Symmetry & The Form of Ethical Objectivity

Santiago De Jesus Sanchez Borboa

Abstract

Many contemporary arguments in moral philosophy appeal to an "objective standpoint" in practical reason. Yet a dialectical gap remains: even if practical reasons admit of some objective characterization, it is unclear why objectivity should require robust impartiality rather than a thinner form of coherence or intersubjective stability. This paper proposes a bridge principle: objectivity in practical evaluation is appropriately understood as invariance under transformations that preserve all normatively relevant structure. I argue that once any nontrivial objectivity constraint is accepted, at least temporal coherence across normatively identical cases, symmetry requirements "bootstrap," i.e., there is no non-arbitrary stopping point short of invariance under person-permutation, perspective, and ultimately, invariance under local transformations of evaluative framing. These invariances yield a compact route to a structural priority of moral reasons over merely prudential or partial considerations: prudential reasons can count only insofar as they are representable as enrichments consistent with the relevant invariances. An appendix offers lightweight formalizations that make the invariance claims and the resulting conditionality of prudence fully explicit.

Introduction

A striking feature of much recent work in moral philosophy is its reliance on the idea that practical reasoning admits of an "objective" standpoint. Constitutivist, rationalist, and impartialist approaches all promise a route to moral requirements that does not depend on parochial moral psychology, local conventions, or the contingencies of desire.[1] If successful, they would vindicate the thought that morality is not merely one optional source of reasons among others, but is bound up with what it is to deliberate well.

A familiar challenge for such attempts is that it can be difficult to explain why impartiality is not a substantive moral add-on. Constitutivist approaches risk grounding normativity in a comparatively thick account of agency. Parfitian approaches can appear to rely on contentious intuitions. Nagelian appeals to an "objective standpoint" can strike readers as suggestive but under-specified. Even if we grant that practical reasons are not wholly subjective, it is not obvious why the relevant constraints must take the form of robust impartiality rather than a thinner ideal of internal coherence, self-constitution, or intersubjective convergence.[2]

This paper proposes a modest answer: the missing bridge is a symmetry constraint. In many domains, paradigmatically, the natural sciences, objectivity is fruitfully understood as invariance under transformations that preserve what is relevant to the phenomenon under study. When we call a description or judgment "objective," we typically mean that it does not depend on features of representation irrelevant to what is being assessed. In ethics, a closely related idea is at work whenever we condemn a practical judgment as arbitrary or question-begging: a change in "point of view" that does not alter the relevant facts nonetheless alters the verdict. I develop this thought as a bridge principle:

> **Objectivity-as-Invariance (OI):** A practical evaluation counts as objective only insofar as it is invariant under transformations that preserve all normatively relevant structure.[3]

This principle is deliberately ecumenical.[4] It does not presuppose consequentialism, contractualism, Kantianism, or any specific moral theory, nor does it require controversial metaphysical commitments.[5] It expresses a familiar constraint on reason-giving: if a transformation changes nothing that rationally matters, then a shift in evaluation is arbitrary. A weak

[1] Notable examples include Korsgaard's constitutivist project (1996, 2009), Parfit's defense of objective reasons (1984, 2011), Nagel's appeal to an impersonal standpoint (1978, 1986), and Sidgwick's attempt to derive rational benevolence from impartiality and diachronic coherence in Book III, Chapter XIII of *The Methods of Ethics* (1981) and Gewirth's argument from the structure of agency to a principle of generic consistency (1978). The relationship between the present argument and Gewirth's is discussed in note 9 below.

[2] For a careful overview of the agent-relative/agent-neutral distinction, see Bykvist (2018). The present paper is neutral with respect to different ways of drawing the distinction.

[3] Throughout the paper, I mean for 'evaluation' to be broad enough to include prescriptions (understood broadly as action-guiding evaluations that something ought to be).

[4] This bridge principle is a practical analogue of a familiar philosophical theme: to call a claim "objective" is, in part, to say that it is stable across a suitable range of admissible variations in perspective or representation (cf. Nozick's (2001) invariance-based discussion of objectivity). I cite Nozick only as a clear precedent; nothing in the argument depends on his particular account.

[5] The argument does not presuppose any particular normative primitive. I often speak of "reasons" because

form of OI is already implicit in diachronic agency: a subject who takes itself to persist over time and to act for reasons must treat normatively identical cases as calling for the same assessment irrespective of when they occur.

The central claim of the paper is that once any nontrivial invariance constraint is admitted, symmetry requirements "bootstrap." If we accept that time-shifts preserving normatively relevant features cannot change practical verdicts, then mere changes in agent labels cannot do so either, nor can mere changes in embodied perspective or evaluative framing that do not alter the relevant structure. The upshot is a compact route from minimal structural commitments about rational agency to stronger impartiality constraints than are usually derived in the contemporary literature.[6]

The argument proceeds in four steps. Section 1 articulates OI and motivates it as a weak but indispensable norm of rational assessment. Section 2 develops the symmetry-bootstrapping argument. Section 3 draws the normative consequence: moral reasons have structural priority over merely prudential or partial reasons.[7] Section 4 clarifies how formal tools from symmetry theory can make invariance constraints explicit, and contrasts the framework's diagnostic structure with Scanlon's contractualism and Rawls's original position. An appendix offers lightweight formalizations.[8]

The claim is conditional and structural: if one accepts that practical objectivity is a genuine ideal, and if one accepts that objectivity is appropriately cashed out as invariance under normatively irrelevant transformations, then substantial impartiality constraints follow, and with them, a principled explanation of why moral reasons cannot be treated as optional competitors to prudential ones. If the argument succeeds, it supplies a compact bridge from a minimal structural rational ideal to robust impartiality without importing substantive moral premises, and it motivates a research program in which we can ask systematically which candidate moral principles respect which invariances, what kinds of "symmetry breaking" are intelligible, and what justificatory costs such breaking incurs.

---

the terminology is widely shared, but the line of reasoning is neutral among axiological, deontic, and aretaic frameworks.

[6]For influential discussions of structural requirements of rationality, see Broome (1999, 2013), Raz (2005), and Kolodny (2005). Nothing in the present argument turns on whether rational requirements reduce to reasons. The appeal to diachronic coherence is only to a minimal constraint of non-caprice implicit in practical deliberation.

[7]In this paper, I use 'moral reasons' as a label for the subset of practical reasons that survive the objectivity-as-invariance constraints defended here. This is not offered as a full analysis of the moral/non-moral distinction.

[8]The idea that moral judgments must satisfy a universalizability constraint has an important precedent in R. M. Hare's prescriptivism. Hare argues that universalizability is a requirement imposed by the logic of moral language (Hare 1952, 1963). The present paper shares Hare's conviction that impartiality constraints can be derived rather than simply postulated, but departs from his approach in three respects. First, the argument here is structural rather than linguistic: the invariance requirements are grounded not in the semantics of moral terms but in what it is for any practical evaluation to count as objective. Second, Hare's universalizability corresponds roughly to one station in the bootstrapping trajectory (person-permutation invariance), whereas the present framework begins from a thinner starting point (diachronic coherence) and extends to stronger constraints (perspective invariance, local invariance) absent from Hare's account. Third, the connection to symmetry reasoning in the natural sciences, which supplies both the methodology and the formal vocabulary, is absent from Hare's linguistic framework. Universalizability, as Hare understood it, emerges here as a special case of a more general pattern.

1 Objectivity-as-Invariance: A Minimal Norm of Rational Assessment

The argument turns on a simple bridge principle about what it is for a practical evaluation to count as objective. The principle is not a substantive moral claim but a claim about rational assessment: when we evaluate reasons, maxims, actions, or policies, we routinely distinguish between features that bear on what is being assessed and features that are merely artifacts of representation. The core thought is that objectivity consists, at least in part, in filtering out these representational artifacts. As formulated above: a practical evaluation counts as objective only insofar as it is invariant under transformations that preserve all normatively relevant structure.

The phrase "normatively relevant structure" is intentionally ecumenical. OI does not presuppose any particular theory of what counts as relevant. Rather, it states a conditional constraint: once we have fixed what matters, we can test whether a purported evaluation is sensitive to anything beyond what matters. OI is best understood as a thin norm governing rational appraisal, akin to familiar constraints against equivocation: do not let your evaluation track what you yourself regard as irrelevant.

The following considerations motivate OI as a weak but indispensable norm.

1.1 OI in scientific reasoning and ordinary practical life

In the sciences, objectivity is routinely associated with invariance, not because scientists fetishize symmetry, but because invariance marks a distinction between what belongs to the phenomenon and what belongs to our description of it. If a model predicted different reaction profiles merely because temperature is expressed in Celsius rather than Fahrenheit, we would treat this as a failure of representation, not a competing empirical hypothesis. The point is modest: where representation can vary while the target remains the same, objectivity requires insensitivity to that variation.

A closely related norm is already operative in ordinary practical reasoning. Most of us take ourselves to be temporally extended agents who form plans and sustain commitments. That self-conception brings with it a minimal demand: if the normative structure of a choice situation remains the same, it should not be evaluated differently merely because time has passed. If yesterday I judged it wrong to break a promise, and today the relevant facts are unchanged, then judging differently simply because it is now today is not updated reasoning; it is arbitrary reversal. Similar points arise with relabeling: if I judge a rule for distributing burdens acceptable when I think of myself as "A" and my colleague as "B," but unacceptable when I swap labels while holding fixed all relevant facts, my judgment is tracking the label rather than the structure. This is precisely what ordinary moral criticism targets when it complains of "double standards."

1.2 OI as a condition on the distinction between reason and caprice

The deepest motivation for OI is that without some invariance constraint, the distinction between reasoning and caprice erodes. To give reasons is to present considerations as counting

in favor of a conclusion in a way responsive to features of the situation, not to arbitrary shifts in description. When an agent's evaluation changes under a transformation that preserves all normatively relevant structure, the agent lacks the resources, by its own lights, to explain the change as reason-responsive. OI helps articulate the contrast between "I changed my mind because something that matters changed" and "I changed my mind for no reason."

This claim does not assume an ambitious picture of rationality. It requires only that the agent's practical outlook support a stable distinction between reasons and mere shifts of attitude. Invariance under normatively irrelevant transformations is one of the weakest constraints capable of doing that work. To reject OI is not merely to adopt a different moral theory but to loosen, perhaps fatally, the rational discipline on which any theory's claims to objectivity depend.

OI is thus best framed as a bridge principle. Most reasonable parties in metaethical disputes can accept that if there is to be anything worth calling "objective" practical evaluation, it must not be hostage to irrelevant representational choices. OI simply makes explicit what is presupposed when we criticize hypocrisy, condemn ad hoc exceptions, or demand consistency over time. The rest of the paper exploits the fact that once OI is accepted even in its weakest diachronic form, the range of transformations that plausibly count as normatively irrelevant expands. That expansion is the subject of the next section.

It's worth emphasizing that OI is not a constitutive condition of agency as such. The argument does not claim that agents who violate impartiality constraints fail to qualify as agents. The argument's conditional character is thus threefold: if one aspires to objectivity in practical reasoning, if objectivity is appropriately understood as invariance under normatively irrelevant transformations, and if one accepts even the minimal invariance of diachronic coherence, then the bootstrapping to robust impartiality follows. The next section's bootstrapping argument does not derive impartiality from the bare existence conditions of deliberation. It shows that once an agent takes the aspiration to objectivity seriously, the range of transformations that must be treated as irrelevant expands in a principled and non-arbitrary way, just as, in physics, the aspiration to empirical adequacy generates disciplined pressure to identify representational artifacts and to ensure that no prediction depends on features recognized as belonging to the description rather than to the phenomenon.

2 A Symmetry-Bootstrapping Argument from Minimal Coherence to Robust Impartiality

Recall that OI does not dictate what counts as normatively relevant. It imposes a structural discipline: whatever you count as relevant, you must be invariant under transformations that preserve it. The bootstrapping argument proceeds as a dilemma: for each candidate transformation, either (i) you identify a normatively relevant structural feature the transformation fails to preserve, in which case invariance is not required, or (ii) you cannot identify such a feature, in which case refusing invariance is arbitrary. The argument's force lies in the justificatory burden it places on those who would resist each bootstrapping step.

2.1 Step one: diachronic coherence as minimal symmetry

Begin with a modest claim about agency. Many practical judgments purport to be reason-guided in a way that extends across time: we plan, sustain commitments, and criticize ourselves for inconsistency. These practices presuppose that temporal location, by itself, is not normatively decisive. If two situations are identical with respect to everything the agent takes to matter, a mere shift in time index is not a rational ground for a different verdict. The relevant norm is not that one must keep every plan; it is that changes in evaluation require changes in what one takes to be normatively relevant. A purely temporal translation, holding those features fixed, does not provide such a change.

This first step establishes an important template: once a transformation is acknowledged to be normatively irrelevant, allowing evaluations to vary under it licenses arbitrariness. That template drives the remaining bootstrapping steps.

2.2 Step two: from temporal invariance to person-permutation invariance

Having granted that temporal location is normatively irrelevant, consider the analogous question about who occupies a role in a moral situation. Hold fixed all facts the evaluator takes to be relevant (interests, harms, benefits, intentions, relationships, commitments, rights) and simply permute agent identities while preserving the relevant structure. If evaluation changes merely because the label changes, it is difficult to avoid the conclusion that the evaluation is tracking something the evaluator cannot rationally defend as relevant.

Consider the "double standard" test. Imagine I endorse a principle permitting me to free-ride on a cooperative scheme when it benefits me. Now consider the structurally identical case where positions are swapped: someone else free-rides on me under the same conditions. If I now judge the act impermissible, what explains the difference? If the explanation is simply "because it is me," then the evaluation depends on a feature (mere identity), which does not articulate a normatively relevant distinction. Identity matters in virtue of relationships, roles, and obligations; those are part of the relevant structure. The claim is narrower: mere label dependence, i.e., a difference in verdict that does not correspond to any relevant difference, looks arbitrary once one has already accepted that temporal location is not, by itself, decisive.

This is the same rational pattern we use in condemning inconsistent self-treatment across time. If "it is now Tuesday rather than Monday" is not, by itself, a reason to reverse my verdict in a normatively identical case, there is no non-arbitrary reason to insist that "it is me rather than you" is, by itself, such a reason. Both are transformations that preserve the features I claim to treat as relevant.[9]

[9]One of the most ambitious prior attempts to derive impartiality from structural features of agency is Gewirth's *Reason and Morality* (1978). Gewirth's argument moves from a first-personal claim ("I have rights to the generic conditions of agency") to a universal claim via a universalization step widely regarded as involving a scope fallacy (see MacIntyre 1981, Ch. 6): the agent can consistently claim rights for themselves without conceding that all agents have such rights. The present argument avoids this difficulty because its logical structure is reversed. It does not begin from an indexed first-personal claim that must be universalized; it begins from an unindexed constraint on evaluation (OI) and shows that bare-indexical privilege is inadmissible by the constraint's own standards. The scope-fallacy objection does not arise because the

2.2.1 Structural identity, bare indexicality, and agent-relative reasons

The previous claim invites an important objection: personal identity is fundamentally different from temporal location because "it is my pain" picks out a relation (undergoing, suffering, living-through) among the facts the evaluator takes to matter. This objection draws on a tradition with considerable force. Williams (1986) argues that an agent's ground projects constitute practical identity and are themselves sources of reasons. Scheffler (1982) develops a related thought: agents are rationally permitted to give disproportionate weight to their own projects, reflecting the normative significance of the standpoint of agency itself.[10]

The symmetry framework can take these concerns seriously without abandoning person-permutation invariance. For any agent i in a case w, we can distinguish: (1) *Structural identity*: the full constellation of relations, roles, projects, commitments, experiential states, and histories that i occupies in w; and (2) *Bare index*: the residual fact that this structural profile is instantiated by i rather than j, once all features in (1) have been specified. The claim is that structural features do the normative work; the bare index is the address at which those features reside.

A sophisticated egoist will resist this decomposition. They will insist that the fact that this pain is happening *to me*, i.e., that I am the subject of this experience, is itself a structural feature, not a bare index; that it is a genuine metaphysical relation of subjecthood that is an irreducible source of practical reasons, not a mere label. Three convergent lines of argument show why this resistance is untenable.

2.2.1.1 The parity argument. The egoist's move proves too much. If "the fact that this pain is happening to me" can be promoted from bare index to structural feature simply by insisting on its experiential or metaphysical significance, then by parity, "the fact that this pain is happening now" can be promoted in the same way. A diachronically capricious agent, call them the *wanton*, can insist with equal conviction: "The fact that this suffering is present, that I am living through it at this moment, is not a mere temporal label; it is a genuine experiential relation. Present pain has an immediacy, a phenomenological vividness, that past and future pain lack. This is a structural feature of my experience, not a representational artifact."

argument never requires the transition from "my rights" to "everyone's rights," only from "invariance under time-shift" to "invariance under person-permutation." A secondary advantage is that the argument does not depend on rights-talk, which critics have argued imports normative content not derivable from the structure of agency alone.

[10] A more recent and in some respects more precise version of this challenge appears in Kolodny (2003), who argues that relationships generate reasons grounded in the distinctive value of a shared history between *these particular* individuals, not merely between any individuals standing in the relevant structural relations. The symmetry framework's response is the same in structure: what makes a shared history normatively significant is its content, i.e., the character of the experiences shared, the commitments formed, the mutual vulnerability undertaken, all of which are functionally characterizable and preserved by person-permutation. What permutation alters is only which individuals instantiate that shared history, not the history's evaluative character. Kolodny's account is therefore best understood not as a counterexample to person-permutation invariance but as a rich description of the structural content that invariance preserves. The challenge would be genuine only if the value of a shared history were irreducibly dependent on the bare numerical identity of the participants, and that would face the same difficulties as the bare-indexical claims examined in the main text.

The wanton's claim has considerable phenomenological plausibility: present experience *does* have a distinctive character that memory and anticipation lack. Yet we recognize (as must the egoist who accepts diachronic coherence) that this phenomenological asymmetry does not generate a normative asymmetry. The vividness of present experience is a psychological fact about how temporal location affects salience, not a rational ground for treating normatively identical cases differently across time. If it were, diachronic coherence would be undermined: the agent could always say "but this is happening *now*," and the constraint would have no purchase.

The egoist is in a precisely parallel position. They point to a genuine phenomenological asymmetry, i.e., first-personal experience has a character that third-personal observation lacks, and attempt to transmute it into a normative asymmetry. But the bootstrapping argument asks: once all the structural features that make my pain normatively significant (its phenomenology, its severity, its relation to my projects and capacities) have been specified, what further normative work does the bare fact that this structural profile is instantiated by *me* do? The egoist must identify such work; the parity argument shows that the most natural way of identifying it (by pointing to the distinctive experiential character of first-personal involvement) equally supports a conclusion the egoist rejects.

A natural resistance at this point appeals to a metaphysical disanalogy: distinct persons are metaphysically separate subjects of experience in a way that distinct temporal stages of a single person are not. The thought is that the unity of a persisting subject gives temporal indices a different normative status than personal indices, and that this difference breaks the parity on which the bootstrapping step depends.

This objection can be met on its own terms. Grant, for the sake of argument, that persons are genuinely metaphysically separate in a way that temporal stages of a single person are not. The question is whether this metaphysical fact does the normative work the partialist needs it to do. The separateness of persons is a structural feature of the situation: it is part of the description of what the moral landscape looks like. And structural features are precisely what person-permutation preserves. After permutation, persons are still metaphysically separate; the boundary between subjects is still real; the full phenomenological and metaphysical character of distinct subjecthood is intact. What changes is only which subject occupies which position in the situation, and that is the bare index. The partialist's mistake is to conflate two claims: the claim that persons are genuinely separate (a structural fact, preserved by permutation) and the claim that the fact that *I* am *this* person is normatively privileged (a bare-indexical claim, not preserved by permutation). The metaphysical separateness of persons, far from blocking the bootstrapping step, is accommodated by it. The separateness is real, but it is a feature of the structure (the kind of world we inhabit, with its plurality of subjects) not a feature that picks out any particular subject as normatively special.

A further difficulty for the partialist who grounds the normative asymmetry between temporal and personal indices in the unity of a persisting subject is that the unity relation itself comes in degrees and varies across cases in ways that do not track the partialist's normative conclusions. The psychological connectedness between me-now and me-five-minutes-ago is far stronger than the connectedness between me-now and me-at-age-three; the unity of consciousness is regularly interrupted by sleep, anaesthesia, and the ordinary discontinuities of a human life; and the phenomenological sense of being "the same person" can be profoundly

disrupted by trauma, conversion, or the sheer passage of decades. If the normative significance of "mine" is grounded in this unity relation, then the strength of self-regarding reasons should co-vary with the strength of the relation. An agent should have markedly weaker prudential reasons concerning their distant future self than their near future self, roughly in proportion to the degree of psychological connectedness. But partialists typically do not accept this implication. They do not hold that my reasons to save for retirement are dramatically weaker than my reasons to eat lunch, nor that an agent's prudential obligations to a radically transformed future self are negligible. The normative conclusions the partialist defends remain stable across cases in which the unity relation varies enormously.

This co-variance failure is itself diagnosable within the symmetry framework. The partialist invokes the unity relation to block the bootstrapping step from temporal to personal invariance: it is the metaphysical unity of a single life, they claim, that makes temporal indices normatively different from personal indices. But they then decline to let that relation do the work it would need to do if it were genuinely the normative ground. They treat the unity relation as load-bearing at the boundary between persons (where it blocks person-permutation invariance) but not load-bearing within persons (where it would generate uncomfortable implications about diminishing self-regarding reasons over time). This selective deployment is precisely the kind of evaluative inconsistency that OI is designed to identify: a feature is treated as normatively relevant when it serves the partialist's preferred conclusion and set aside when it does not. The partialist's own normative commitments come apart from the metaphysical fact they invoke, which strongly suggests that the unity relation is functioning as a *post hoc* rationalization for a conclusion arrived at on other grounds rather than as the genuine source of the normative asymmetry between persons and times.

It should be acknowledged that this is the point at which the bootstrapping exerts its most intense pressure and at which the argument's conditional character is most salient. A partialist who is prepared to insist, at this juncture, that the bare personal index is normatively irreducible, i.e., that being this subject is not a structural feature but a primitive normative fact, has found the precise location at which to resist. The costs of that resistance, and the full landscape of options available to the partialist at this juncture, are mapped in the clarificatory note at the end of Section 2.

2.2.1.2 The functional characterizability criterion argument. A second line of defense offers a principled criterion for distinguishing structural features from bare indices, one that does not simply encode the desired conclusion. The criterion is *functional characterizability*: a feature is structural if and only if it is specifiable in terms of roles, relations, and qualitative properties without residual reference to a particular index.

Consider what makes my pain normatively significant: its phenomenological character (what it is like to undergo it), its severity (how much it disrupts functioning and well-being), its relation to my projects (whether it threatens commitments I have reason to sustain), my capacity to bear it (given my history, resilience, and resources). All of these are functionally characterizable. They describe a role and a situation: any agent who occupied this role and found themselves in this situation would instantiate the same features. The bare index "me" is precisely what remains once all functional characterization has been given. It is the answer

to the question "who fills this role?" not to the question "what is the role?".

This criterion is not an independent premise of the argument but an inference from OI's non-vacuity. The inferential structure is as follows. OI imposes a non-trivial constraint: objective evaluation must be invariant under transformations that preserve all normatively relevant structure. For this constraint to have critical force, the category of "normatively relevant structure" must itself be principled, i.e., it cannot be open to unrestricted stipulation by the agent whose evaluation is under scrutiny. If any agent could simply declare their own index to be part of the relevant structure, OI would be vacuously satisfied: every evaluation would count as objective, since no transformation that altered the agent's favored index would count as preserving "all normatively relevant structure." The constraint would lose all capacity to diagnose double standards, ad hoc exceptions, or capricious reversals. Functional characterizability is what prevents this collapse. The point bears emphasis: the philosophical burden here is the reverse of what it might initially appear to be. Functional characterizability is not an aggressive stipulation that indexical properties are normatively inert. It is the *minimal* condition required to prevent OI from collapsing into vacuity. The partialist who rejects it does not thereby adopt a more permissive or more modest position; they adopt a position on which the very constraint they invoke when criticizing others' evaluations as arbitrary or ad hoc (the constraint that evaluation should track what matters rather than what is idiosyncratic) loses all critical force. The controversial position is not functional characterizability but its denial.

By requiring that normative relevance claims be articulable in terms of roles, relations, and qualitative properties without residual indexical reference, the criterion preserves OI's ability to distinguish genuine normative sensitivity from mere positional privilege. It is therefore not a stipulation that bare indices are normatively inert, nor a discovery about the metaphysical status of indexical properties, but an inference from what OI must demand in order to function as the rational constraint it purports to be. The criterion does exclude bare-indexical reasons from the category of structural features, but it does so in a way that is uniform across all indices (temporal, personal, and perspectival) and that is motivated at each point by the same consideration: without it, the objectivity constraint that both the impartialist and the partialist invoke when criticizing arbitrary evaluations would have no purchase. The partialist who resists functional characterizability is therefore not quarreling with an auxiliary premise; they are quarreling with OI itself, or more precisely, with the conditions under which OI is non-vacuous. That is a coherent position, but its costs should be recognized: it amounts to accepting that any agent can immunize any evaluation from objectivity constraints simply by insisting that their own standpoint is part of the normatively relevant structure.

It is worth making the dialectical burden fully explicit. The partialist who resists functional characterizability does not merely need to assert that certain indexical relations (for instance, evaluative authority over one's own future) are structural rather than bare-indexical. They need a criterion for which indexical relations count as structural, one that meets three constraints simultaneously. First, the criterion must be *non-question-begging*: it cannot simply enumerate the indexical relations the partialist antecedently wishes to privilege. Second, it must be *non-vacuity-inducing*: it cannot be so permissive that any agent can immunize any evaluation from objectivity criticism by declaring their favored index part of the relevant structure, for that would collapse OI into triviality. Third, it must be *independently*

motivated: it must be defensible on grounds other than the desire to preserve the partialist conclusion. The difficulty is that candidate criteria tend to fail at least one of these constraints. A criterion that admits "the relation of being the subject who will undergo this experience" as structural, for instance, must explain why it does not equally admit "the relation of being the agent who occupies this spatial location" or "the relation of being the person who belongs to this group," relations whose promotion to structural status would generate conclusions the partialist does not accept. If the criterion excludes these by stipulation, it is question-begging. If it admits them, it is vacuity-inducing. If it draws the line by appeal to the phenomenological distinctiveness of first-personal experience, then the parity argument of Section 2.2.1.1 applies: temporal immediacy is phenomenologically distinctive in precisely the same way, yet the partialist who accepts diachronic coherence does not treat it as structurally privileged. The absence of a criterion meeting all three constraints is not merely a gap in the partialist's argument; it is evidence that the decomposition into structural features and bare indices offered by functional characterizability tracks a genuine rational distinction.

This criterion directly addresses the egoist's resistance. When they insist that "being the subject of this experience" is a structural feature, they are pointing to something real, but the something real is a functional role: *the one who undergoes the suffering*. This role is indeed part of the structural profile, and it is preserved by person-permutation. After permutation, the occupant of the role still undergoes the suffering, still experiences it first-personally, still bears its phenomenological character. What changes is only which particular individual fills the role, and that is the bare index.

The criterion also explains why the decomposition is not ad hoc. It applies uniformly: to temporal indices (the functional role of "the present moment in deliberation" is structural; the bare fact that the present moment is this particular Tuesday is not), to personal indices (the functional role of "the agent who must choose" is structural; the bare fact that the agent is this particular person is not), and to perspectival indices (the functional role of "the standpoint from which the evaluation is made" is structural; the bare fact that it is this particular standpoint is not). The decomposition tracks a single underlying distinction across all three cases.

2.2.1.3 The explanatory idleness argument. A third line of defense presses a question the egoist has difficulty answering: what normative work does the bare index do that the structural profile does not already do?

Suppose we have a complete structural description of a case of suffering: the phenomenological character of the pain, its intensity and duration, the sufferer's history and projects, the relation of the suffering to those projects, the sufferer's capacity to bear it, the availability of alternatives, and so on. Call this the *complete structural profile*. Now ask: if we learn, in addition, that the sufferer is *me* rather than *you* (while holding the entire structural profile fixed), does this additional piece of information change what is normatively required? Does it change which responses are rationally appropriate, which duties obtain, or which trade-offs are permissible?

The egoist says yes: the fact that it is me who suffers generates additional reasons, reasons that would not obtain if the same structural profile were instantiated by someone else. But this claim faces a dilemma. Either the egoist can articulate what normative difference the

bare index makes, in which case the difference is characterizable in functional terms and is therefore part of the structural profile after all, or the egoist cannot articulate the difference, in which case the bare index is doing no work that can be rationally defended. In the first case, the egoist has not identified a genuine bare-indexical reason; they have identified a structural feature that was omitted from the profile, which can be added without breaking person-permutation invariance. In the second case, the bare index is explanatorily idle: it is posited as normatively relevant but cannot be shown to make a normatively articulable difference.

Allowing explanatorily idle features to shift evaluations is a form of the very caprice that OI is designed to rule out. To accept that a feature is normatively relevant while being unable to articulate what normative difference it makes is to allow one's evaluation to track something one cannot defend as reason-giving. This is precisely the contrast between reasoning and caprice that motivates OI in the first place.

2.2.1.4 The convergence of the three lines. These three arguments converge on the same conclusion from different directions. The parity argument shows that the egoist's strategy for promoting the bare index to structural status, if successful, would equally undermine diachronic coherence (a constraint the egoist aspiring to objectivity by accepting diachronic coherence accepts). The functional characterizability criterion argument closes the partialist's most promising escape route (rejecting the criterion while replacing it with an alternative that admits the bare personal index without collapsing OI into vacuity) by showing that any such alternative must be non-question-begging, non-vacuity-inducing, and independently motivated, constraints that candidate criteria have not been shown to meet. The explanatory idleness argument shows that the bare index, even if posited as normatively relevant, cannot be shown to do normative work that the structural profile does not already do, and that allowing such idle features to shift evaluations is incompatible with the rational discipline that OI articulates.[11]

The upshot is that the most defensible forms of agent-relativity in the Williams–Scheffler tradition are not counterexamples to person-permutation invariance but illustrations of what it already accommodates. What the framework rules out is not the normative significance of personal identity. My projects, my relationships, and my experiential situation all matter, but the claim that the bare fact of being *me*, once all such structural features have been specified, is itself an independent and irreducible source of practical reasons. Identity matters in virtue of its structural content. The bare index, stripped of that content, cannot do the normative work that the egoist would have it do.

[11] A related objection holds that deliberation is constitutively first-personal in a way that temporal location is not: when I deliberate, I exercise practical authority that is mine in a manner irreducible to the occupancy of a functional role. But this objection is addressed by the convergence of the arguments already given. The parity argument applies directly: deliberation is constitutively present-tensed in precisely the same way (I deliberate now about what to do next), yet the agent who accepts diachronic coherence has already conceded that this constitutive temporal presence does not generate normative privilege. And the functional characterizability criterion identifies the constitutive first-personality of deliberation as a universal structural feature of agency (every agent deliberates first-personally), which is precisely what person-permutation preserves.

2.3 Step three: perspective invariance

The second step of the argument yields impartiality under person-permutation. But practical reasoning faces another common threat to objectivity: sensitivity to mere differences in perspective or framing. Consider an agent who endorses a policy when described abstractly ("it increases efficiency") but rejects it when described concretely ("it will predictably burden this vulnerable subgroup"), despite granting that the two descriptions refer to the same policy with the same effects. Sometimes this reflects learning. But sometimes the agent acknowledges that the descriptions pick out the same relevant facts and nonetheless treats the shift in salience as reason-giving. In those cases, the evaluation tracks psychological contingencies rather than normatively relevant structure.

Perspective invariance follows from OI without controversial premises. If I am committed to giving reasons that are not merely self-advertisements of my current attitudes, I must distinguish "my judgment changed because I learned something relevant" from "my judgment changed because the same facts were presented differently." The latter is again caprice relative to one's own standards.

2.4 From global impartiality to local invariance

The bootstrapping argument of the preceding sections is now complete in its essentials. Beginning from the minimal invariance of diachronic coherence (Section 2.1), the argument has shown that symmetry requirements extend, without arbitrary stopping point, through person-permutation invariance (Section 2.2) and perspective invariance (Section 2.3). These three invariances jointly yield a robust form of impartiality: objective practical evaluation cannot depend on mere differences of temporal location, agent identity, or perspective when those differences preserve all normatively relevant structure. Section 3 will draw the normative consequence of these results for the relationship between moral and prudential reasons.

Before turning to that consequence, however, it is worth showing that the invariance framework does not exhaust itself at the global level. Evaluations satisfying all three global invariances may still vary in their action-guiding structure because of differences in agents' local evaluative calibrations and representational 'frames.' The discussion that follows is explicitly programmatic. Its aim is not to carry the same argumentative weight as the bootstrapping steps but to demonstrate that the objectivity-as-invariance framework generates a structured research program extending beyond the global symmetries that constitute the paper's central argument.

2.4.1 The scaling case

Two agents, A and B, both accept that suffering is bad and that person-permutation invariance holds. They face a resource-allocation problem involving patients with moderate and severe suffering. Agent A's deliberative economy registers these at weights of 3 and 10; Agent B registers them at 30 and 100. Both agree on all first-order facts.

If both agents agree on the ordering of reasons, the normative significance ratios, and the pattern of admissible trade-offs, then the tenfold difference in absolute magnitude is a calibration artifact, analogous to the difference between Celsius and Fahrenheit. The dynamical

content is the same. But suppose B's higher absolute weights generate a structural consequence that A's do not: a lexical priority of severe over moderate suffering that A rejects. Here, the scaling difference is no longer merely representational but reflects a genuine normative disagreement. Local invariance is not violated; the requirement simply does not apply, because the transformation from A's frame to B's is not a representational recalibration but a substantive divergence.

The structural test is this: if two evaluative framings yield convergent verdicts and convergent rational pressures across the full range of cases, residual differences are representational, and local invariance requires that objective evaluation not depend on which framing is employed. If they yield divergent verdicts, the divergence is substantive and local invariance does not demand equivalence.

2.4.2 Cross-cultural evaluative vocabularies and the general requirement

The above scaling case involves numerical calibration. A richer class of cases involves evaluative vocabularies: the conceptual frameworks through which agents organize practical reasoning. Different ethical traditions (e.g., a Confucian vocabulary organized around role-constituted obligations and a Kantian vocabulary organized around duties and respect for persons) may differ substantially in conceptual structure while yielding convergent practical verdicts across a wide range of cases. Where such convergence obtains under conditions of internal coherence, informational adequacy, and competent deployment of each tradition's evaluative resources, the residual differences are representational: the two vocabularies are different gauges for encoding the same normative content. Where convergence breaks down, the framework identifies a substantive normative disagreement and locates it precisely as a disagreement about which features of the situation are normatively relevant.

These cases motivate the local invariance requirement in its general form. 'Local' means that the transformations may vary from agent to agent and context to context: unlike person-permutation, local invariance concerns transformations of an agent's internal evaluative frame that may differ across agents. The requirement is that objective evaluation be invariant under such transformations when they are representational, i.e., when they preserve the structure of rational pressure, the ordering of reasons, and the pattern of admissible trade-offs.

The analogy with gauge theory in physics is instructive. In physics, the distinction between gauge degrees of freedom and physical degrees of freedom is not always obvious in advance; identifying the true physical content of a gauge theory is itself a substantive theoretical achievement. Similarly, determining whether two rich evaluative traditions converge under idealized reflection is a substantial empirical-philosophical undertaking. The local invariance requirement does not answer every such question in advance, but it ensures that the question is asked precisely and that the justificatory burden is clearly located. For any cross-traditional disagreement, the framework poses a structured question: locate the normatively relevant difference that explains the divergence, or acknowledge that one or both evaluations depend on a representational artifact. The difficulty of drawing this distinction in hard cases is not a deficiency of the framework but a reflection of the complexity of the normative terrain it maps.

The strongest form of the local invariance requirement demands stability not only of

evaluative outcomes across representational recalibrations but of the normative structure governing deliberation: which considerations count as reasons, how they constrain choice, and which trade-offs are admissible. This is the endpoint of the symmetry-bootstrapping trajectory: objectivity, understood as invariance under normatively irrelevant transformations, naturally extends from global symmetries of time and person to a local invariance constraint governing the action-guiding structure of practical reasoning itself.

A clarification about the character of the foregoing argument is in order. The bootstrapping argument does not claim to refute every possible form of partiality. What it does is map the dialectical terrain with precision, revealing that a partialist who accepts diachronic coherence faces three options.

The first is to accept that the bare personal index is normatively irrelevant, conceding person-permutation invariance (and the structural priority of moral reasons that follows from it, as I argue in the next section).

The second is to insist that being this particular subject is not a structural feature but a primitive source of practical reasons. This position is technically coherent, but the arguments of Section 2.2.1 reveal its costs to be high. The parity argument shows that the same strategy would equally license the wanton's rejection of diachronic coherence. The position requires positing an explanatorily idle normative primitive, and pushed to its limit, it amounts to accepting that any agent can immunize any evaluation from objectivity constraints by declaring their own standpoint part of the normatively relevant structure (which is to accept OI in letter while emptying it of critical force). At that point, the partialist has not exited the argument's scope but must hold that OI can do no diagnostic work, a consequence whose costs are borne not only in ethics but in every other domain where the partialist may wish to invoke objectivity to criticize arbitrary evaluations.

The third is to retreat to the wanton's position and reject the aspiration to practical objectivity that the argument's conditional structure presupposes. This option is coherent (the argument does not claim otherwise), but it is not a counterexample to the bootstrapping argument; it is a confirmation of its conditional character. The only way to block the extension from temporal to personal invariance without incurring the costs of the second option is to exit the argument's scope entirely.[12]

The upshot is a shift in burden. The argument does not demonstrate the impossibility of principled partiality; it shows that resisting the bootstrapping pressure requires a specific kind of justification, and it makes the costs of the available strategies explicit. This is a genuine advance over familiar routes to impartiality that either build substantial impartiality assumptions into the setup, or lean primarily on case-driven intuition.

[12] A related sub-option is to accept only a graded form of diachronic coherence, treating temporal proximity as part of the normatively relevant structure rather than as a bare temporal index. This concedes less than the full retreat to the wanton's stance, but it is an instance of precisely the promotion of a bare index to structural status that OI diagnoses as caprice. This sub-option thus concedes, by a different route, what the wanton's retreat concedes directly: the aspiration to practical objectivity that the argument presupposes. In either case, the partialist exits the scope of the bootstrapping argument entirely.

3 The Normative Upshot: Structural Priority and the Conditionality of Prudential Reasons

If objective evaluation is constrained by invariance under the relevant transformations, then moral considerations have structural priority over merely prudential or partial considerations. Prudential reasons can matter, but only conditionally, i.e., only insofar as they can be integrated into an evaluation that respects the invariances required by objectivity.

I am not claiming that prudence is always irrational, that agents must be indifferent to their own welfare, or that morality always overrides every other consideration regardless of content. The claim is narrower: once we adopt OI, we can distinguish between (i) considerations representable as part of an objective evaluation and (ii) considerations that cannot be, because they depend essentially on a symmetry-breaking privilege of time, identity, or perspective. The former have standing as reasons within objective practical deliberation; the latter do not.

3.1 Integrable and non-integrable prudence

1. *Integrable prudence*: an agent takes its own welfare into account in a way that could be incorporated into an impartial evaluation, i.e., treating its welfare as one instance of a type of consideration that would count similarly for any agent in relevantly similar circumstances.[13]

2. *Non-integrable prudence* (privileging): an agent treats its own welfare as counting more simply because it is its own, holding fixed all relevant structure.

The symmetry-bootstrapping argument implies that only the first kind has standing within objective practical reasoning. The second is disallowed not because it is "immoral" by stipulation, but because it is arbitrary relative to the standards that define objectivity.

3.2 More than coherence, less than a full first-order theory

One might worry that this is merely a dressed-up coherence requirement. But coherence is compatible with systematically self-favoring outlooks so long as they are stable. What

[13]The integrable/non-integrable distinction requires a test to prevent a form of bad faith: an agent who redescribes bare self-privileging in structural language ("I am devoting disproportionate resources to the agent best positioned to execute these particular projects") without genuinely tracking the structural features they invoke. The test is counterfactual robustness under the same person-permutation logic that generates the distinction. Hold fixed all structural features of the case and permute the agent into a different position (with a different agent's projects, history, and relationships, while someone else now occupies the agent's original position with its structural profile). If the agent's evaluative verdicts track the structural features, if they judge that the agent now occupying the structurally favored position (whoever that turns out to be) merits the disproportionate resources, then the self-concern is genuinely integrable: it depends on the structural profile, not on the bare index. But if the agent's resource allocation consistently tracks their own position across permutations, regardless of the structural profile attached to it, then the structural redescription is *post hoc* rationalization of bare self-privileging, and the evaluation fails the invariance test. This counterfactual robustness test is not a supplement to the person-permutation analysis of Section 2.2 but an application of it: the double-standard test already deployed there, now made explicit as a diagnostic for the integrable/non-integrable boundary.

the invariance framework adds is a principled test for arbitrariness: it identifies distinctions (mere differences of time index, label, or perspective) that do not track normatively relevant structure. An outlook that depends essentially on those distinctions is not merely coherent or incoherent; it is non-objective.

The invariance constraints do not tell us which actions are right or which outcomes are best. They constrain the form that objective evaluation can take and thereby rule out certain patterns of justification, in particular, any purported reason whose force depends essentially on a symmetry-breaking privilege.

3.3 Lexical priority: a qualified but principled sense

Invariance constraints generate a hierarchy among candidate considerations.[14] Considerations invariant under minimal objectivity requirements (temporal coherence) are admissible in any rational deliberation aspiring to be reason-guided across time. Those invariant under person-permutation are admissible in deliberation aspiring to impartial objectivity. Those stable under local invariance tests are admissible even relative to variation in evaluative framing. Considerations that fail at any stage are not merely outweighed; they are excluded as reasons within the corresponding notion of objective deliberation.

A natural worry is that structural priority is excessively demanding.[15] The framework's answer is no. An agent's projects, relationships, and personal commitments are part of the normatively relevant structure, which is preserved, not erased, by the invariance tests. What structural priority rules out is not self-concern but bare self-privileging: the treatment of one's own welfare as counting more simply in virtue of being one's own, once all structural features have been accounted for. An agent who devotes significant energy to personal projects does not thereby violate the invariance constraints, so long as that devotion can be represented as an instance of a kind of concern that would count similarly for any agent in a relevantly similar structural position.

3.4 What the argument rules out

The symmetry-bootstrapping argument rules out an outlook on which there are objective reasons that fundamentally privilege a particular agent's standpoint, for instance, that "my pain matters more because it is mine" is an objective reason. If "mine" functions merely as a label, privileging it violates OI as soon as we accept that "now" cannot function merely as a label in diachronic reasoning. To insist on one but not the other is to adopt an arbitrary asymmetry.

[14] I understand 'priority' here as a structural constraint on what counts as an objective reason: prudential considerations that depend essentially on symmetry-breaking privilege fail the invariance tests and so do not qualify as objective reasons. I remain neutral on whether such considerations may still have motivational pull or constitute reasons in some other sense.

[15] The worry is forcefully articulated in Wolf (1982), who argues that a life fully governed by moral demands would lack the space for personal projects, deep relationships, and individual excellence that make a human life worth living.

4 Symmetry Tools and Diagnostic Contrasts

The preceding sections advanced the philosophical claim without technical machinery. Formal tools from the theory of symmetry are nonetheless useful at a programmatic level. This is not because it somehow "reduces" ethics to physics but because the symmetry framework provides a disciplined vocabulary for making objectivity constraints explicit, diagnosing where competing views differ, and generating structured questions for future work.[16]

Two ideas from physics are especially illuminating. First, in mature sciences, symmetry often functions not as a property discovered after the fact but as a constraint shaping admissible theory form, a sieve that narrows the space of candidate dynamics. The methodological lesson for ethics: invariances can regiment objectivity constraints as constraints on admissible forms of evaluation. Second, Noether's theorems make vivid how different kinds of symmetry requirements yield different downstream consequences.[17] Global invariances (transformations applied uniformly) are associated with stable, shareable constraints, i.e., features that remain fixed across transformations and serve as common currency of objectivity. Local invariances (transformations that may vary point to point) demand additional structure that tracks how locally varying descriptions are related. In ethics, this suggests that global invariances generate shared evaluative constraints, while local invariances demand the additional structure needed to keep evaluative judgments coherent across agents whose deliberative frames may vary. (The appendix makes this formal.)

4.1 What formal tools contribute: clarity, diagnostics, and extension

Explicitness. Philosophers often invoke "the objective standpoint" or "impartiality" without fully specifying which transformations are treated as normatively irrelevant. A symmetry framework forces us to state the invariances explicitly. The result is not increased metaphysical ambition but decreased dialectical ambiguity. Some substantive normative disagreements can be fruitfully understood as disagreements about which transformations count as irrelevant.

Diagnostics. Once invariances are explicit, the framework provides a simple way to diagnose disputes. Two theorists who both endorse "impartiality" may accept different symmetry groups and so be committed to different conceptions of objectivity. The bootstrapping argument can then be seen as an attempt to answer precisely the question: what justifies stopping at this symmetry rather than proceeding to that one?

Disciplined extension. The framework suggests a natural research program: catalogue candidate transformations that might be normatively irrelevant; argue for or against their irrelevance; derive constraints on admissible evaluations from the resulting invariances; and identify defensible forms of symmetry breaking. Just as symmetry breaking in physics can be

[16]For an illuminating philosophical discussion of the methodological role of symmetry in physics, see Brading & Castellani (2006). For a classic physicist's statement, see Wigner (1979a, 1979b).

[17]Noether's 1918 paper proves two foundational results: continuous global symmetries of an action yield conservation laws, while local symmetries generate differential identities that motivate the modern gauge perspective. See Noether (1918; English trans. Tavel). For a pedagogical exposition, see Neuenschwander (2017).

controlled and explanatory, departures from strict invariance in ethics (special obligations, role responsibilities, epistemic asymmetries) may be justified, but they should be located precisely and defended as such.

4.2 A diagnostic contrast with Scanlon's contractualism

It is worth briefly noting how the symmetry framework relates to an alternative route to impartiality. Scanlon's contractualism derives moral constraints from a test of reasonable rejectability: an action is wrong if it would be disallowed by principles that no one could reasonably reject (1998). The two approaches overlap in their conclusions but differ in derivational structure. Scanlon's test is grounded in a substantive ideal of interpersonal justification; the symmetry framework derives impartiality from a thinner structural source: the non-arbitrariness requirements internal to objective evaluation as such.

A brief example shows that this diagnostic difference is not merely verbal. Consider a principle governing aggregative trade-offs:

> **(Aggregate):** When redirecting a harm would save five lives at the cost of one, it is permissible to redirect.

Both frameworks generate grounds for scrutinizing this principle, but they locate the source of difficulty in different places. For Scanlon, the problem is interpersonal: the one person who would bear the redirected harm has a claim that cannot be adequately answered by pointing to the aggregate benefit enjoyed by the five. The principle fails because it cannot be justified to that individual on grounds they could reasonably accept.

(Aggregate) passes all three global invariance tests: it is temporally invariant, person-permutation invariant, and invariant under global perspective shifts. The question the symmetry framework raises is whether the principle survives local invariance, specifically, whether its verdict depends on a particular evaluative framing (an aggregative calculus in which five units of welfare straightforwardly outweigh one) that may not be the uniquely correct representation of the normative structure. An agent employing a rights-based framing (in which each person's claim against being harmed is not directly commensurable with aggregate welfare) may reach a different verdict while agreeing on all first-order facts. If this disagreement is substantive, local invariance does not apply and the divergence is real. But if, under idealized reflection, the divergence is traceable to a calibration difference rather than a difference in which considerations the agents take to bear, then (Aggregate)'s dependence on the aggregative framing is a form of evaluative partiality: it privileges one representational scheme without establishing that the scheme tracks a genuine structural distinction.

The two frameworks thus converge on the conclusion that naive aggregation faces serious objections, but they identify different structural sources. Scanlon's diagnosis is irreducibly relational: the principle fails because of what is owed to a particular person. The symmetry framework's diagnosis is representational: the principle's verdict may depend on a choice of evaluative frame whose status as uniquely tracking the normative structure has not been established. These are complementary insights, and cases where they come apart, i.e., where a principle is justifiable to each individual but depends on a contested evaluative calibration, or is locally invariant but reasonably rejectable, mark genuinely different regions of normative space that neither framework, on its own, fully maps.

4.3 A diagnostic contrast with Rawls's original position

Rawls's original position provides the most celebrated device for generating impartiality in moral and political philosophy. Parties behind the veil of ignorance choose principles of justice without knowledge of their particular position in society: their talents, social class, conception of the good, or generation. The veil thereby ensures that the chosen principles are not tailored to anyone's particular advantage. It is natural to ask how the symmetry framework relates to this construction.

The relationship is illuminating in both directions. From the standpoint of the present framework, the veil of ignorance is best understood as a device for enforcing person-permutation invariance. By stripping away knowledge of which position one occupies, the veil ensures that the chosen principles do not depend on the bare index of the chooser: any party behind the veil, regardless of which actual person they turn out to be, faces the same deliberative situation and is subject to the same rational pressures. The veil does not merely illustrate person-permutation invariance; it *implements* it, by removing the information that would allow evaluations to track bare identity rather than structural features.

But there is a crucial difference in derivational structure. Rawls introduces the original position as a representation of a prior commitment to fairness. The veil's design (what information it excludes, what it permits) is justified by appeal to a conception of persons as free and equal and to the role that principles of justice are meant to play in a well-ordered society. The original position is a construction built to express antecedently accepted moral commitments. It does not derive those commitments from thinner premises; it presupposes them.

The symmetry-bootstrapping argument, by contrast, derives the need for something like the veil from thinner premises. It begins not from a substantive conception of fairness or of persons as free and equal, but from two structural commitments: (i) that practical objectivity is a genuine aspiration, cashed out as invariance under normatively irrelevant transformations (OI), and (ii) that diachronic coherence is the minimal instance of such invariance. From these premises alone, the bootstrapping argument generates person-permutation invariance as a rational requirement, not because we have antecedently accepted that persons are free and equal, but because refusing to extend the logic of invariance from temporal indices to personal indices is, by the standards of OI, arbitrary.

This difference matters for the dialectical situation. Rawlsian constructivism is powerful but faces a familiar challenge: the design of the original position appears to build in the moral conclusions it is meant to justify. Critics have pressed this point in various ways: the choice of what information to exclude behind the veil already reflects a substantive moral judgment about what is and is not relevant to justice. The symmetry framework offers a response to this structural worry, not by defending Rawls's particular construction, but by providing an independent argument for why something with the structure of the veil is rationally required. If the bootstrapping argument succeeds, the veil of ignorance is not an arbitrary or question-begging device; it is a natural implementation of a constraint that any agent aspiring to objectivity in practical reasoning has reason to accept.

5 Conclusion

The argument has moved from a thin bridge principle to a substantive normative consequence. Objectivity-as-invariance articulates what is presupposed when we criticize double standards and demand consistency: if a transformation preserves everything an evaluator takes to be normatively relevant, a shift in verdict is arbitrary. Diachronic coherence supplies the initial foothold, and from there, bootstrapping pressure extends through person-permutation, perspective invariance, and local invariance to yield a principled structural priority of moral reasons over merely prudential considerations. As the dialectical mapping at the end of Section 2 makes explicit, the partialist who resists the argument must either empty OI of the critical force that underwrites objectivity claims across all domains (including the partialist's own capacity to criticize evaluations as arbitrary) or abandon the aspiration to practical objectivity altogether. Both routes are ones that most participants in the debate are already foreclosed from taking by their own critical practices: anyone who has ever condemned a double standard, demanded consistency, or dismissed an evaluation as ad hoc has implicitly endorsed the aspiration and the constraint that the bootstrapping argument exploits.

Several avenues for future work are especially salient. First, the symmetry framework invites a more systematic taxonomy of candidate transformations in practical reasoning (temporal shifts, agent permutations, perspective shifts, and variations in evaluative calibration) and of the justificatory burden associated with treating them as normatively irrelevant. Second, it invites a principled account of justified symmetry breaking in ethics. Many plausible moral considerations, e.g., special obligations, role responsibilities, epistemic asymmetries, and the moral relevance of relationships, appear to introduce agent-relativity. The symmetry lens offers a way to separate cases where such agent-relativity reflects genuine relevant structure from cases where it collapses into mere label dependence. Third, the framework suggests a promising point of contact with the Sidgwickian "dualism of practical reason." One natural reading of the dualism treats prudence and morality as competing sources of reasons without an internal rational ordering. The symmetry-bootstrapping argument provides a structural rationale for resisting that picture: prudential considerations have rational standing only insofar as they can be represented within an objective evaluation that respects the invariances constitutive of practical objectivity. In that sense, prudential reasons are conditional: admissible as enrichments of, but not competitors to, the reasons that survive impartiality constraints. This bears a structural resemblance to recent work (notably, Ralf Bader's (2015) discussion of 'silencing' in the context of this dualism) that aims to explain how the claims of prudence can be subordinated to, rather than weighed against categorical moral reasons.

If there is a single guiding moral of the paper, it is this: once we take objectivity in practical reason seriously, invariance is not an optional embellishment but a constitutive constraint. Making that constraint explicit reveals a powerful and underappreciated route from minimal coherence to robust impartiality, and with it, a principled account of why moral reasons can claim priority within objective deliberation.

Appendix: Formalization of Objectivity = Invariance

This appendix supplies a lightweight formal vocabulary for the bridge principle used in the paper: that objectivity in practical evaluation can be understood as invariance under transformations that preserve normatively relevant structure. The aim is not to burden the main argument with mathematics, nor to suggest that moral philosophy must be mathematized in order to be rigorous. Rather, the aim is clarificatory. A small amount of formalism helps us say, cleanly and without metaphor, what it means to treat certain differences (time shifts, relabelings, perspective changes, etc.) as normatively irrelevant, and what it means for an evaluation to be "objective" relative to those irrelevancies.

The argument proceeds in three stages. First, the relevant formal definitions are introduced and explained. Second, a structural lemma shows that two formal definitions of objectivity are equivalent. Third, substantive ethical conclusions are derived as a theorem (and corollary) from the definitions and lemma, together with the philosophical argument of the main paper.

A.1 Transformations, irrelevance, and the basic idea

Fix a domain of ethically relevant "cases" (or "situations") that we evaluate: actions, maxims, policies, social arrangements, etc. Call this domain W. An evaluation is any map

$$E : W \to V$$

where V is a set of possible ethical evaluative outputs (e.g., permissible, impermissible, or a richer set of verdicts, rankings, or reason-weights).[18] The formal details of V do not matter for our purposes; what matters is that E assigns an evaluative status to each case.

Now suppose we have a family of transformations of cases (e.g., relabelings of agents, time shifts, perspective shifts, etc.) that we take to preserve everything normatively relevant (in the sense defended in the main text). Intuitively, each such transformation takes a case $w \in W$ and produces another case $g \cdot w \in W$ that is "the same for evaluative purposes." The central formal move is to represent these transformations abstractly and then state invariance as a simple condition on E.

A.2 Symmetries, groups and group actions in plain terms

A group G is, informally, a collection of "moves" (i.e., symmetry transformations) that can be composed and reversed. Each element $g \in G$ is a transformation. There is an identity move e that changes nothing. Transformations can be composed ($g_1 g_2$), and each transformation has an inverse g^{-1} that undoes it. A group action of G on W is a rule that tells us, for each $g \in G$ and $w \in W$, what the transformed case $g \cdot w$ is, subject to two natural constraints:

1. **Identity does nothing** : $e \cdot w = w$.

[18] As with the main paper, I mean for the term 'evaluative output' here to be broad enough to include prescriptions, understood as action-guiding evaluations, as well as axiological, deontological, and aretaic evaluations.

2. **Composition matches composition** : $(g_1 g_2) \cdot w = g_1 \cdot (g_2 \cdot w)$.

This is simply a compact way of encoding the idea that we can apply transformations to cases in a systematic way. We can formalize the examples discussed in the paper as follows:

- Time translations: $G = \mathbb{Z}$ (or $\mathbb{R}$) acts on temporally indexed cases by shifting the time coordinate, holding fixed the facts taken to be normatively relevant.
- Permutations of agents: $G = S_n$ acts by relabeling agents, holding fixed the relevant roles, relations, and facts.[19]
- Global perspective shifts: $G = G_P$ acts by shifting point of view (or mode of description) in ways that preserve the relevant structure of the case. Intuitively, these are transformations that change how the case is represented (what is salient, whose standpoint is used to describe it, which descriptive coordinates are employed) without changing what the evaluator takes to matter.

These are all global symmetries, that is, transformations $g \in G$ are applied uniformly across cases (e.g., a time-translation, a person-permutation), yielding a transformed case $g \cdot w$. Global invariance is the condition:

$$E(g \cdot w) = E(w) \quad \text{for all relevant } g, w,$$

where "relevant" means: g preserves the normatively relevant structure of w.

The main text also appeals to a stronger impartiality constraint that is naturally expressed as a local invariance requirement: objective permissibility should be stable not merely under global relabelings (times or persons), but under transformations that may vary by agent and context—changes in evaluative calibration that do not (by assumption) track genuine normatively relevant differences.

A simple way to model this is to allow transformations indexed by agent i and context x. Let G_L be a family of transformations $g_{i,x}$ and suppose $g_{i,x}$ acts on cases by changing the local evaluative framing under which the case is represented (for example, by rescaling a numerical proxy for value, renormalizing salience, or otherwise reparameterizing the agent's representation of the same underlying structure). Then local invariance can be expressed schematically as:

$$E(g_{i,x} \cdot w) = E(w) \quad \text{for all relevant } g_{i,x}, w$$

[19]The distinction between structural identity and bare indexicality defended in section 2.2.1 can be given a precise interpretation within this framework. For any case $w \in W$ involving agents indexed by $i \in 1, ..., n$, the normatively relevant structure of w determines a structural profile for each agent position, i.e., the full constellation of roles, relations, projects, commitments, and experiential states that the occupant of that position bears. A permutation $\sigma \in S_n$ acts on bare indices (which agent occupies which position) while preserving these structural profiles. Invariance under S_n therefore amounts to the claim that evaluation depends only on the structural profiles and their relations, not on which bare index is attached to which profile. This is the formal counterpart of the main text's claim that the bare demonstrative "me" cannot do normative work once all structural features have been specified. The main paper's decomposition of cases into structural identity and bare index could be formalized as a map $\pi : W \times I \to S$ that extracts the structural profile $s \in S$ of agent $i \in I$ in case $w \in W$.

Remark (optional intuition for local invariance): The local invariance requirement can be given a more fine-grained formal expression by representing objective ethical evaluation not only as a pointwise static map $E : W \to V$, but as inducing a directed transition structure on cases: $w \rightsquigarrow w'$, expressing hat the evaluation rationally guides or constrains a transition from w to w' in deliberation. Local invariance is then naturally understood as requiring invariance of this transition structure, not merely of pointwise verdicts, under locally varying recalibrations $g_{i,x}$. This makes vivid why local symmetry, as in the physical case, concerns the **dynamical** or **action-guiding** content of a theory rather than static descriptions. The remainder of the appendix does not depend on this extension; it is noted here as a programmatic direction that connects the ethical framework to the richer mathematical setting of gauge theory, where local symmetry constrains not individual quantities but the laws governing change.[20]

A.3 Invariance: the formal core of OI

Once a group G of normatively irrelevant transformations is fixed, the simplest formal expression of "objectivity as invariance" is:

$$\boxed{\textbf{Invariance} : E \text{ is } G\text{−objective } \Leftrightarrow E(g \cdot w) = E(w) \text{ for all } g \in G,\ w \in W.}$$

This essentially states that applying an irrelevant transformation should not change the evaluation. The formulation here is intentionally minimal. It does not say which evaluation is correct; it says what it is for an evaluation to count as objective relative to the irrelevancies represented by G. If an evaluator claims that time-shifts, relabelings, perspective shifts, and evaluative recalibrations are irrelevant, then invariance is simply what it means to respect that claim.

Two quick interpretive remarks help connect this to familiar philosophical practice:

- **Double-standard diagnoses** can be read as identifying violations of invariance under some relabeling transformation.
- **Diachronic incoherence** can be read as identifying violations of invariance under a time-shift transformation, in cases where relevant facts are held fixed.

Note that nothing in the foregoing discussion commits us to a particular substantive account of "normative relevance." The formalism merely gives us a way to state: *given* a set/group of transformations we agree are irrelevant, what should objectivity amount to?

A.4 Equivalence relations: "the same for evaluative purposes"

The invariance condition can also be expressed without explicitly using groups, using only an *equivalence relation*, which may feel even more philosophically familiar.

[20] Although this appendix is not the place to develop this remark, a companion paper does so, explicitly formalizing ethical local invariance and arguing that Kant's philosophy can be illuminated in light of this formalization (Sanchez Borboa, manuscript)

Given a normatively irrelevant symmetry transformation σ in a family of such transformations Σ, we can state the objectivity as invariance condition as:

$$\boxed{\textbf{Invariance} : E \text{ is } \Sigma\text{−objective } \Leftrightarrow E(\sigma \cdot w) = E(w) \text{ for all } \sigma \in \Sigma,\ w \in W.}$$

We can then define a relation

$$w \sim w' \Leftrightarrow \exists \sigma \in \Sigma \text{ such that } w' = \sigma \cdot w.$$

This says: two cases are equivalent if and only if one can be obtained from the other by a symmetry (i.e., irrelevant) transformation. Because the symmetry transformations in Σ include the identity and are closed under composition and inversion, the relation $\sim$ is an equivalence relation (it is reflexive, symmetric, and transitive). As an equivalence relation, it partitions W into equivalence classes, where an equivalence class $[w]$ is just the set of all versions of the same case you can get by applying irrelevant transformations, formalized as $\{w' \in W | w' \sim w\}$. Each class contains all the cases that differ only by normatively irrelevant variations formalized by Σ. Call the set of equivalence classes $W/\!\sim_\Sigma$. Philosophically, you can think of $W/\!\sim_\Sigma$ as "the space of cases modulo irrelevant differences," i.e., the space of objective case-types relative to the chosen irrelevancies.

Now **Invariance** can be stated using this equivalence relation as:

$$\boxed{\textbf{Invariance} : E \text{ is } \Sigma\text{−objective } \Leftrightarrow w \sim w' \Rightarrow E(w) = E(w').}$$

This basically says: if two cases are the same for evaluative purposes, they must receive the same evaluation. This can be recognized as a formal sharpening of a very common philosophical thought.

A.5 Factoring through the quotient: the cleanest statement of objectivity

The most perspicuous way to state the idea that objectivity = invariance is in terms of a simple "factorization" condition that is logically equivalent to **Invariance**. Let $q : W \to W/\!\sim_\Sigma$ be the quotient map sending each case w to its equivalence class $[w]$. This map plays the role of a quotient by the equivalence relation $\sim$ that counts two cases as the same iff they differ in an irrelevant way specified by a symmetry transformation $\sigma \in \Sigma$. Thus, q takes the space of cases to the space of cases modulo irrelevant differences formalized by Σ. (This is analogous to the more familiar example of clock arithmetic, in which we quotient the integers by the relation that counts two integers as the same iff they differ by a multiple of 12; here, the quotient takes the integers $\mathbb{Z}$ to the integers modulo 12: $\mathbb{Z}/12\mathbb{Z}$ (often written $\mathbb{Z}_{12}$)). Then:

$$\boxed{\textbf{Factorization} : E \text{ factors through } q \Leftrightarrow \exists \overline{E} : W/\!\sim_\Sigma \to V \text{ such that } E = \overline{E} \circ q.}$$

That is, an evaluation E on W factors through q iff there exists an evaluation $\overline{E}$ from the space of cases modulo irrelevant differences to the set of possible evaluative outputs V such

that $\overline{E}$ composed with q is equal to the original evaluation E. This is the formal expression of a familiar idea: an objective evaluation does not depend on the full, messy, representation-laden description of a case; it depends only on its objective structure: its equivalence class under irrelevant transformations.

If E factors through the quotient, then $E(w)$ is determined entirely by the equivalence class $[w]$. Put differently: E can be understood as first "throwing away" irrelevant information (via q) and then evaluating what remains (via $\overline{E}$). This factorization is the formal heart of the main paper's bridge principle. It helps make precise what it means to say that objectivity consists in invariance: the evaluation is a function of the case's structure up to irrelevant differences, not of its arbitrary representational dressing.

The proof of the lemma below makes good on the above claim that these two ways of expressing objectivity (requiring invariance under irrelevant transformations and requiring evaluation to depend only on equivalence classes) are logically equivalent.

Lemma (Factorization $\Leftrightarrow$ Invariance)

Proof: Let $E : W \to V$ be any candidate evaluation. Then:

- (**Factorization $\Rightarrow$ Invariance**): Assume E factors through the quotient, i.e., $E = \overline{E} \circ q$. If $w \sim w'$, then $q(w) = [w] = [w'] = q(w')$. Hence $E(w) = \overline{E}(q(w)) = \overline{E}(q(w')) = E(w')$. Thus, E is invariant, i.e., $\Sigma-$objective.

- (**Invariance $\Rightarrow$ Factorization**): Assume E is invariant. Define $\overline{E} : W/\sim_\Sigma \to V$ by $\overline{E}([w]) := E(w)$. This is well-defined because (by assumption) E is invariant (i.e., $\Sigma-$objective): if $[w] = [w']$, then $w \sim w'$, so by **Invariance**, $E(w) = E(w')$, and thus $\overline{E}([w])$ does not depend on the representative chosen. This implies that for any $w \in W$, $(\overline{E} \circ q)(w) = \overline{E}([w]) = E(w)$, so $E = \overline{E} \circ q$, i.e., E factors through the quotient. $\square$

Philosophically, the proof of this lemma establishes that evaluating cases in an objective way is equivalent to evaluating their equivalence classes under the symmetry transformations that encode normatively irrelevant differences. This equivalence between invariance and factorization reflects a very general structural principle: whenever a system possesses symmetries, quantities that count as objective must be invariant under those symmetries and therefore depend only on the corresponding equivalence classes (often called symmetry orbits). This is the same structural logic that underlies symmetry reasoning across mathematics and the sciences, where identifying invariant structure allows one to separate genuine content from artifacts of representation.

Remark (why the quotient is canonical): The quotient construction above does more than provide a convenient reformulation of the invariance condition; it identifies the precise structure that objective evaluation is allowed to depend on. The map $q : W \to W/\sim_\Sigma$ removes exactly those differences between cases that the symmetry transformations in Σ treat as irrelevant, sending each case to its equivalence class under those transformations. The above lemma thus shows that any evaluation that respects those symmetries must operate entirely at this level: it assigns values not to individual descriptions of cases, but to their equivalence classes. In this sense, $W/\sim_\Sigma$ captures the objective structure of the situation relative to the chosen irrelevancies, and evaluating objectively amounts to evaluating cases only via their equivalence classes under the irrelevant transformations.

A.6 Formal derivation of the conditionality of prudence

It's worth emphasizing that the formal result below is not intended as a free-standing derivation of the conditionality of prudence from logic alone. Its premises (the choice of symmetry group and the identification of its transformations as normatively irrelevant) are the product of the philosophical argument in sections 1–2 of the main paper. What the formalism contributes is explicitness: it states, without ambiguity, what follows once those premises are accepted, and it provides a precise criterion (factorization through the quotient) for when a prudential consideration qualifies as incorporable into objective evaluation. Its value lies in making the argument's commitments and consequences fully transparent.

With the above caveat in mind, we can derive the normative upshot of the main paper, i.e., the conditionality of prudential reasons, as a straightforward consequence of the factorization condition and the structural lemma.

Theorem (Conditionality of Prudence)

Proof: Let $E_p : W \to V$ by any candidate prudential evaluation (i.e., an evaluation that assigns practical significance in virtue of an agent's welfare, projects, or other prudentially indexed considerations). Then:

1. **Admissible Prudence**: If E_p factors through q, then there exists an invariant evaluation $\overline{E}_p$ such that $E_p = \overline{E}_p \circ q$. In that case, the prudential evaluation has a symmetry-invariant core: it can be represented as depending only on its equivalence class $[w]$. Intuitively, the prudential consideration is thereby incorporable into an objective evaluation (its practical significance does not depend essentially on normatively irrelevant self-privileging, but can be expressed in a form that survives the relevant invariances).

2. **Non-admissible Prudence**: If E_p does not factor through q, then there is no invariant $\overline{E}_p$ such that $E_p = \overline{E}_p \circ q$. In that case, the prudential evaluation is not objectively incorporable: its practical significance depends essentially on features treated as normatively irrelevant by the objectivity constraints (paradigmatically, mere self-location, label, or standpoint). Such prudential reasons therefore fail to qualify as objective reasons under OI.

Observation (witness of non-objectivity): Suppose E_p doesn't factor through q. Then (by **Factorization** and the **Factorization $\Leftrightarrow$ Invariance Lemma**), this implies E_p is not $\Sigma-$objective (i.e., not invariant under all symmetry transformations $\sigma \in \Sigma$). This implies (by **Invariance**) that there exist $w, w' \in W$ with $w \sim w'$ such that $E_p(w) \neq E_p(w')$. This formally anchors the diagnosis in the discussion of **Non-admissible Prudence** above by providing a concrete formal witness of non-objectivity: E_p's failure to factor implies it assigns different outputs to cases that differ only by Σ-irrelevant transformations. In other words, the fact that a prudential evaluation fails to factor through q/be invariant under Σ implies that this evaluation fails to treat Σ-equivalent cases in equivalent ways, and so depends on features deemed normatively irrelevant by the objectivity constraints.

Putting together the results established so far, we can derive:

Cond. Prud. : E_p is $\Sigma-$objective $\Leftrightarrow$ E_p factors through $q \Leftrightarrow \exists \overline{E}_p$ such that $E_p = \overline{E}_p \circ q$. $\square$

This derivation of the conditionality of prudence as a theorem from the above formal definitions and structural lemma (together with the philosophical argument of the main paper) yields, in a precise structural sense, the conditionality of prudential considerations: prudential reasons have standing within objective practical reason if and only if they admit a symmetry-invariant representation (a relevant $\overline{E}_p$), i.e., if and only if they can be recovered as (or embedded within) an evaluation that factors through the objectivity quotient q. A helpful way to summarize what this theorem establishes is that factorization is the formal mark of "incorporability" or compatibility with objectivity: what can be carried over unchanged into the objective quotient counts as an objective reason; what cannot is revealed as dependent on normatively irrelevant privileging.

As a suggestive classical application, note that the formal derivation of this theorem also helps illuminate Socrates's famous claim that it is worse to commit than to suffer injustice.

Corollary (Socratic Ordering of Committing and Suffering Injustice)

Proof: Assume, first, that committing injustice constitutes an objective moral defect: it is a failure in the agent's action or practical orientation and so falls under the symmetry-based objective evaluation characterized above. Assume, second, that suffering injustice is bad primarily in prudential terms: it consists in damage, loss, or suffering undergone by the victim and so belongs to the class of considerations whose evaluative force is conditional on their compatibility with objective structure. Then, by the **Conditionality of Prudence Theorem**, prudential evils lack structural priority and are admissible only insofar as they remain subordinate to the objective core of evaluation. It follows that the badness of suffering injustice cannot outrank, or otherwise take priority over, the badness of committing injustice. For committing injustice falls within the objective core of evaluation, whereas suffering injustice is a prudential evil whose badness is conditional on, and therefore subordinate to, that objective core. Hence, under these natural interpretive assumptions, committing injustice is worse than suffering injustice. □

This is not to say that the formalism alone yields all of Socratic ethics from nowhere. Rather, the point is that once one accepts (i) the symmetry-based priority of objective evaluation and (ii) the classification of perpetrated injustice as an objective moral defect and suffered injustice as a prudential harm, the famous Socratic ordering follows as a straightforward consequence of the structural asymmetry already established.

A.7 What this appendix does (and does not) claim

This minimal formalism is deliberately neutral. It does not claim that normativity is group theory or that formal invariances replace substantive moral judgment. It also does not claim that the correct G (or alternatively, Σ) of normatively irrelevant transformations is obvious (although the bootstrapping argument in the main paper gives reasons why G (or Σ), plausibly includes the groups/transformations discussed in this appendix). What it does claim is more modest:

1. If we take some transformations to be normatively irrelevant, then invariance is the natural formal expression of objectivity relative to those irrelevancies.

2. The factorization condition $E = \overline{E} \circ q$ makes explicit what is usually implicit in ordinary

reasoning when we object to double standards, arbitrariness, and capricious reversal.

3. The symmetry-bootstrapping argument in the main text can be understood as an argument about how far we should extend G (or Σ) if we take objectivity seriously, beginning from temporal coherence and extending to local transformations.

In short, the formalism provides a compact language for a philosophical idea already doing work in the paper: objective evaluation tracks what matters, not how it is merely represented. If you ignore all the technical terms, the logic that leads from widely acceptable structural requirements of rational agency to the substantive normative upshot in the paper is simple:

1. Pick which differences are normatively irrelevant (encode them as symmetry transformations in a family Σ).

2. Form the quotient that forgets those differences ($q : W \to W/\!\sim_\Sigma$).

3. Call an evaluation objective if and only it factors through this quotient ($E = \overline{E} \circ q$) and thus does not depend on normatively irrelevant details.

4. Recognize that in order for prudential evaluations to count as objective, they must factor through this quotient ($E_p = \overline{E}_p \circ q$) and thus they must be suitably incorporable into an objective evaluation.

Acknowledgments. I thank Jacob Barrett, Bryan Chambliss, and Jenann Ismael for first discussing with me the idea that Kant's moral law can be understood as a kind of symmetry principle that grounds the objectivity of moral dynamics and that symmetry principles can be fruitfully applied in ethics when I was in graduate school. Their feedback, engagement, and encouragement that the fundamental idea was sound helped motivate me to continue pursuing it for years.

The core idea motivating this paper and its companions (i.e., that structures in ethics can be fruitfully thought of as analogous to structures in physics) was inspired by Kant's own twofold metaphysics of nature and morals, as well as by Stephon Alexander's *The Jazz of Physics* (2016), which compellingly argues for the fruitfulness of drawing interdisciplinary analogies between music and physics. My earlier work on the relation between Kant's natural philosophy and moral philosophy (Sanchez Borboa, 2023) inspired this paper's philosophical analogical project, and Alexander's work inspired me to formalize the analogy and explore the disanalogies as opportunities for innovation. Eugenia Cheng's discussion of how mathematics can illuminate the structure of practical and political phenomena in *Is Math Real?* (2023), and her discussion of equivalence relations in particular in *Unequal* (2025), inspired the appendix's formalization of ethical equivalence relations and the invariance and factorization conditions using tools from abstract algebra.

Finally, I thank my wife, Taylor Boulware, PhD for years of conversation about the philosophical ideas in these papers, for her willingness to think through half-formed arguments and ideas with me, and for her steadfast support of a research program whose development has been anything but conventional.

References

[1] Alexander, S. (2016). *The jazz of physics: The secret link between music and the structure of the universe*. Basic Books.

[2] Bader, R. M. (2015). Kantian axiology and the dualism of practical reason. In I. Hirose & J. Olson (Eds.), *The Oxford handbook of value theory* (pp. 175–201). Oxford University Press.

[3] Brading, K., & Castellani, E. (2006). Symmetries and invariances in classical physics. In J. Butterfield & J. Earman (Eds.), *Philosophy of physics* (pp. 1331–1367). Elsevier.

[4] Broome, J. (1999). Normative requirements. *Ratio*, *12*(4), 398–419.

[5] Broome, J. (2013). *Rationality through reasoning*. Wiley-Blackwell.

[6] Bykvist, K. (2018). Agent-relative and agent-neutral reasons. In D. Star (Ed.), *The Oxford handbook of reasons and normativity* (pp. 821–838). Oxford University Press.

[7] Cheng, E. (2023). *Is math real? How simple questions lead us to mathematics' deepest truths*. Basic Books.

[8] Cheng, E. (2025). *Unequal: The math of when things do and don't add up*. Basic Books.

[9] Gewirth, A. (1978). *Reason and Morality*. University of Chicago Press.

[10] Hare, R. M. (1952). *The Language of Morals*. Oxford University Press.

[11] Hare, R. M. (1963). *Freedom and Reason*. Oxford University Press.

[12] Kolodny, N. (2003). Love as valuing a relationship. *The Philosophical Review*, *112*(2), 135–189.

[13] Kolodny, N. (2005). Why be rational? *Mind*, *114*(455), 509–563.

[14] Korsgaard, C. M. (1996). *The sources of normativity*. Cambridge University Press.

[15] Korsgaard, C. M. (2009). *Self-constitution: Agency, identity, and integrity*. Oxford University Press.

[16] MacIntyre, A. (1981). *After Virtue*. University of Notre Dame Press.

[17] Nagel, T. (1978). *The possibility of altruism*. Princeton University Press. (Original work published 1970).

[18] Nagel, T. (1986). *The view from nowhere*. Oxford University Press.

[19] Neuenschwander, D. E. (2017). *Emmy Noether's wonderful theorem* (Rev. & updated ed.). Johns Hopkins University Press.

[20] Noether, E. (2011). Invariant variational problems. In Y. Kosmann-Schwarzbach & B. E. Schwarzbach (Eds.), *The Noether theorems* (pp. 3–22). Springer.

[21] Nozick, R. (2001). *Invariances: The structure of the objective world*. Belknap Press of Harvard University Press.

[22] Parfit, D. (1984). *Reasons and persons*. Oxford University Press.

[23] Parfit, D. (2011). *On what matters* (Vol. 1). Oxford University Press.

[24] Rawls, J. (1971). *A Theory of Justice*. Harvard University Press.

[25] Raz, J. (2005). The myth of instrumental rationality. *Journal of Ethics & Social Philosophy*, *1*(1).

[26] Sanchez Borboa, S. D. J. (manuscript). The geometry of the moral law.

[27] Scanlon, T. M. (1998). *What we owe to each other*. Harvard University Press.

[28] Scheffler, S. (1982). *The rejection of consequentialism*. Oxford University Press.

[29] Sidgwick, H. (1981). *The methods of ethics* (7th ed.). Hackett Publishing. (Original work published 1907).

[30] Wigner, E. P. (1979a). Invariance in physical theory. In *Symmetries and reflections* (pp. 3–13). Ox Bow Press. (Original work published 1949).

[31] Wigner, E. P. (1979b). The role of invariance principles in natural philosophy. In *Symmetries and reflections* (pp. 28–37). Ox Bow Press. (Original work published 1964).

[32] Williams, B. (1986). *Ethics and the limits of philosophy*. Harvard University Press.

[33] Wolf, S. (1982). Moral saints. *The Journal of Philosophy*, *79*(8), 419–439.