations, including love and friendship, and so of what sort of life is worth living'. Gaus develops accounts of mature love and moral friendship in *Value and Justification*. Twenty years later, in *The Order of Public Reason*, he appears to endorse the arguments he developed in the earlier work.

For Gaus, not all loves are moral relations. I might love my dog; yet, the reactive attitudes do not govern our relationship. Love between normal adult human beings nonetheless constitutes a moral relation because it consists in following principles of mutual respect and support. While we seldom articulate those principles, they remain present. To count as loving others, we must observe love's constitutive norms and our reactive attitudes must be activated when our beloved violates those norms (Gaus 1990: 362; Gaus 2001). If our beloved deceives us or betrays us by having an affair, we experience overwhelming resentment. This 'mature love' is a moral relation (Gaus 1990: 291, 360).

Friendship is typically a moral relation. True 'moral friendship' in Immanuel Kant's (and Gaus's) sense requires that good friends follow certain norms over long periods of time, such as norms of trustworthiness (Gaus 1990: 287–88; Kant 2009: 233). Otherwise, as with love, the negative reactive attitudes activate, threatening the relationship (though some degree of negative reactive attitudes might hold relationships together). To sustain a moral friendship, we must rationally commit to maintaining it over time. While positive affect does not always accompany our commitment to our friends, violating the commitment will still generate negative reactive attitudes.

Gaus (2001: 38–40) would eventually describe our fundamental values—including relations of love and friendship—as 'principle-grounded'. To count as a good lover or friend one must comply with certain kinds of principles. We cannot enjoy the value of the relationship without observing principles of concern and respect for friends and lovers. But since the value of these thicker moral relations forms a fundamental part of the way we see the world, we value following these principles *as such* and not just in our particular relationships.

To become a mature lover and friend, an individual must morally mature by developing a decentered perspective regarding other persons. Moral maturity consists in seeing others as distinct agents who take themselves to have their own reasons to act (Gaus 1990: 198–203; 1996: 46, 118; 2011: 226). Once we see our social world in this way, the value we place on thicker moral relations requires that we adopt principles of justice and respect for all. Gaus argues on this basis that human beings have strong reason to be moral (1990: 359–67, 379–80), an argument he affirms in later works (1990: 13–4; 1996: vii; 2011: 192–93).

Why should one think that valuing principles that constitute love and friendship requires that we value following them in other contexts? Once we reach a mature, decentered moral perspective, we will experience cognitive dissonance if we value a principle differently in different settings. That is why social morality is ‘deeply embedded in our view of the world’ (or perhaps in the morally mature person’s worldview). Its principles pervade social life. Once we inhabit a decentered perspective, we recognize noninstrumental reasons to follow the principles embedded in our values, and we experience negative reactive attitudes in response to perceived moral violations. Gaus (2011: 192) claims we then have no reason to step outside of that decentered perspective. To renounce the practice of social morality ‘would be to renounce most of what we care for and value. But how could we have reason to do *that*?’

So far, I have only explained why Gaus thinks we should follow these principles in all moral relations. I have not identified those with whom we have moral relations. Gaus admits that we do not have moral relations with all other human beings. He identifies social dynamics that lead societies to expand the moral community (Gaus 2021a: 248–49). One may still wonder whether Gaus provides us with grounds to treat those outside the moral community according to the same principles (Vincent 2021).

If my reconstruction works, we have noninstrumental reasons to value moral relations. Valuing love and friendship rationally requires valuing their constitutive principles in our interactions with everyone, and so we have reason to observe those principles in all our moral relationships.

3. Institutional Threats to Moral Relations

Gaus identifies three features of modern societies that threaten moral relations: (i) coercion (ii) pluralism, and (iii) complexity.

3.1 Coercion

As a liberal, Gaus (1990: 382–99; 1996: 165–66; 2011: xvi, 344–46) worried about the justification of state coercion given its power to harm and destroy. Drawing on Joel Feinberg (1987: 9) and Stanley Benn (1988: 87), Gaus endorsed a presumption against coercion or interference. Attempts to coerce others or interfere with them requires justification. Refraining from doing so does not (Gaus 1990: 382–99; 1996: 165–66; 2011: 345; 2016: 187–98; 2021a: 161–64).

Gaus (2011: 344–45) endorsed Benn’s claim that we think interference requires justification, but that noninterference requires no justification. He offers a similar argument that practical rationality contains a presumption against coercion (Gaus and Nichols 2017). If the presumption is not met, interference and coercion become manipulative, degrading, and harmful. Given that modern nation-states use spectacular amounts of coercion, determining how to justify it is pressing (Gaus 2011: 511–20).

This is not to say that Gaus was a libertarian. He periodically distanced himself from libertarianism (1996: 267; 2011: xv; 2021a: 245n5). His concern with justifying coercion was not sectarian but consonant with the liberal tradition as a whole (2011: xv). He did argue that public justification has a ‘classical tilt’ (2011: 521–28). It favors limited government and the market. Libertarian states, though, cannot adequately supply public goods (2011: 529–44). Thus, coercion threatens moral relations, but it is not always wrong and has legitimate nonlibertarian uses.

3.2 Pluralistic Reasoning

In *The Open Society and Its Complexities*, Gaus claims that disagreement about morality pervades human societies, even tribes and clans, and that such disagreement stretches back tens of thousands of years (2021a: 68). Deep moral disagreement seems an ineliminable feature of social life, especially the social life of open societies. Disagreement challenges justifying moral and political authority in any society (Gaus 2003a: 1–24). Most people agree that political power must be justified, but they disagree about the reasons that justify it. Thus, human beings disagree about what political orders should do. Gaus eschewed John Rawls’s term ‘reasonable pluralism’ (Rawls 2005: xvii) to describe this phenomenon, preferring ‘evaluative pluralism’, which connotes that even exemplary evaluations will not produce agreement about the good, justice, religion, and so on. Indeed, diversity tends to expand; it is ‘autocatalytic’ (Gaus 2021a: 123–50). In his *Contemporary Theories of Liberalism* (2003a), Gaus called this a post-Enlightenment insight. The Enlightenment stressed rationally justifying political authority but erroneously assumed that public justifications would consist in homogeneous reasons (2003a: 1–24).

Gaus first attempted to overcome evaluative pluralism by developing a bargaining model of public justification. Public reason aims at compromise between political perspectives. Gaus (1990: 439–75; 1986) characterized compromise via a David Gauthier-inspired Kalai-Smorodinsky bargaining solution. In *Justificatory Liberalism*, Gaus (1996: vii) admitted that his bargaining model did not provide the degree of normative guidance he had hoped. Gaus (1996: 179–89) acknowledged both ‘inconclusive’ and ‘indeterminate’ public reasoning. Public justification cannot vindicate particular laws and policies nor rules of public or social morality. Because we have distinct reasons, we may disagree about what can be publicly justified; hence Gaus has been associated with ‘convergence liberalism’ (Gaus and Vallier 2009; Gaus 2011: 38–42). Rawls (2005: 133–72) acknowledged this, but Gaus presses the point. He allows reasonable disagreement about justice to run as deep as reasonable disagreement about the

good, something Rawls (2005: xxxvi) was loath to admit (Gaus and Van Schoelandt 2017).

As Gaus was writing *The Order of Public Reason* (2011), he came to believe even *Justificatory Liberalism* (1996) had not taken pluralistic reasoning seriously enough. It had mistakenly stressed shared reasoning as a criterion of sincere deliberation, which barred diverse reasons from public deliberation in some contexts (2011: 288–92). The earlier work had also placed too much weight on political and judicial mechanisms to resolve disputes (2011: 24n48), and it had allowed societies to choose among potentially justified proposals by using adjudicative procedures such as democratic voting (1996: 215–45), but *The Order of Public Reason* observes that we also dispute which procedure to use. We cannot respond by publicly justifying second-order decision procedures to solve these problems because that risks a justificatory regress. To avoid the regress, societies must employ nonrational methods of coordinating on laws, policies, and moral rules (Gaus 2011: 391–93, 409–17). The apex of *The Order of Public Reason* then is Gaus's 'Kantian Coordination Game', which shows how individuals, appealing to both their own evaluative standards and their desire to cooperate with others, can spontaneously coordinate on a moral rule. So long as the rule lies within a set of justifiable proposals, which Gaus (2011: 395–409) called the 'optimal eligible set', it achieves public justification (D'Agostino 2013: 138). *The Order of Public Reason* makes social evolution central to public justification by helping us adopt moral rules when moral reasoning runs dry (2011: 409–23).

Introducing evolutionary mechanisms into public justification means that we cannot provide a complete rationalization of our social institutions (Gaus 2011; 2012: 111). Justifications for social morality, law, and politics exhibit path-dependence. Rules and institutions arise somewhat arbitrarily based partly on historical contingencies, but they achieve public justification all the same (2011: 45). Spontaneous human action selects a member of the eligible set, completing a public justification. Here David Hume and F. A. Hayek figure in Gaus's work (2011: 45–46) as much as Kant and Rawls.

Gaus's next two books push pluralistic reasoning even further. A central idea in *The Tyranny of the Ideal* (2016) is that no one knows what justice is or how to institutionalize it. That is, we not only disagree about justice, but we ought to admit our collective ignorance about it. Public justifications for certain institutions must now also include ideas about how to discover what justice requires (Gaus 2016: 105–49). In *The Open Society and Its Complexities* (2021a), pluralistic reasoning runs deeper still. Theorists may never determine whether we have public justifications for our institutions, given their great complexity (2021a: 142–53).

Gaus initially engaged pluralism to prevent it from undermining moral relations. In the above-mentioned works (2016, 2021a), Gaus turns to harness pluralistic reasoning as an asset for preserving moral relations, creating the 'New Diversity' research program (as discussed in Gaus 2018a).

3.3 Complexity

Gaus (2021a: 7–15) thought large-scale institutions threaten moral relations because modern institutions exhibit complexity. They not only have many parts, but those parts interact. A social system is complex when:

1. The system has a large number of constituent parts.
2. There is feedback between the parts, where the behavior of one part affects the others.
3. The parts are not homogeneous in that the properties of the parts vary between the parts (or at least that the parts display different behaviors in response to the same signal).
4. The behavior of agents can be understood as complying with rules (since Gaus's complex systems are rule-governed systems).
5. The interdependence of the parts is not so great that the system becomes chaotic.
6. Finally, the rules that govern the parts themselves respond to how the parts behave; specifically, the behavior of the parts shapes the rules that govern their behavior. Briefly, there is system-level feedback.

Complexity confounds predicting system outcomes and even grasping the causes of outcomes once they occur (Gaus 2021a: 252–55). Similarly, complexity means we often cannot determine whether a rule is publicly justified for diverse perspectives. To justify liberal institutions, we must nonetheless determine their public justification. *The Open Society and Its Complexities* vigorously addresses this problem by arguing that societies reach public justification because publicly justified institutions help manage complexity even if we do not fully grasp the effects of our moral rules. Pluralism and complexity can have feedback relations; thus, complexity is not a distinct problem from pluralism (2021a: 142–50).

4. Psychological Threats to Moral Relations

Moral relations face another, but equally important, challenge: the psychological threat of the mere first-personal moral point of view (Gaus 2011: 225–32). The mere first-personal point of view consists of a mental orientation where an individual only notices or cares about her or his own values and reasons. It is not necessarily egotistical or selfish. One can both care about others and one's moral commitments (Gaus 1990: 199–200, 275–77). This perspective instead does not recognize other persons as distinct valuers with their own perspectives and reasons, as is the case with psychopaths (1990: 292–99). The mere first-personalist ignores the perspective of others. To put it in Jean Piaget's (1965) terms (which Gaus repeats from 1985 to 2011), the person who reasons only from the first-personal perspective has not 'decentered' or refuses to do so (Gaus 1985: 199; 1990: 199–200; 1996: 23, 46, 118; 2011: 214–16, 226, 256).

Human beings always reason from a first-personal perspective. We act from our own values and reasons. Yet, we can integrate the perspective of others into our own.

If we do so, we take a *compound* first-personal perspective: we view social life from our own perspective and that of others. And we care about both. Gaus (1985: 199; 1990: 298; 1996: 23, 305; 2011: 214–15, 256–57; 2016: 171) follows Lawrence Kohlberg in arguing that decentering is an ordinary stage of psychological maturation, even if some persons never reach it. Decentering also distinguishes normal human beings from psychopaths. Psychopaths cannot care about the perspective of others for others' sake. They cannot understand the affective states of others or value them for noninstrumental reasons.

In many cases, normal human beings choose to ignore the perspective of others and refuse to inquire about their values and reasons. But when they choose to act from this mere first-personal perspective, problems arise. First, the mere first-personalist will experience cognitive dissonance (Gaus 1990: 278–86). She must fight to view others as mere patients despite the fact that she already recognizes them as distinct agents with their own values and reasons (2011: 192). Such dissonance means that she will both affirm her reactive attitudes when she sees them as agents and suppress them when she sees them as patients (2011: 214). The mere first-personal valuer also loses her ability to enjoy moral relations with those who reject her beliefs and values. Her refusal to take their perspective makes her a worse friend and, in some respects, a worse lover (1990: 287–92).

The mere first-personal actor also disrespects others (2018b) though respect for persons is enabled by public justification. If she does not take others' perspective seriously, she cannot rationally hold them responsible for moral violations (2011: 205–31). To blame and punish moral violations properly, we judge the culpability of transgressors (2011: 193–204). And culpability judgments are intrinsically perspectival: we reason about what others believe and value. To hold others responsible aptly, we must determine which reasons they can and cannot see as applying to them. But if people refuse to take the perspective of others into account when deciding how to act, our practice of holding accountable breaks down (Gaus 2011: 183–204; 1990: 13–22). Holding others accountable morphs into disrespectful browbeating and authoritarianism (1996: 123–33, 176, 319; 1996: 182–89; 2011: 32–34, 45, 264, 389, 414, 452). We fail to treat persons as our moral equals (2011: 14–21). These activities damage moral relations because others will resent us if we hold them accountable for actions that they cannot see as wrong (2011: 183). The mere first-personalist irritates moral relations and undermines their benefits.

The Tyranny of the Ideal (2016) addresses another problem for the mere first-personal perspective, or as Gaus calls it, the optimizing stance. This stance insists on a social life organized around what the optimizer views as the best moral rules (2016: 215–19, 222, 232, 244ff.). The optimizing perspective faces epistemic challenges. We have our own perspective on what justice requires and how it can reorganize our institutions. Yet, we seldom know how to apply justice to institutions, especially without help from persons with different ideals (2016: 218). Gaus (2016: 248–50) next argues that if an individual successfully conforms her institutions to her idea, and hers alone, she will know less about how to sustain her ideal. She closes her institutions to prevent other ideals from taking

hold, she prohibits diverse searches, and she learns less about which institutions realize justice. Her optimizing perspective damages the practice of holding accountable. Without this practice, our capacity for joint inquiry weakens (2016: 218). The mere first-personalist undermines her pursuit of justice by acting only from her perspective.

In *The Open Society and Its Complexities*, Gaus (2021a: 189–90) argues that populations that include persons who reason primarily but not exclusively from their own perspective may yield more ‘reconciliation’ than social systems containing only persons who care about reconciliation (though only quasi-Kantians figured into the model, not full-blown optimizers). But even so, if mere first-personalists take power and close open institutions, everyone loses in the ensuing conflict (2021a: 296). As Gaus puts it, ‘a major source of moral conflict is hyper-individualized moral judgment’. The mere first-person perspective can cause even more conflict than ‘tribalism’ (2021a: 196–97).

The mere first-personalist therefore creates the following problems:

1. She cannot stably enjoy moral friendship and mature love with others (Gaus 1990).
2. She browbeats others (Gaus 1996), and she acts as a small-scale authoritarian (Gaus 2011).
3. She becomes overconfident in her grounds for imposing her ideals (Gaus 2011, 2016).
4. She will oppress and tyrannize others if she comes to power (Gaus 2016).
5. And she undermines our capacity to reconcile and self-govern (Gaus 2021a).

5. A Solution Implicit in Moral Relations

Fortunately, moral relations contain the key to their own survival. Gaus (1985; 1986; 2011: 193–205; 2021a: 50–62) long argued (1990: 284; 1996: 319; 2011: 3–5, 45, 101, 163, 178–80, 183–94, 205, 208, 210–11, 213, 218, 225, 261, 293, 299; 2016: 181–82; 2021a: 196) that our reactive attitudes presuppose that we believe they respond appropriately to moral transgressions (Strawson 1974: ch 1). If our belief is correct, then we may blame and punish the transgressor.

We can aptly blame and punish others only when we think they *knew better* or *should have known better* than to act as they did. We must deem them culpable, and culpability is again perspectival: we can only blame those who can see their actions as moral violations based on their own beliefs and values. I can only hold Reba responsible for violating a moral requirement if I think she would affirm the requirement after reflecting on her moral commitments. (Here moral rules might be objective but their interpersonal authority requires joint recognition.)

The rational structure of culpability judgments contains a standard that tells us when our moral demands are appropriate, permissible, and authoritative for others. This is not an external moral standard, one grounded in some response-independent domain of moral truth. Rather, the standard arises from within our moral psychology and

moral emotions. It implies that moral justifications for the terms of social life consist in the reasons and values each person affirms from her own perspective. Gaus calls this ‘public justification’. When we follow a public justification standard, we can sustain moral relations and the benefits they bring. Public justification specifies the appropriateness of our reactive attitudes and thus when moral relations can stably exist.

Moral relations presuppose public justification, and moral practices within those relations then commit us to the standard. Public justification thereby grounds *immanent* criticism: reforming moral practices that presuppose the public justification standard. Public reason ‘liberalism’ evaluates social institutions interpersonally, not from the point of view of the universe, but within social standards of criticism (Gaus 2021a: 191–93). Otherwise, applying the public justification principle could become sectarian and disrupt moral relations, much like religions and ideologies sometimes do.

Social philosophy articulates these standards of ‘reconciliation’ (Gaus 2021a: 56–60). We reconcile through moral relations based on mutually acceptable rules that govern our choices.

6. Public Justification as the Solution

From 1985 to 2020, Gaus insisted that public justification is not an agreement. Agreement is a heuristic for determining which moral rules each person has sufficient reason to accept. As Gaus (1990: 22) says, ‘In place of “What moral principles can we all agree upon?” I ask “What moral principles do we all have reason to accept?”’ We justify a ‘public morality’ when we show it ‘provides all moral persons with good reasons to accept it’ (Gaus 1990: 17). Gaus sought to justify a public morality first, coercion second.

Beyond this basic point, Gaus’s conception of public justification changes considerably across his work. In this section, I formulate a four-part schema for characterizing those changes: (i) diversity of justificatory reasons, (ii) level of idealization, (iii) degree of determinacy of the set of justifiable rules, and (iv) mode of rule choice within that set. I set aside the distinction between telic and deontic justification Gaus uses in *Value and Justification* (1990: 17). *Justificatory Liberalism* (1996) abandons it a mere six years later.

Before I begin, it is important to recognize why Gaus made these changes. His encounter with real-world diversity was the engine of his evolving project. Gaus always wanted to respect persons as they are, and he always recognized the need to reconcile diverse perspectives. At each stage of Gaus’s career, embracing further diversity led him to recast his project. As Gaus studied real-world orders, he found more and more sources of diverse reasoning. Public justification, then, loses much of its abstract and homogenizing character. In this way, Gaus was on a similar developmental path as Rawls, but Gaus’s project begins with less homogeneity and innovates faster. Rawls always thought that political liberalism could embrace justice as fairness. Gaus goes so far as to reject traditional social contract theory at the end of his career. In *The Open Society and Its Complexities* (2021a), he models rule choice as emerging from socially interdependent decisions rather than being a single collective choice.

6.1 Diverse Reasoning

A central cleavage in the public reason literature concerns the set of reasons that publicly justify moral authority, coercion, and so on. ‘Consensus’ theorists restrict the set to shared reasons or to reasons all can evaluate with shared evaluative standards. ‘Convergence’ theorists reject both requirements. They allow diverse reasons into public justification.

Gaus was the leading convergence theorist, but this becomes completely clear only in *The Order of Public Reason* (2011). *Value and Justification* (1990) allowed diverse reasons to figure into public justification: moral personhood requires that we recognize that others have distinct reasons and values. Public morality must be justified for each person on her own terms (1990: 320). But Gaus (1990: 321) sometimes slips into shared reasons talk, saying ‘if I accept that a consideration is a reason for me and not for a fully rational you, it cannot enter into our justificatory argument’.

After *Value and Justification* (1990), Gaus grew increasingly friendly to diverse, pluralistic reasoning. He never rejected it, but diverse reasoning is not central to *Value and Justification*. However, *Justificatory Liberalism* (1996) allowed diverse reasoning, but adopted a sincerity constraint that requires using shared reasons in public deliberation (1996: 139). Even in 2003, Gaus (2003b) characterized justificatory reasons as shared. But by 2009, Gaus forever dispensed with any shared reasons requirement for deliberation or justification (Gaus and Vallier 2009). From *The Order of Public Reason* (2011: 263) forward, Gaus permitted diverse reasoning in public discussion and insisted that public justifications can include diverse reasons (2011: 38–42, 276–92). He embraced pluralistic reasoning and, eventually, used it as a resource for building moral relations. Shared reasons requirements frustrate that aim.

Gaus also revised his model because he adopted a ‘functionalist’ approach to political philosophy. On a functionalist picture, justice adjudicates our competing claims on one another and our shared institutions (1996: 183–84; Van Schoelandt 2020). Philosophers should assess principles of justice based on how well they perform this adjudicating function. By adopting functionalism, one must consider all the reasons that members of the public acknowledge. Otherwise, selected principles of justice cannot resolve their disputes. If we attend only to shared reasons, the functionalist constraint will not be met because people have diverse reasons. Functionalism reinforces the convergence approach.

The functionalist refinement complicates efforts to interpret *The Tyranny of the Ideal* (2016). There Gaus acknowledges nonfunctionalist conceptions of justice, ones that measure the justice of states of affairs without a conflict-adjudicating standard. A conception of justice might solely select the most equal state of affairs. Given the plurality of *concepts* of justice in that work, one could argue that it is not principally concerned with public justification. *The Tyranny of the Ideal* rather shows that pursuing nonfunctionalist ideals of justice can undermine relations of accountability. On this reading, *The Tyranny of the Ideal* internally critiques ideal theories of justice by arguing that they cannot achieve their aims. However, Gaus stresses that we must maintain our practice of accountability,

which is another way of talking about moral relations. Public justification underlies the argument developed in *The Tyranny of the Ideal* (2016).

6.2 Idealization

Theorists must idealize persons to determine their justificatory reasons. Gaus always maintained that we must not attribute reasons to persons based simply on whatever they actually affirm. Our present acceptances depend on bias, ignorance, and limited processing time. But the form and ground of idealization changes throughout Gaus's work. In *Value and Justification* (1990), idealization approached full rationality. Gaus repeatedly says that the only reasons that figure in public justifications are those that fully rational valuers acknowledge. The only alternative requires appealing to reasons that actual persons affirm. In *Justificatory Liberalism* (1996), however, Gaus (1996: 30–32, 130–43; 2011: 235–44) introduces imperfectly idealized agents, agents that are then foregrounded in *The Order of Public Reason* (2011). *Justificatory Liberalism* (1996) initiates Gaus's efforts to develop a theory of *moderate idealization*. At first, he calls this 'open justification' (1996: 30–32, 130–43). Theorists ascribe reasons to agents only once they clean up some of the agents' ignorance, inferential error, and logical and evidential incoherence. Full idealization ascribes foreign reasons to persons, ones they cannot recognize as such. The deliberations of fully idealized agents sharply contrast with excellent real-world human reasoning. *The Order of Public Reason* goes further and describes full rationality as 'a myth' (2011: 235).

The Order of Public Reason also adopts moderate idealization, but Gaus justifies it through his theory of the reactive attitudes. Our practice of accountability presupposes moderate idealization because we ascribe reasons to persons they *should have* known, but currently do not know. These accountability judgments implicitly idealize, but not radically so. Gaus (2011: 232–58) calls this the 'Reasons One Has' standard, which identifies when our moral demands have authority. (In contrast with the 'Reasons There Are', public justification is neutral about the metaphysics of reasons.) Gaus ties idealization to the moral emotions in *The Tyranny of the Ideal* (2016) and in *The Open Society and Its Complexities* (2021a), and thus arguably maintains one account of idealization from 2011 to 2020.

6.3 Determinacy of the Rule Set

Value and Justification ((1990) presumes that rule choice is determinate. Theorists can determine with great specificity whether a rule is publicly justified. Readers find little indication that public justification is inherently inconclusive. It is unclear whether decision procedures must settle which laws and rules authoritatively apply to each person (1996: 130–58, 179–83). But in *Justificatory Liberalism*, Gaus (1996: vii) changes his mind: he advocates publicly justifying social morality struggles to advance 'beyond abstract principles' and so focuses on 'how political institutions cope with this inconclusiveness'. He continues: 'political institutions not only express what can be morally justified, but also respond to our pervasive inability to provide decisive justifications'.

Many laws and policies lack conclusive justifications (Gaus 1996: 151–58). Diverse people rank proposals differently, leaving the public with a range of justifiable proposals for governing behavior. Yet, because no one proposal is vindicated, the set contains inconclusively justified proposals. As noted, societies must appeal to justified decision procedures to select from the set, which then provides a conclusive justification (1996: 223–45). Because democratic choice helps select from the ‘eligible set’ of proposals, Gaus (1996: 258–74) defends ‘adjudicative democracy’.

In *The Order of Public Reason* (2011) Gaus acknowledges pluralism about political decision procedures. Rational deliberation will not produce consensus about which procedures we should use (2011: 391–93). As noted, Gaus addresses this problem by introducing moral evolution into public justification, drawing on Brian Skyrms’s (1996) and Loren Lomasky’s (1987) models of the evolution of the social contract. Gaus’s aim here was to show how members of the public, deliberating and acting freely, converge on rules in the ‘optimal eligible set’ of proposals (2011: 323). The Kantian Coordination Game remains central: each person chooses whether to comply with an eligible moral rule without using any top-down procedure and they then converge on a rule through social interaction (2011: 323–33; 391–409). Free individuals coordinate when collective deliberation cannot lead to agreement. The evolution of morality thus plays a central role in producing public justifications and preserving moral relations (2011: 409–17).

In *The Tyranny of the Ideal* (2016) Gaus grapples with *ignorance* about justice. Even idealized members of the public struggle to grasp the truth about justice and how they should use justice to reform institutions (2016: xix, 90, 92, 96). Theorists must now identify public justifications for social spaces that allow social experiments to uncover what justice requires (2016: 184–207). Rules can achieve public justification by aiding social discovery (though learning from social experiments faces difficulties in a complex system; Gaus 2016: 89f.).

The Tyranny of the Ideal does not alter the Kantian Coordination Game of *The Order of Public Reason* (2011), or at least not much. Still, the former stresses norm ambiguities that allow moral innovators to choose new behaviors while remaining within the eligible set, and this in turn helps uncover new publicly justifiable rules. *The Tyranny of the Ideal* (2016) then emphasizes that open societies improve their moral ‘constitutions’ by adopting these new rules.

In *The Open Society and Its Complexities* (2021a) Gaus stresses that we often do not know whether rules are publicly justified. Some institutions are so complex that we cannot determine whether diverse persons have sufficient reason to obey them (2021a: 202). One reason is that open societies engender ‘autocatalytic diversity’ (2021a: 128–41). Our disagreements expand over time (though some modes of reducing complexity can reset diversity). *The Open Society and Its Complexities* combats indeterminacy by arguing that traditional liberal institutions, like markets (2016: 202–05) and democracy (2021a: 288–94), help us discover how to self-govern. In that book, Gaus also draws special attention to polycentric orders, that is, orders where collaborating, distinct governance systems produce social

outcomes (Ostrom 1990; Gaus 2021a: 183–84). Gaus thereby strengthens the case that liberal institutions are publicly justified.

At the same time, in *The Open Society and Its Complexities* (2021a) Gaus also diversifies his conception of emergent order. In *The Order of Public Reason* (2011), he had sometimes confused evolved social rules with emergent rules. But some orders, like social media, emerge by design (2021a: 120).

6.4 Rule Selection within the Set

Gaus (2011: 264–75) uses a deliberative model to narrow the optimal eligible set. The model resembles Rawls's original position. The theorist first idealizes representative persons to reason through shared considerations while ignoring their diverse reasons. The theorist then lifts these limitations to see whether the agreement is justified for each diverse perspective based on the idealized representative persons' total set of reasons (Rawls 1971; Gaus 2011: ch 5). Gaus here does not endorse a shared reasons requirement. Rather, the deliberative model remains a heuristic for determining what people can converge upon. Theorists determine full convergence only in the second stage of justification. This second 'rule choice' stage depicts how idealized agents agree on shared rules.

The deliberative model presented in *Value and Justification* shows compromises between ideological perspectives (1990: 439–72; following Gauthier 1986). Contractualist justification finds bargaining points between competing ideologies. Contractors fairly divide social space to realize their diverse values, but Gaus (1990: 439–42) never limited public justification to conceptions of the good, unlike Rawls (1971, 2005). In *Justificatory Liberalism* (1996) Gaus develops an inconclusively justified set of proposals and uses publicly justified decision procedures to then conclusively justify proposals (1996: 215–45). In *The Order of Public Reason* (2011: 471ff.) Gaus notes continuity with his earlier arguments in *Justificatory Liberalism* (1996) for political institutions. But the Kantian Coordination Game sharply distinguishes the two works. Rule choice in *The Tyranny of the Ideal* (2016) tracks the account given earlier in *Order of Public Reason* (2011), with the caveat that moral innovators can exploit rule ambiguity to find new ways of living together. Norm ambiguity matters because norms and rules do not tell us what to do. They instead prohibit actions, leaving other lines of conduct open (Gaus and Nichols 2017).

In *The Open Society and Its Complexities* (2021a) Gaus revises the Kantian Coordination Game. In the earlier *The Order of Public Reason* (2011) Gaus had implicitly represented agents as assigning the same weights to their desire to follow their own standards and their desire to reconcile with diverse others. But in *The Open Society and Its Complexities*, Gaus recognizes that some people place greater weight on satisfying their own standards, whereas others will care more about reconciliation (2021a: 177–91). The question whether members disagree about how to weigh these factors (2021a: 159) was not addressed in *The Order of Public Reason* (2011).

If members of the public differ in how they balance these two factors, how can they coordinate on an eligible rule? To answer this question, Gaus in *The Open*

Society and Its Complexities (2021a) deploys an agent-based model, a model that simulates interactions between independent agents to examine system-level effects. The model attempts to show that societies with diverse weightings may be *more* likely to coordinate on a moral rule than societies with homogeneous weightings (2021a: 189–91)—a surprising result. The new model adopts an element of self-organization that previous models lacked: agents in the model interact and self-organize not only despite their different commitments, but because of their differing desires for reconciliation (2021a: 187–89). Gaus then argues that open institutions, such as markets and democratic governance, achieve public justification because they harness diversity to aid social cooperation. Sectarian institutions undermine the self-organizing features of open societies, and thus they fail the test of public justification (2021a: 248–94). In general, complexity can obscure public justification, but Gaus argued that some institutions become publicly justified because they manage complexity in such a way as to reveal other public justifications. Democracy has public justification because it uncovers public justifications for laws and policies. That is, complexity challenges public justification and manages it. The same holds true for pluralism.

Display 1 contrasts Gaus's five models of public justification. His five books lie on the x-axis, his conceptions of public justification on the y-axis.

Gaus describes *failures* of public justification as raising the problems discussed above, but notes that they are all problems for having moral relations with others. *Justificatory Liberalism* (1996) says that browbeating is what occurs without public justification, while *The Order of Public Reason* (2011) says that without public justification one becomes a small-scale authoritarian. That too undermines moral relations. And while neither 'browbeat', 'authoritarian', or their cognates appear in *The Tyranny of the Ideal* (2016) or *The Open Society and Its Complexities* (2021a), the ideas of moral relations and accountability appear frequently.

7. Public Justification Helps Neutralize Institutional and Psychological Threats

Public justification neutralizes the institutional and psychological threats to moral relations.

7.1 Institutional Threats

When coercion and political power are publicly justified, they sync with each person's reasoned perspective. Coercion no longer oppresses us because we impose it on ourselves through our moral emotions and free deliberation (1996: 122, 184; 2011: 479–508). Coercion enhances our positive freedom, the freedom of our reason to endorse its own felt constraints (2011: 31–36). Under conditions of diversity, only liberal institutions reach public justification because only liberal institutions protect each person's or group's ability to pursue what they value. Liberal institutions also guarantee negative freedom by limiting government power and stopping interference from others (1990: 466–69; 1996: 199–203; 2011:

Display 1: Gaussian Public Justification Models

Public Justification ↓	VJ	JL	OPR	TI	OSC
Diverse reasons?	Unclear	Yes, but . . .	Yes	Yes	Yes
Idealization	Full Rationality	Moderate, Open Justification	Moderate, Accountability	Moderate, Accountability	Moderate, Accountability
Rule Set Determinacy	High	Moderate, Political	Low, social (OES)	+ norm ambiguity	+ justification ignorance
Rule Choice	Collective Bargainers	Collective Adjudicators	Kantian Coordinators	+ norm innovators	+ diverse reconcilers

VJ = *Value and Justification* (1990); JL = *Justificatory Liberalism* (1996); OPR = *The Order of Public Reason* (2011); TI = *The Tyranny of the Ideal* (2016); OSC = *The Open Society and Its Complexities* (2021a).

334–86). This argument undergirds Gaus’s case for the public justification of rights in *Value and Justification*, *Justificatory Liberalism*, and *The Order of Public Reason*.

Publicly justified moral and legal rules overcome the challenge of pluralism. Our practice of accountability, blame, and punishment make feasible a shared social life (2011: 232–59; 1996: 173). Even if we cannot live by shared values, we nonetheless reconcile by adopting shared rules for different reasons. Gaus (2021a: 152–59) always stresses the feasibility of reconciliation even as he allows more and more diversity into his conception of public justification, but by *The Open Society and Its Complexities* (2021a), he worries that public justification is feasible only in certain contexts.

Finally, public justification can overcome the threats posed by institutional complexity. Publicly justified rules help self-governance through self-organization. Self-organized self-governance requires that people see their shared moral rules as ones they can internalize as their own (2021a: 269–72). Societies can stabilize social norms across diverse populations only when people can internalize those norms. In contrast, imposing unjustified norms and rules leads people to defect from them. Such impositions may even create blowback against governing institutions (Barrett and Gaus 2020). Good government will then deteriorate. To reverse this deterioration, state officials must employ destructive and oppressive state power that can easily backfire. If governors stick to imposing publicly justified norms, however, they can gradually improve social outcomes like public health and education at the micro-level or meso-level, if not the macro-level. It is the reflexivity of policy that drives complexity: government is but one actor, and its effects depend on how citizens respond to it, and on how government responds in turn. But interactions governed by publicly justified rules reduce complexity and increase predictability. People tend to follow rules they view as justified. In the process, they become more governable.

7.2 Psychological Threats

For Gaus, public justification arises naturally from a decentered moral perspective—the compound first-personal point of view. His major works from 1990 to 2021 speak with one voice (Piaget 1965; Gaus 1990: 275–78; 1996: 118; 2011: 214). From a decentered perspective, we see others as separate individuals with distinct reasons, and yet as we decenter, we build their perspective into our own. We shape our personal choices while we take others’ perspectives into account, rather than just considering their well-being. Public justification flows from how decentered persons think and act.

If we hold others to publicly justified rules and also abide by them ourselves, we secure two great goods: (i) recognizing the moral personality of others and (ii) avoiding cognitive dissonance and feelings of inferiority. We maintain moral relations with others consistent with our own moral psychology.

To identify the first good, imagine Alf imposes an unjustified moral rule on Betty. Since Alf has decentered, he knows that the rule is not acceptable to Betty. He can no longer coherently hold Betty accountable for moral violations. Moral relations break

down. But if Alf confines himself to publicly justified rules, he takes Betty seriously. Alf decenters and acknowledges Betty's moral personality (Gaus 2011: 225–32). As a result, moral relations revive.

To identify the second good, suppose Alf endures subjection to unjustified rules. He must abide by rules that violate his evaluative standards. This causes Alf much frustration and resentment. He allows himself to be bullied. Here the psychological costs are plain: who wants to be pushed around? And who wants to be psychologically vulnerable to a repressive regime? On the other hand, if Betty's society's rules are publicly justified, holding her to those rules fits her evaluative standards. Yes, Betty endures guilt and shame, but her guilt is self-imposed. Betty's shame matters because her critics only hold her accountable if they take Betty's perspective seriously. Such a community is not always pleasant, but it is functional and even psychologically healthy. Gaus (2015) expressly acknowledges the good of recognizing moral personality. He does not explicitly acknowledge the good of avoiding dissonance and feelings of inferiority, but the text strongly supports the attribution.

Public justification also neutralizes the threats of browbeating and authoritarianism. When Alf demands that Betty follow the rules of social morality, he disrupts moral relations if his reactive attitudes target Betty inaptly (Gaus 1996: 123). As Gaus stresses in *The Open Society and Its Complexities*, punishing persons for violating norms they reject invites countersanction and resistance (2021a: 260–63). If Alf restrains his behavior by demanding only that Betty follow justified rules, he avoids browbeating her. Alf and Betty remain free despite being under one another's authority (Gaus 2011: 25, 30) because justified authority is not browbeating, only unjustified authority is. Alf and Betty also stabilize their practice of accountability, and they can jointly discover what justice requires (2016: 218, 230); they can even govern themselves well (2021a: 74, 194, 200, 234, 279, 282).

Public justification neutralizes the psychological and institutional threats to moral relations. Alf and Betty overcome the mere first-personal point of view and help ensure that their social order withstands coercion, pluralism, and complexity.¹

8. Reforming Political Philosophy

Gaus sharply criticizes contemporary political philosophy. Philosophers use methods of inquiry, such as armchair reasoning, that obscure the social dimensions of their judgments. They may not realize that they use their theories of justice to control and direct others' lives in ways that do not make sense to those others. This leads to authoritarianism, tyranny, and closed societies.

As noted, Gaus rejected the optimizing stance in social life. People should not impose social rules that advance their values at the expense of others. The optimizing stance also drives Gaus's criticism of contemporary political

¹ One hopes Alf and Betty live happily ever after.

philosophy. Gaus thinks philosophers are the ultimate optimizers: they divine a theory of justice and impose it on those who disagree.

Gaus's kickstarted his critique of the political philosophy profession in *Justificatory Liberalism* (1996: 46), but that is not a central theme of the book. *The Order of Public Reason* contains the full-fledged critique (2011: 549). The philosopher intuits principles of justice as moral blueprints that arrogantly and dangerously place everyone under his authority: 'The moral philosopher, impressed by her own rectitude, who proclaims that as we move away from her ideal we move away from a justified true morality, becomes the unwitting enemy of a free moral order' (2011: 549).

In *The Tyranny of the Ideal* (2016), Gaus turns up the heat. That book arguably concerns methodology as much as political philosophy itself. From the first page, it criticizes philosophers for downplaying formal modeling while using less rigorous informal models (2016: xv–xx). Gaus here claims that political philosophers can make more progress by working in diverse teams to discover what justice requires. Gaus (2016: 105–49) thought political philosophers did not grasp how ignorant they are about justice and how to establish it.

In *The Open Society and Its Complexities* (2021a) Gaus makes this point in a different way. Political philosophers appeal to the moral intuitions of their social subgroup: those with WEIRD (Western, Educated, Industrialized, Rich, and Democratic) psychologies (2021a: ii). Political philosophy 'has degenerated into reports of the intuitions and judgments of a section of upper-middle class Americans and Europeans' where 'there is clear evidence that this is a remarkably biased group'. Since WEIRD psychologies are so unusual, our moral intuitions may be biased in favor of some moral considerations and against others. Gaus does not reject the egalitarian morality of the WEIRD. Instead, studying WEIRD psychologies should humble political philosophers because such study exposes their moral judgments as parochial. Recognizing this, political philosophers should not philosophize based on intuitions alone.

The problems with the mere first-personal point of view imply not merely that we must *live* differently, but that we must *philosophize* differently. Political philosophers can reform and improve only by taking the social sciences seriously. Gaus tried to practice what he preached. All five of Gaus's treatises richly engaged the sciences, each time with somewhat different sciences. For example, *Value and Justification* (1990: 25–78) draws on moral psychology and theories of psychological development, while *Justificatory Liberalism* draws on cognitive psychology (1996: 38–62), and *The Order of Public Reason* draws on game theory and evolutionary theory (2011: 53–181). *The Tyranny of the Ideal* draws on complexity theory and the modeling of discovery in the natural sciences (2016: 105–49). And *The Open Society and Its Complexities* adds evolutionary biology and anthropology (2021a: 19–109).

9. Gaus's Project in His Own Words

In Gaus's final paper, he concisely describes his project and adds a critical observation:

For much of my career I have developed an account of how people who deeply disagree about the basis of normativity and have serious disagreements about what is right and wrong can nevertheless converge on common social-moral rules for cooperative living. The core idea is ‘convergent normativity’: while we disagree on many of the grand issues of morality we can, in the interests of achieving a cooperative order based on relations of mutual moral accountability, reconcile on common rules that each of us, for her own reasons, endorses. This tale draws on empirical literature concerning moral psychology, norms, social cooperation, punishment and practice of accountability. So, I trust it is not a ‘just so’ story. But its lesson, that each can simultaneously affirm her own moral perspective while reconciling with others on common rules, abstracts from numerous issues. It assumes, first and foremost, that most people do not simply dismiss the normative perspectives of others as befuddled crooked thinking, but acknowledge them as intelligible ways to understand social morality. In the current environment, many people apparently prefer the joys of anger and aggressive self-righteousness to reconciliation with others. Why reconcile when the enemy can be crushed? I am pessimistic about the future of social orders with this ethos, but if that is the ethos the itch my story seeks to relieve—how can a truly diverse society share a moral framework?—is not felt. These views are attracted to stories about life among truly straight-thinking people, and damning those crooked thinkers who stand in the way of true justice. (Gaus 2021b)

We can retreat into the mere first-personal point of view by categorizing others as enemies. Many gleefully do so. But for those of us who lack the conquistador’s soul, Gaus shows us a better way. He leaves us with a pressing personal question: what kind of people do we want to be?

KEVIN VALLIER 

BOWLING GREEN STATE UNIVERSITY

kevinvallier@gmail.com

References

- Barrett, Jacob, and Gerald Gaus. (2020) ‘Laws, Norms, and Public Justification: The Limits of Law as an Instrument of Reform’. In S. A. Langvatn, M. Kumm, and W. Sadurski (eds.), *Public Reason and Courts* (New York: Cambridge University Press), 201–28.
- Benn, Stanley. (1988) *A Theory of Freedom*. New York: Cambridge University Press.
- D’Agostino, Fred. (2013) ‘The Orders of Public Reason’. *Analytic Philosophy*, 54, 129–55.
- Feinberg, Joel. (1987) *Harm to Others: The Moral Limits of the Criminal Law*. Vol. 1. New York: Oxford University Press.
- Gaus, Gerald. (1985) ‘On Community and Justice: A Reply to Professor Golding’. *Bulletin of the Australian Society of Legal Philosophy*, 9, 197–204.
- Gaus, Gerald. (1986) ‘Subjective Value and Justificatory Political Theory’. *Nomos*, 28, 241–69.
- Gaus, Gerald. (1990) *Value and Justification: The Foundations of Liberal Theory*. Cambridge: Cambridge University Press.

- Gaus, Gerald. (1996) *Justificatory Liberalism: An Essay on Epistemology and Political Theory*. New York: Oxford University Press.
- Gaus, Gerald. (2001) 'What Is Deontology? Part Two: Reasons for Action'. *Journal of Value Inquiry*, 35, 179–93.
- Gaus, Gerald. (2003a) *Contemporary Theories of Liberalism: Public Reason as a Post-Enlightenment Project*. Minneapolis: Sage Publications.
- Gaus, Gerald. (2003b) 'Liberal Neutrality: A Radical and Compelling Principle'. In George Klosko and Steven Wall (eds.), *Perfectionism and Neutrality* (Lanham, MD: Rowman & Littlefield), 137–65.
- Gaus, Gerald. (2008) 'Reasonable Utility Functions and Playing the Cooperative Way'. *Critical Review of International Social and Political Philosophy*, 11, 215–34.
- Gaus, Gerald. (2011) *The Order of Public Reason*. New York: Cambridge University Press.
- Gaus, Gerald. (2014) 'The Good, the Bad, and the Ugly: Three Agent-Type Challenges to the Order of Public Reason'. *Philosophical Studies*, 170, 563–77.
- Gaus, Gerald. (2015) 'On Dissing Public Reason: A Reply to Enoch'. *Ethics*, 125, 1078–95.
- Gaus, Gerald. (2016) *The Tyranny of the Ideal*. Princeton: Princeton University Press.
- Gaus, Gerald. (2018a) 'The Complexity of a Diverse Moral Order'. *Georgetown Journal of Law and Public Policy*, 16, 645–80.
- Gaus, Gerald. (2018b) 'Respect for Persons and Public Justification'. In Richard Dean and Oliver Sensen (eds.), *Respect for Persons* (New York: Oxford University Press), 1–22.
- Gaus, Gerald. (2021a) *The Open Society and Its Complexities*. New York: Oxford University Press.
- Gaus, Gerald. (2021b) 'Philosophical Fables' Unpublished manuscript fragment.
- Gaus, Gerald, and Chad Van Schoelandt. (2017) 'Consensus on What? Convergence for What? Four Models of Political Liberalism'. *Ethics*, 128, 145–72.
- Gaus, Gerald, and John Thrasher. (2021) *On Philosophy, Politics, and Economics*. 2d ed. Princeton: Princeton University Press.
- Gaus, Gerald, and Kevin Vallier. (2009) 'The Roles of Religious Conviction in a Publicly Justified Polity: The Implications of Convergence, Asymmetry and Political Institutions'. *Philosophy and Social Criticism*, 35, 51–76.
- Gaus, Gerald, and Shaun Nichols. (2017) 'Moral Learning in the Open Society: The Theory and Practice of Natural Liberty'. *Social Philosophy and Policy*, 34, 79–101.
- Gauthier, David. (1986) *Morals by Agreement*. New York: Oxford University Press.
- Hayek, F. A. (1973) *Rules and Order*. Vol. 1 of *Law Legislation and Liberty*. Chicago: University of Chicago Press.
- Kant, Immanuel. (2009) *The Metaphysics of Morals*. Translated by Mary Gregor. Edited by Mary Gregor. New York: Cambridge University Press.
- Lomasky, Loren. (1987) *Persons, Rights, and the Moral Community*. New York: Oxford University Press.
- Ostrom, Elinor. (1990) *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge: Cambridge University Press.
- Piaget, Jean. (1965) *The Moral Judgment of the Child*. Translated by Marjorie Gabain. New York: Free Press.
- Rawls, John. (1971) *A Theory of Justice*. New York: Oxford University Press.
- Rawls, John. (2005) *Political Liberalism*. 2d ed. New York: Columbia University Press.
- Skyrms, Brian. (1996) *Evolution of the Social Contract*. Cambridge: Cambridge University Press.
- Strawson, P. F. (1974) *Freedom and Resentment and Other Essays*. New York: Routledge.
- Van Schoelandt, Chad. (2020) 'Functionalist Justice and Coordination'. *Social Theory and Practice*, 46, 417–40.
- Vincent, Michael. (2021) 'A Sympathetic Critique of Gaus's *The Order of Public Reason*'. PhD diss., University of Queensland, Brisbane.