

An Unconditional Duty to the Conditional: [Draft]

An Immanent Critique of Kant's *Pflichtenkollision* Thesis and his Preclusion of the Moral Dilemma

Mark Raftery-Skehan

MANUSCRIPT DRAFT — NOT FOR CITATION OR FURTHER DISTRIBUTION WITHOUT AUTHOR'S PERMISSION

Circulated for comment only. March 2026.

Abstract

Kant's preclusion of genuine *Pflichtenkollision* — his insistence that a conflict of duties is 'inconceivable' (*Metaphysics of Morals*, AA 6:224) — is not a peripheral remark but the architectonic guarantee on which the entire system of distinctions constituting Kantian practical philosophy depends. This paper argues, by immanent critique, that the preclusion fails. Three moments in Kant's own corpus — the contradiction-in-conception demonstration, in which the immorality of false-promising and property-theft is established by showing the logical destruction of institutions the acts presuppose; the *Metaphysics of Morals* doctrine of right, which acknowledges that property, contract, and rightful external relations require the civil condition for their actuality; and *The Conflict of the Faculties*, in which university autonomy proves conferred and bounded by the political authority it would stand apart from — each disclose, without gathering, a constitutive dependence of categorical acts upon worldly conditions that the *Groundwork*'s founding gambit was designed to exclude. What the universalisation test itself reveals is that what is ordinarily taken to be the imperfect duty subordinated to and dissolved by the perfect duty is in fact an unconditional duty to the conditional: the quasi-transcendental conditions of the act's possibility bind the will with categorical force. Taking as its site the university ethics classroom, the paper demonstrates that two perfect duties — each generating a contradiction-in-conception upon universalisation, each grounded in the same concept of rational self-legislation — are rendered simultaneously binding and mutually exclusive under determinate conditions. The paper concludes that the *Critique of Practical Reason* must be re-read not as the demonstration of pure practical reason's self-sufficiency but as the involuntary disclosure of its constitutive exposure to the conditions of a world it cannot command.

Introduction: The Paradox of Teaching Ethics

To teach ethics is not simply to operate within a cloistered space emptied of ethical decision-making, absolved of the normative commitments the ethics seminar seeks to entertain as philosophical objects and submit to theoretical inquiry. The pedagogical acts through which ethics is taught are themselves modes of praxis, and as such cannot but be implicated in the very ethical issues, the very quandaries, the seminar may elect to treat of thematically and theoretically: the principles the teacher submits to the critical reason of their students may serve as the justification for certain pedagogical choices, or be in fact betrayed by others; such principles — say, the cultivation of student autonomy or the social ends the seminar is to serve — may even become themselves the locus of dispute in the face of options between which the teacher, reflecting upon whether the *raison d'être* of their teaching lies in the one or the other, must decide.

This auto-implication of ethics in the ethics seminar manifests itself most acutely in the antagonism between the duty to preserve and cultivate the individual autonomy of the student, and the institutional or social duty to ensure that the teaching of ethics contributes to a more just or ethical society. On the one hand, the teacher must respect and protect the autonomy of the student — the student's right to think, to reason, and to decide freely; the cultivation of their moral reasoning is the end animating the teaching of ethics, and to teach ethics ethically cannot mean to inculcate, to coerce, or to command, but to awaken in the student a capacity for self-legislation, for moral judgment grounded in reason alone — a cultivation that, as Kant's own account of moral pedagogy insists, proceeds not by lecture or prescription but by the erotetic method of guided questioning in which the teacher draws from rather than supplants the student's own reasoning.† On the other hand, the teacher cannot help, however forlornly, but wish that their teaching should bear tangible fruit in the world — that it should not just form citizens capable of morally well-reasoned choices or well-led lives but prove transformative with regard to the social and political conditions in which justice and the good thrive. What is good formally and potentially here sets itself against the good as realised *in concreto*. The first duty is predicated on a certain withdrawal or abstention, the circumscribing of a protected space; the second demands active intervention and engagement with the effects wrought by one's teaching, demands anticipations and calculations measured to the effects it would bring about. This is the double bind of ethical pedagogy: to respect the student as an end in themselves while seeking, through them and beyond them, not merely to

interpret but to change the world. To attempt to resolve this bind by appealing to balance or moderation (*argumentum ad temperantiam*) would be to miss what is at stake — not a practical difficulty of calibration but a philosophical conflict: between a Kantian deontology, for which the duty owed to the student's rational autonomy is unconditional and admits of no qualification by heteronomous social consequence, and a utilitarian consequentialism, for which a pedagogy indifferent to and unanimated by its effects upon the world mistakes the preservation of the agent's good conscience for the stakes that are alone of genuine moral consequence.

Kant's moral system appears to offer an incisive and definitive response. For Kant, the moral law derives its authority not from its effects but from its form: its universality and necessity as expressed in the Categorical Imperative. The good will is good 'not because of what it effects or accomplishes... but because of its volition' (*Groundwork*, AA 4:394). The teacher, then, might take comfort in a Kantian model of ethical instruction that privileges autonomy above all — that forbids the instrumentalisation of students for any political or social purpose, however noble. Empirical derivation produces only conditional obligation; unconditional obligation derives solely from the law requiring the will to conform to its own rational form. To instrumentalise students for social ends — to make the seminar a mechanism for, say, producing democratically committed citizens or for begetting a society prioritising this or that set of just ends, rather than a space for cultivating autonomous moral reasoning — falls foul of the humanity formula of the Categorical Imperative (*Groundwork*, AA 4:429); indeed, even if we were to recognise those students as members of the society at whose improvement one aims — and thus to consider them not thereby reduced to mere means — such instrumentalisation would contravene the universalisation formula by determining the content of the values and principles students hold, rather than fostering the formal modes of reasoning by which they arrive at them. This putative dilemma, between respect for autonomy and utilitarian concern for social efficacy, dissolves on Kantian terms: no genuine collision arises between a perfect duty and a heteronomous end which, by dint of its contingency and arbitrariness — of its being drawn from a domain in which what counts as good or beneficial is empirically and variably determined — can generate only conditional and latitude-admitting obligation, not the categorical necessity morality requires. What is drawn from an empirical world imports its conditionality into what the will that would be moral requires to be absolute and binding in form alone.

The *Groundwork's* opening gambit is to insulate the moral law from precisely such dependence. The moral law governs a noumenal will — a will considered as purely rational, abstracted from all empirical condition, inclination, and circumstance — and its authority is accordingly unconditional: it derives from the formal self-legislation of reason alone and relies upon no particular configuration of the empirical world for its validity as law. The noumenal will legislates from a standpoint that transcends whatever contingent institutions — promising, property, the civil order — happen to obtain in the sensible world, and it is this transcendence that is supposed to guarantee the moral law's authority as categorically binding regardless of circumstance. Yet Kant himself erodes this very transcendence when his own philosophical procedure requires him to acknowledge — however silently, and without drawing the conclusions an immanent critique cannot leave undrawn — the necessary and quasi-transcendental conditions that make possible the acts the moral law is designed to govern. The universalisation test, as will now be shown, is the first and most revealing site of this erosion.⁸

The standard Kantian response to the second pedagogical difficulty — that the conditions making autonomy's exercise possible are the object of a merely imperfect duty, broad and latitude-admitting, which must yield to the perfect duty of formal non-interference — cannot survive scrutiny of the very test by which Kant establishes what a perfect duty is. This is the paper's opening immanent claim, which the paper's title names: an unconditional duty to the conditional. The contradiction-in-conception that demonstrates the immorality of false-promising consists precisely in showing the act to entail the *logical* destruction of the institution of promising — that is, of the worldly, intersubjectively maintained, and hence contingently existing convention of trust upon which the act of promising (and for that matter, false-promising) depends; the act's violation of a perfect duty is established by showing that its universalisation destroys the conditions necessary to its own possibility. The destruction is logical, hypothetical — and thus, in itself, not actual; but the destructibility and the necessity of those conditions must be actual. It is because those conditions are genuinely operative and genuinely necessary to the act that their hypothetical destruction carries categorical moral weight. A contradiction-in-conception applied to a non-existent or merely contingent institution would carry no more force than a prudential warning; the CIC compels unconditionally only because what it shows to be logically destroyable is already at work as the indispensable ground of the act. The same structure governs the case of indoctrination: the conditions of autonomous moral education are already operative, as necessary presuppositions,

within Kant's own demonstration of what constitutes their violation. These conditions, whose *logical* destruction is brought to light in the test of universalisation, are therefore not simply empirical and contingently existent in the sense of being optional or exterior to the moral domain and the locus of duty; they are necessary to the act or institution whose violation the perfect duty prohibits — necessary in the precise sense the false-promising case makes plain. This should hardly surprise us: Kant's philosophical method is defined not by the classical question "what is *X*?" but by the transcendental question "what are the necessary conditions of possibility of *X*?" — and it would be inconsistent for that method to bracket, from the domain of categorical obligation, conditions whose destruction renders *X* itself impossible, conditions inseparable from the act in question. That Kant himself recognises an analogous structural dependence in the political domain — where, in *The Conflict of the Faculties*, university autonomy proves conferred, bounded, and capable of suspension by the political order it would stand apart from — is a concession, however contained, that autonomous acts can depend on conditions extraneous to them. The implication, which an immanent critique cannot leave unexplored, is that the duty to those conditions may carry not the latitude-admitting force of an imperfect duty but an unconditional obligation with the categorical weight of a perfect one: that what appeared as one, singular duty — the duty to autonomy — is in fact divided within itself, bifurcating into a duty to autonomy in its formal identity and a duty to the conditions of its realisation, each capable of placing rival demands upon the educator simultaneously. The collision will be concrete: upholding the formal autonomy of one student may require compromising the socio-political conditions that make autonomous speech possible for others; or, conversely, maintaining those conditions for the many may require constraining the formal autonomy of the one.

Three moments in Kant's corpus, distributed across different domains and never gathered by Kant himself into explicit acknowledgement, each disclose the same constitutive dependence the opening gambit was designed to exclude. The first is the universalisation test itself, as already shown: the CIC of false promising — and of property-theft, the case that even more transparently presupposes a worldly institution — consists in showing that the act's universalisation destroys the institution on which the act depends for the very possibility of its being performed. The moral will in an act of fulfilling a perfect duty of promise-keeping can no more extricate itself from the worldly, contingently instantiated conditions of promising than the gift-giving act can be guaranteed, a priori, to be extricable from the conditions — of recognition, of obligation, of the

circuit of exchange — within which gift-giving is a phenomenon and can be entered into as an act. The second moment is more explicit and more striking: the *Metaphysics of Morals*' doctrine of right acknowledges that property, contract, and rightful external relations cannot be derived a priori from the moral law alone; they require the civil condition for their actuality (MM 6:230–256). Kant concedes, in the domain of juridical rather than ethical philosophy, that the worldly, institutionally constituted conditions of the very acts the universalisation test treats as canonical cannot be generated by pure practical reason. The quarantining of this admission to the doctrine of right — its systematic exclusion from the ethics, where it would press against the moral law's claim to unconditional authority — is itself a structural symptom: the suppression is visible precisely because it must be performed, and performed in a cordoned-off domain. The third moment is the *Conflict of the Faculties*, already noted: the philosophical faculty's autonomy is conferred by the political authority it would stand apart from. None of these admissions is gathered by Kant into the acknowledgement they jointly constitute. What the immanent critique discloses is the formula that Kant's own procedure entails but his architectonic suppresses: the conditions cannot be excluded from the act, because the act's constitutive exposure to them is what makes it the act it is — and a moral act in the second instance. The proximity of this finding to Hegel's empty formalism objection should be acknowledged: in the *Philosophy of Right* (§135), Hegel argues that the categorical imperative presupposes, in its canonical cases — deposit, property, promise — precisely the institutions whose moral status it claims to evaluate. The present paper's immanent critique confirms, from a different angle, the worldly dependence Hegel identifies; but the conclusion drawn differs fundamentally. Where Hegel's response is to abandon the formalist project for *Sittlichkeit* — obligation grounded in historically actualised ethical life — the present paper shows that the dependence on conditions generates, from within the formalist project itself, a collision between two simultaneously binding categorical obligations that no move to concrete ethical life can dissolve. Whether *Sittlichkeit* itself escapes the practical antinomy identified here, or merely relocates it, is pursued in a companion essay.⁶

Kant, confronted with such a dilemma, would deny that it is possible. 'A conflict of duties (*Pflichtenkollision*) is inconceivable,' he insists (*Metaphysics of Morals*, AA 6:224). Reason cannot legislate contradictory necessities; apparent collisions arise only from misclassification — one duty will prove to be merely imperfect, grounded in contingent ends, and must yield to the perfect duty that reason commands unconditionally. This preclusion is not a peripheral remark at

the margin of Kantian ethics but the guarantee upon which the solidarity of the entire architectonic of pure practical reason depends: the system of distinctions constituting Kantian morality — form/matter, categorical/hypothetical, autonomy/heteronomy, perfect/imperfect duty, the contradiction-in-conception and the contradiction-in-will — holds together as a system precisely insofar as genuine dilemma is excluded from within it. To demonstrate that the duty to the quasi-transcendental conditions of autonomy's possibility and realisation carries categorical force is therefore to show not merely that one element has been misclassified but that the imperfect has been elevated into the perfect — that the hierarchical distinction upon which the system's coherence depends has been inverted from within, unsettling the very purity of the distinctions the exclusion of dilemma requires. That the pedagogical scenario this paper analyses may disclose a strain within this guarantee is what the following argument undertakes to show. Here the Kantian system turns upon itself. The categorical imperative, which binds the will purely by reason, already presupposes the empirical conditions of its possibility and realisation in acts — conditions whose logical destruction in the universalisation test simultaneously renders impossible the very acts the imperative is designed to govern. Autonomy is not the privilege of an isolated self but the form of a shared rationality, dependent upon mutual recognition, communication, and trust. These are not hypothetical ends but the very conditions of autonomy's possibility and realisation. For a philosophy defined not by the classical question 'what is *X*?' but by the transcendental question 'what are the necessary conditions of possibility of *X*?', it would be inconsistent to exclude from the scope of categorical obligation conditions whose systematic destruction renders the exercise of *X* itself impossible, conditions Kant recognises to be intrinsic to the thing's very identity. To preserve them may thus be not an imperfect duty of beneficence but one that responds to an obligation with all the force of an unconditional command. The teacher's dilemma then ceases to be a merely pedagogical difficulty and becomes a philosophical one: the exposure of a contradiction within the Kantian denial of contradiction itself. This paper proceeds from that recognition. It argues that the Kantian preclusion of genuine dilemmas depends on the purity of distinctions — between autonomy and heteronomy, the categorical and the hypothetical, form and matter — that cannot, in practice, be maintained. The ethics classroom, as a microcosm of moral life, exposes their impurity: the law of autonomy depends upon the conditions of its realisation, and the respect for freedom demands the preservation of the world in which freedom can exist. The teacher's double bind thus reveals the fragility of Kant's edifice: its reliance on the exclusion

of what is indispensable. To teach ethics, then, is to inhabit the very fault line of reason — to enact, in the modest sphere of pedagogy, the paradox that Kant's system must deny: that practical reason, in order to be consistent with itself, must will not only the moral law's formal validity but the conditions of possibility in which alone that law can be realised — and discovers in those conditions the dilemmas it had declared inconceivable.

In a stable liberal order, this paradox may remain latent, masked by the tacit assumption that the university, the seminar room, and the wider polity already guarantee the freedoms upon which moral education depends. That those guarantees can be objects of live political contest — that the conditions of autonomous rational discourse are never historical achievements safely secured once-and-for-all but can be sites of ongoing erosion to whose degradation the conscientious teacher of ethics must attend — means that the dilemma re-emerges, at intervals and with varying intensity, in its pure form. The ethics teacher may under such conditions confront situations where they must decide whether to defend the autonomy of their students by enforcing the norms of democratic dialogue, or to respect autonomy by allowing the expression of views the aim of which is the destruction of such would destroy those norms. In either case, autonomy — that of the emboldened speaker or of those whose autonomy is denied — is both served and betrayed.

† Kant distinguishes the acroamatic method — lecture and the prescription of moral content — from the catechistic or erotetic method, in which the teacher 'guides his young pupil's course of thought merely by presenting him with cases in which his predisposition for certain concepts will develop' (*MM* 6:478), functioning as 'the midwife of the pupil's thoughts.' The method resembles Socratic maieutics in form but differs from Socratic elenchus in aim: where Socrates proceeds toward *aporia*, Kant's teacher proceeds toward a determinate telos — the student's subjective autonomy, their capacity to give themselves their own maxims, such that 'your own reason teaches you what you have to do and directly commands you to do it' (*MM* 6:481). The distinction is developed in Section 1 below.

* On *quasi-transcendental*: the term is Derrida's, first deployed in his reading of Husserl in the *Introduction to Husserl's Origin of Geometry* (trans. Leavey, University of Nebraska Press, 1989, pp. 87–90). The argument there concerns not the ideality of the geometrical object as such — which transcends any act of cognising or failing to cognise it — but geometrical ideality as a *possession of the tradition*: as transmissible across generations, reactivatable as knowledge, available to a community that transcends the finitude of any particular group or individual. For this possession, writing is necessary — not this or that particular text, and thus not any or even all contingently existing texts. The logic is precise: geometrical ideality requires neither this manuscript nor that textbook; the destruction of any particular text has no bearing on it; and since the destruction of any particular text has none, the destruction of all contingently existing texts equally has none. What geometrical ideality as a possession of the tradition requires is the *perpetual possibility of being vouchsafed to writing* — a possibility ensured not by the existence of any empirical text but by the very conditions of possibility of writing as such: the sign, language, and the originary iterability that constitutes them. For a sign to be a sign, it must be capable of being repeated — of functioning in the absence of any particular empirical context, any particular sender or addressee.

This reproducibility is not a contingent property of signs but their constitutive condition: a mark that could not in principle be reinscribed would not be a sign at all. The perpetual possibility of geometrical ideality's being vouchsafed to writing is therefore indestructible so long as there is language — so long as the sign's originary iterability obtains. Writing is *quasi-transcendental* in this precise sense: necessary to the tradition's possession of ideality, yet not a priori derivable from the concept of ideality itself, since it requires the material structure of the iterable sign rather than being analytically entailed by ideality's own concept. The conditions the present paper analyses are quasi-transcendental in the same sense: what is required is not any particular institution of free speech or any specific university but the structural possibility of mutual recognition, free rational discourse, and equal discursive standing — conditions necessary to the acts the Kantian framework governs, neither this nor that particular empirical instantiation being strictly necessary, yet requiring empirical maintenance and susceptible to systematic foreclosure. That Kant himself recognises an analogous structure within his practical philosophy — in the *Doctrine of Right* (MM 6:230–256), where the conditions of rightful external freedom require positive institution and empirical maintenance rather than being a priori given — constitutes an internal precedent: the quasi-transcendental structure is not an external imposition on the Kantian framework but a tension it has always harboured.

1. The Condition of Autonomy

The term *Autonomie* — Kant's coinage from the Greek *auto-nomos*, self-lawgiving — names the property of the will by which it is a law to itself, independently of any object of desire (*Groundwork*, AA 4:440–442), and stands opposed to *Heteronomie*, the determination of the will by empirical motives or contingent ends, which yields only hypothetical imperatives incapable of providing the unconditional law that morality requires.

Yet even as Kant insists on its purity, his own corpus shows that autonomy's exercise presupposes relations that enable a will to recognise itself as lawgiving, and others as likewise free and self-legislating. If the Categorical Imperative binds all equally, it also binds each to all others, since the capacity for self-legislation must be mutually recognised. Perfect duties — 'do not lie', 'do not steal', 'do not make false promises' — protect this reciprocity by prohibiting acts that undermine another's autonomy. In Kant's example of the false promise, the would-be borrower deceives the lender, depriving them of the knowledge required to decide freely; the wrong lies not merely in the loss of property but in its having compromised the other's self-legislating power to decide with all relevant information necessary to a genuinely free decision having been provided. Autonomy, then, is not the possession of isolated subjects but a structure of intersubjective commerce, sustained by trust, reciprocity, and respect — the minimal architecture of any community of rational agents. These conditions are not moral embellishments, tacked on to beings legislating for themselves in isolation, but the formal prerequisites of a self-legislating moral life itself.

Christine Korsgaard's reading of the Formula of Humanity (*Creating the Kingdom of Ends*, ch. 3) sharpens this point: to treat rational nature always as an end is not merely to refrain from violations of agency but to sustain the conditions under which rational agency can be exercised and expressed — the Formula of Humanity already harbours a positive, not merely prohibitive, dimension. The present paper draws the conclusion Korsgaard herself does not: that this positive dimension generates, under determinate conditions, a duty to the conditions of autonomy's intersubjective exercise capable of colliding, as a perfect duty, with the duty to formal autonomy.

The progressive formulations of the Categorical Imperative make this intersubjective dimension increasingly explicit. The Formula of Universal Law requires that every rational will be able to will its maxim as universal law, with the other implicit as counterpart in the universalisation test; the Formula of Humanity names the other explicitly as an end in themselves — the being in whom my self-legislation finds its confirmation; the Formula of the Kingdom of Ends gives autonomy a communal form, in which each will legislates for a whole of rational beings bound under laws they themselves have given. In the ethical commonwealth (*Metaphysics of Morals*, 6:94) that Kant envisions as the telos of moral life, these relations of mutual recognition and self-legislation are not a supplement to morality but that upon which the realisation and flourishing of each person's autonomy depends.

These relations — trust, reciprocity, communicative freedom — are the quasi-transcendental conditions under which autonomy can appear and be realised. Though not derived from experience, they must obtain in experience if moral agency is to be possible at all. To violate them through deceit, coercion, or denial of discourse is not only to act immorally but to undermine the very space of moral action. What the universalisation test discloses in its canonical cases is this: the immorality of false promising rests not on the harm done to any particular person but on the *de jure* destruction of the intersubjectively held conventions of trust upon which the institution of promising depends. The moral law already draws upon, and cannot dispense with, the very conditions it seeks to transcend.

Unlike Kant's own transcendental conditions — the categories, the forms of intuition — which are established a priori and cannot be contingently absent, these conditions exist and are instantiated historically and politically, susceptible to erosion, denial, and systematic foreclosure; it is this combination — necessary as a structural possibility, yet contingently instantiated rather

than a priori given — that the term *quasi-transcendental* is designed to capture, and that explains why the duty to secure them, though it cannot be derived a priori from the moral law, may nonetheless carry categorical rather than merely latitude-admitting force. A practical reason that wills the moral law but remains indifferent to the conditions without which that law cannot be realised contradicts itself — it wills an end while refusing what that end requires (*Groundwork*, AA 4:417) — and this self-contradiction is itself grounds for recognising the duty to those conditions as carrying unconditional rather than merely discretionary force.

This recognition reappears in Kant's own account of moral pedagogy in the Doctrine of the Methods of Ethics (*MM* 6:477–481). The very concept of virtue, Kant insists, already implies that it must be acquired rather than innate, and this acquisition cannot proceed by exhortation or the imposition of moral conclusions — virtue 'cannot be taught merely by concepts of duty or by exhortations (by *paraenesis*)' (6:477). The method adequate to the aim of moral education is accordingly the erotetic or dialogic method, in which 'the teacher, by his questions, guides his young pupil's course of thought merely by presenting him with cases in which his predisposition for certain concepts will develop'; the teacher is, in Kant's own phrase, 'the midwife of the pupil's thoughts' (6:478). The telos of this practice is made explicit at 6:480: 'a maxim of virtue consists precisely in the subjective autonomy of each man's practical reason' — the student is to reach a point at which imitation, the first stage of will-formation, has given way to maxims 'he afterward makes for himself' (6:480), such that, as the model catechism has the teacher say, 'your own reason teaches you what you have to do and directly commands you to do it' (6:481). Moral pedagogy, on Kant's own account, is therefore not the transmission of moral conclusions but the cultivation of the student's capacity for rational self-legislation — a cultivation whose method is accordingly one that draws from, rather than supplants, the student's own reasoning. The teacher therefore bears a dual responsibility: to respect the student's formal autonomy and to sustain the conditions within which that autonomy can be meaningfully exercised.

2. The Impossibility of Dilemma

What might appear as a conflict between respect for students as ends-in-themselves and the social ambitions of ethical pedagogy is, within Kant's system, a pseudo-dilemma: it opposes a perfect duty grounded in the formal law of pure practical reason to a heteronomous end whose

conditionality is imported from the empirical domain. Since no necessity derivable from contingently selected ends can match the categorical authority of the moral law, the apparent conflict dissolves asymmetrically — the perfect duty commands unconditionally; the social ambition, as a merely conditional end, duly yields.

If the humanity formula lays emphasis on students' inherent autonomy, precluding them from being reduced to mere instruments, the universalisation formula highlights (i) the inconsistency by which the agent could never logically will their own autonomy be so compromised, and (ii) the logical consequence, entertained within the hypothesis of the maxim's universal adoption, that the very conditions of possibility of education pursued in accordance with its proper end would be destroyed. To aim to inculcate with determinate content is to undertake an enterprise incompatible with the fostering of moral autonomous reasoning; education is brought to the point of incompatibility with its very concept.

Kant's doctrine that 'a collision of duties and obligations is inconceivable' (*collisio officiorum s. obligationum; obligationes non colliduntur*, *MM*, AA 6:224) expresses that all such putative ethical dilemmas dissipate upon inspection and specification of the duties supposedly in collision. This preclusion follows as a structural corollary of the whole system of distinctions that constitutes Kant's moral architectonic: the categorical as opposed to hypothetical imperative; the perfect as opposed to the imperfect duty; the rational law that binds the autonomous will absolutely as opposed to the heteronomous end that obligates conditionally. Apparent *Pflichtenkollisionen* are, for Kant, either (a) misapplications of what is truly and categorically obligatory, or (b) a clash between a perfect duty — the absolute prohibition on compromising students' autonomy — and an imperfect duty, the discretionary and latitude-admitting duty the educator might conceive themselves as having regarding the effects of their teaching upon the broader society. A pseudo-dilemma of this kind dissolves once the hierarchy is correctly identified: the perfect duty commands unconditionally; the imperfect duty duly yields.

The distinction between perfect and imperfect duties corresponds to the two distinct modes of failure produced by the universalisation test. A contradiction-in-conception (CIC) occurs when the universalised maxim conceptually destroys the very conditions of possibility the act requires, rendering it self-defeating — and yields perfect duties: strict, exceptionless, categorical prohibitions. The instrumentalisation of students that would compromise their autonomy is

precisely such a case: the adoption of a maxim by which every educator renders students mere means to predetermined social ends would deprive education of the conditions in which its aim — the cultivation of autonomous reasoning — is conceivable.

The logic of the CIC, and of the universalisation test that generates it, is clarified by Onora O'Neill's analysis in *Constructions of Reason* (chs. 5–8). O'Neill argues that the universalisation test is not merely a check for formal logical consistency but a test for what she calls principles that can be 'action-guiding for all' — principles that can be adopted and acted upon by all rational agents in a shared world without structural self-defeat. The failure of non-universalisable maxims, on her account, is precisely their parasitic structure: a maxim of deception requires, for its execution, that others maintain honesty — that they not also deceive; a maxim of false promising requires that others honour the institution of promising — that they not also make false promises. The maxim draws on a condition — trust, the institution of promising, the standing of others as credulous rational agents — that its universalisation destroys. This is a practical self-defeat prior to any formal logical contradiction: the maxim cannot be 'a principle for all' because its efficacy in any individual case depends on its not being everyone's principle.¹ The CIC, understood in O'Neill's terms, is the disclosure of this structural parasitism: the moment at which the universalisation test reveals that the act presupposes, as the condition of its possibility, precisely what its universalisation would destroy. This reading underwrites and deepens the argument of the present paper: what the teacher's dilemma discloses is not a formal contradiction in the abstract but the parasitic self-defeat of two maxims, each of which presupposes, as the condition of the educative practice it would govern, precisely what its universalisation destroys.

O'Neill's complementary argument in 'The Public Use of Reason' (*Political Theory*, 14.4, 1986) extends this to the institutional conditions of moral reasoning: Kantian ethics requires not merely formal respect for rational agency but the communicative conditions under which principles can be publicly articulated and subjected to rational scrutiny — conditions constitutive, not merely instrumental, to the practice of public reason.

In the case of the contradiction-in-will (CIW), it is a matter of a self-frustrating volition: the act being not logically self-annihilating but inimical to the ends willed by a rational being (the neglect of one's talents frustrating one's own development; the maxim of never helping others being at odds with the contingency by which one will on occasion oneself need help). The CIW

does not annihilate the conditions of the institution in question; it merely impedes an end no rational agent could will to be thwarted. Were a maxim adopted by which one political ideology were tacitly favoured over all others, the rational educator could not coherently prefer such a maxim to be universally adopted; but the world in which it is so adopted is not logically inconceivable: education can persist, however imperfectly. Thus does there arise not a CIC but an imperfect duty to cultivate conditions favourable to student autonomy.

The CIC against indoctrination is therefore not merely against the external conditions of education but against its constitutive concept: as Kant's own account of moral pedagogy in the *Doctrine of the Methods of Ethics* establishes (Section 1 above), a practice that forecloses the student's capacity to give themselves their own maxims contradicts education's concept by denying the student the standing of the self-legislating rational subject that *Bildung* presupposes and cultivates.

Since perfect duties express the form of rational autonomy itself, their status as commanding absolute duty must be recognised; since the imperfect duty bears only on a heteronomous, exterior, and contingent end, no genuine dilemma can ever arise between the two. The perfect duty delimits the space within which the imperfect duty can be fulfilled (*MM*, AA 6:224; *Groundwork*, AA 4:421–424). Imperfect duties afford a certain latitude in the manner and timing of their execution — what Kant calls *latitudo*, the space of free choice in complying with the law (*MM*, AA 6:390). The unconditional excludes the conditional. What appears, then, as a dilemma between the perfect duty to student autonomy and the imperfect duty to social efficacy is, on Kantian terms, a pseudo-dilemma: the former commands absolutely, the latter gives way. The task this paper sets itself is to demonstrate, from within the Kantian system, that this resolution is insufficient — that the duty to autonomy can, under determinate conditions, generate not a conflict between the perfect and the imperfect but a collision between two perfect duties, two CICs grounded in the same concept. It is to that demonstration that we now turn.

Before that demonstration proceeds, the argument's ground should be made explicit in the form it takes from Kant's own practical philosophy. A practical reason that wills the moral law necessarily wills the conditions of its realisation: to will an end while remaining indifferent to the means without which it cannot be achieved is, as Kant himself acknowledges, a form of practical self-contradiction (*Groundwork*, AA 4:417). Applied not to contingent ends but to practical

reason's own necessary end — the realisation of the moral law in the world of agents subject to it — this principle yields a conclusion Kant does not draw but cannot without inconsistency deny: that practical reason necessarily wills the quasi-transcendental conditions of the moral law's exercise, and that a duty grounded in practical reason's own necessary willing cannot be relegated to the latitude-admitting discretion of an imperfect duty. The *Critique of Practical Reason's* own doctrine of the highest good corroborates this at a deeper level: Kant acknowledges there that pure practical reason must will not only the moral law's formal validity but the proportionality of virtue and happiness in the sensible world — that practical reason cannot be self-sufficient with respect to the conditions of its own realisation (*CPrR*, AA 5:110–114). The pedagogical dilemma the following phases demonstrate is the precise, concrete, and technically verifiable form this insufficiency takes when the quasi-transcendental conditions of autonomous moral education are themselves placed in jeopardy.

3. The Conditions of the Dilemma's Realisation

Phase 1: Democracy as One Ideology Among Others

If democracy is conceived as one more ideology among others, and solely as theoretical content that can be imparted rather than as a form that can be instituted, then the educator who surreptitiously favours this ideological content in their teaching falls foul of the universalisation test; they do so on account of their usurping or undermining the very conditions by which material is to be set before the free critical evaluation of students in keeping with the educational aim of their autonomy being fostered. Beyond the case of unconscious bias — the mere incidental favouring of democracy through unexamined pedagogical assumptions, which would qualify as a CIW — its conscious imposition upon students would contravene education's very *raison d'être*. In either case, no dilemma emerges: the fulfilment of the duty to non-indoctrination stands against a violation of a CIC or a CIW respectively.

The structural concession Kant makes here, and which will prove decisive in what follows, is this: the immorality of the indoctrinatory maxim rests not on the formal violation of the autonomy of any given, particular student, nor on the empirical likelihood that indoctrination will actually destroy its contingently existing conditions, but on the *de jure* damage that the act, as

universalised, would in principle wreak upon the quasi-transcendental conditions upon which the act of education depends: the essentially empirical yet necessary conditions whose logical destruction constitutes the CIC as such. The practice of moral education is possible, viable, and intelligible only by virtue of intersubjective and institutional conditions whose existence the practice presupposes; just as the false no less than the genuine promise can only be what it is so long as the intersubjectively held conventions of promising are in place, so indoctrination and genuine education alike presuppose the conditions of free rational self-legislation — the student's capacity to question, dissent from, contest, and revise what is presented to them. These conditions — institutional, intersubjective, and essentially empirical in character, yet necessary to the acts they enable — are what we may designate, in a precise sense, as quasi-transcendental.

Phase 2: Democracy as Conducive to Autonomy

To confine democracy to being ideological content would be to overlook how, within an educational setting, it may be implicated as a set of norms governing behaviour: civility in disagreement, the submission of contested claims to reason as ultimate arbiter, the illegitimacy of *ad hominem* argument, the right to dissent from an orthodoxy and to voice an opinion without intimidation, tolerance of plurality, and so on — each norm having its social analogue in a democratic polity. These norms are not the exclusive creation of any individual teacher: their provenance lies variously in the university's code of conduct, in the unwritten protocols governing academic exchange, in the performative assumptions implicit in the practice of the seminar itself. What the teacher does, in implementing them, is less to found than to enact — to enforce, and in enforcing, to repeat and sustain, while assuming responsibility for, a structure of imposition whose origins exceed any particular pedagogical relation. Yet the moment of enforcement harbours a unilateral dimension that the norms' ordinary operation tends to conceal: these remain norms the students will not themselves have legislated in any direct or immediate sense, whose adoption by the class is presupposed rather than solicited. Student autonomy is thus, in the first instance, bypassed: the institution of these norms has not issued from a process of communal deliberation in which the students' consent was first secured.

The democratic norms in question close off any meaningful contestation of themselves even as they seek to ground the right to meaningful contestation of all discourse conducted within the framework they erect. Their authority rests not on the consent of those they govern but on an

imposition that precedes any such consent and that constitutes the very framework within which consent could subsequently be given: a structure whose underlying violence is rendered visible only perhaps at the moment of its contestation. The educator's passing non-recognition of autonomy in the imposition of democratic norms must nonetheless be distinguished from the double violation incurred in the institution of non-democratic, hierarchical, or exclusionary protocols. That the democratic norms install conditions conducive to autonomy's preservation disbars us from determining the initial violation as a CIC.

Against any charge of heteronomy in Reath's sense — acting on the authority of an institution rather than reason — the teacher's imposition of democratic norms is not grounded in institutional command. The teacher imposes the framework not because the university's code of conduct requires it but because no rational agent could will a world in which the conditions of autonomous self-legislation are left entirely exposed: the universalisation test itself reveals a CIW. The ground is rational, not institutional.

Kant is by no means unaware of the possibility of autonomy being conferred upon the agent from without. University autonomy, as he acknowledges in *The Conflict of the Faculties*, is conferred by the state that lies beyond it: the philosophical faculty acquires its rights and freedoms from the state, enjoying the freedom to judge upon truth — to exercise its rational judgment on whatever subject it fancies — on condition that it refrain from issuing such judgements in the public domain, judgement and its publication being deemed irreducibly different, as belonging to the orders of truth and action respectively. This distinction between the purely rational verdict and its public expression is one whose stability the nature of language itself, as Derrida notes, cannot underwrite, since language is constitutively oriented toward iterability and an absent addressee: the moment thought takes linguistic form, it has already entered the circuit of potential public reception that Kant's distinction presupposes it can contain.² Where the condition is breached — where the philosophical faculty carries its corrective judgment upon a tenet of a higher faculty into the public sphere — what had been a legal conflict within academic walls becomes, when it spills over into the public realm and may exert an influence on the people, an illegal act.³ The educator who imposes democratic norms within their classroom inhabits a structurally analogous position: their right to do so is implicitly sanctioned by the broader institutional and political order, a sanction that is no less conditional for ordinarily remaining unremarked upon.

This *a-legality* will prove consequential in Phase 3, where the emboldened student's act of contesting the norms will be shown to expose, rather than simply violate, the founding violence that the norms' ordinary operation keeps invisible. The structure of the founding act itself — as distinct from the conferred character of what it founds — is more precisely described by the concept Derrida identifies, in his commentary on *The Conflict of the Faculties* and in *Mochlos*, as *a-legality*. The act by which any institution of freedom is established is neither legal nor illegal vis-à-vis the order of legality it founds: since that order does not yet exist at the moment of foundation, the act cannot be authorised or legitimated by it; the founding act therefore precedes and makes possible the very distinction between legal and illegal, without itself being simply either. The imposition of democratic norms within the classroom is *a-legal* — not authorised by the democratic order it seeks to produce, prior to the very distinction upon which any judgement of its legality would depend. It is this *a-legality* that the student who contests the norms — wittingly or unwittingly — invokes: in refusing to be bound by a framework they did not consent to, they appeal to a principle — democratic self-determination — that the framework itself upholds, and that the framework's own founding act was unable to secure.

The violation of student autonomy does not, for all that, escape amounting to a CIW. The depriving of the student of a certain self-legislative capacity vis-à-vis their governance qua participating student constitutes an irreducible infraction which, however teleologically justifiable and however transitory, cannot simply be rationally preferred over no such infraction. The paradox is precise: the transgression is at once at odds with the act's end — the securing of autonomy — and a necessary means to it. It is from this irreducible tension that there arises an imperfect duty: the duty, with all the latitude that thereby pertains to it, to foster conditions conducive to the end of education, including that of autonomous expression.

Now, the status of the initial violation of student autonomy — as an infraction of a perfect or imperfect duty — proves conditional upon the ongoing maintenance and subsequent enforcement of the norms originally instituted. We said that the violation amounted to a CIW rather than a CIC precisely because those norms would in fact serve to secure student autonomy. If, however, this securing fails to take place — if the teacher allows a student's right to speak to be drowned out by *ad hominem* argument — then the initial violation is retroactively transformed in character: it becomes a CIC. It is no longer a momentary infraction within a broader horizon of

securing autonomy; it is an imposition of norms that, having failed to deliver the conditions they promised, are indistinguishable, functionally, from the non-democratic imposition of Phase 1. Kantian thought can countenance the deferral of an imperfect duty in order to fulfil a perfect one, but not the status of an act as a violation of a perfect or an imperfect duty being determined by the contingent fulfilment of a further imperfect duty. This conditionality of the unconditional is, for Kant, inconceivable.

Kant's system of distinctions begins here to unravel. (The full account of what this unravelling implies — its systematic implications for the interconnected series of distinctions that constitutes Kant's moral architectonic: form/matter, autonomy/heteronomy, law/condition, perfect/imperfect duty, CIC/CIW, conditional/unconditional, categorical/hypothetical — is reserved for the Conclusion, where each of these distinctions is shown to be implicated in the others' instability.) What the difficulty discloses is that autonomy has to be conceived as always already dislocated, as conditionally subject to the supposedly extraneous factors that facilitate or impinge upon its free exercise. The inner freedom to legislate for oneself remains abstract and impoverished until the communal conditions within which such self-legislation tends to be practised and accommodated are taken into account.

This shift in our conception of autonomy alters, in turn, the scope of the duty we owe to it. It will not suffice for the teacher to ensure that they in no way infringe upon the formal and abstract autonomy of the student; their duty will extend to those conditions — including the democratic norms by which autonomous participation in the seminar is safeguarded — without which that formal autonomy remains inert. We have already seen, in the analysis of the CIC and its constitutive dependence upon the quasi-transcendental conditions of the acts it governs, that this very hierarchy is impure on Kant's own terms: the means by which Kant demonstrates the categorical force of the duty not to violate formal autonomy already draws upon, and cannot dispense with, the very quasi-transcendental conditions of autonomy's exercise that Kant might otherwise seek to demote to a secondary, merely imperfect status. In conceiving of autonomy as extending to an intersubjective, social, and politico-institutional setting on which the exercise of autonomous acts in fact depends, we may discern the bifurcation of the duty owed to others as autonomous beings: such that the duty not to violate autonomy in its formal sense may stand alongside a no less binding duty to the conditions within which the person can actually and

meaningfully exercise their autonomy. It is when faced with these demands, such as they may obligate one simultaneously and thus in the form of a double-bind, that we can glimpse the formal structure of a dilemma contravening the Kantian preclusion of *Pflichtenkollisionen*.

Before we arrive at this dilemma in concrete form, it remains to ask whether the educator might simply refrain from implementing democratic norms altogether. Non-implementation avoids one infraction of student autonomy, but only at the cost of another: in the lawless vacuum left by non-implementation, infractions by which students curtail the autonomy and rights of others meet with no resistance. The teacher's non-act becomes complicit with a rationale no educator can rationally will — for to will a world in which every teacher abstains from implementing any framework of norms is to will a world in which the stronger voice always silences the weaker: a world the universalisation test condemns, since no rational being could will that the conditions allowing for their own free participation in discourse be left entirely exposed. There arises here, then, an imperfect duty to promote the conditions conducive to autonomy. We find, in consequence, a perfect symmetry of imperfection: to act is to compromise autonomy in its form; to refrain is to allow for its annihilation in exercise. The will can form a coherent conception of each world — neither scenario generates a CIC, neither is logically inconceivable — yet neither world can be simply or purely rationally willed, since each involves a CIW.

To act is to will a world in which autonomy is formally violated as a condition of its being secured; to refrain is to will a world in which its violation by others is left unopposed. This is not yet a dilemma in the Kantian sense — the collision of two perfect duties — but its antechamber: the collision of two imperfect duties. Marcia Baron's analysis of *latitudo* (*Kantian Ethics Almost Without Apology*, ch. 3) is instructive here: imperfect duties afford latitude in execution — the freedom to determine how, when, and to what degree the duty is fulfilled — because the situations to which they respond are general, ambient, and temporally open. Each world in Phase 2 is conceivable; neither forecloses the conditions of education as such; and the teacher retains latitude to act or refrain, to time and calibrate the implementation of norms, to correct subsequent failures. Baron herself acknowledges that latitude never eliminates the obligation it qualifies; what it governs is only its particular execution. Phase 3, however, does not merely narrow this latitude to a vanishing point — it does not represent a situation in which the imperfect duty has narrowed to the point where it must be executed this very instant. What Phase 3 introduces is a qualitative shift

of a different kind: not an exhaustion of latitude within the domain of imperfect duty, but the arrival of a situation in which the conditions at stake are shown to be not merely conducive to autonomous practice in the educational setting but its *sine qua non*. This is a transformation in the logical type of the duty — from imperfect (whose violation generates a CIW) to perfect (whose violation generates a CIC) — and not a matter of imperfect duty becoming more pressing. Baron's account, operating entirely within the domain of imperfect duty and its internal gradations, has no resources to account for this kind of type-shift.

It is important to register precisely why latitude is available in Phase 2 but foreclosed in Phase 3. In Phase 2, the teacher's imperfect duty to promote the conditions of autonomous discourse admits of calibration over time — of incremental imposition, of subsequent correction, of opening space in later sessions where earlier constraints were too severe. The horizon of fulfilment is extended. Phase 3 removes this temporal opening entirely. The scenario it stages concerns a specific, immediate, and irreversible act: a particular student in the *hic et nunc* of a class in progress, actively denying other students the standing to speak and reply — not through some ambient hostility to open discourse but through a determinate act whose effects are immediate and whose occasion for remedy, once passed, cannot be restored. No future session, no subsequent correction, can return to those students the moment of participation that has been destroyed. This singularity is what collapses the latitude of the imperfect duty and forces the teacher into a choice that bears the structure of a collision between two categorically binding obligations. The type-shift from Phase 2 to Phase 3 is therefore not the imperfect duty becoming more urgent while remaining imperfect; it is the conditions themselves being revealed — through the immediacy and specificity of the scenario — as constitutively necessary rather than merely conducive, and thereby as the ground of a perfect duty.

Phase 3: Democracy as Meta-Ideology and the Genuine Dilemma

The analysis of the preceding phases has forced upon Kant two concessions that are not peripheral admissions but axioms upon which the present phase's immanent critique proceeds. From Phase 1: the immorality of the indoctrinatory maxim was shown to rest on the *de jure* damage that universalisation would wreak upon the quasi-transcendental conditions upon which the act of education depends — the essentially empirical yet necessary conditions whose logical destruction constitutes the CIC as such. From Phase 2: to demonstrate that the teacher's imposition of

democratic norms does not itself generate a CIC, Kant must implicitly concede that autonomy's expression is subject to conditions — conditions that are material, intersubjective, and institutional, whose presence enables and whose absence frustrates the full exercise of autonomous acts. These two concessions are now operative as axioms.

Phase 3 will show that this concession has not been granted in its full implication. As the Doctrine of the Methods of Ethics establishes, the telos of moral pedagogy is the student's subjective rational autonomy — their capacity to give themselves their own maxims, to contest and revise positions under the pressure of reasons. This subject is not merely causally assisted by a free discursive environment; the autonomous self-legislating subject is constitutive of what it means for *Bildung* to take place at all. A 'practice' that systematically denies the student the standing to express, contest, and revise their own rational positions is not merely education performed sub-optimally — it is no more education in the relevant sense than an act of lending is performed by the person who secures a 'loan' with no intention to repay it. The free self-legislating subject that can voice and revise its own positions without fear of sanction or silencing is not merely conducive to education: it is its conceptual presupposition and its end, in precisely the way that trust in the sincerity of assertion is not merely conducive to the institution of promising but constitutive of it, such that its destruction would leave not a damaged practice of promising but the concept of promising itself logically annihilated. The move from 'conducive' to '*sine qua non*' is therefore not a matter of degree but of kind.

The third phase does not so much shift the conception of democracy as arrive at a less naïve, more philosophically adequate understanding of what democracy, as it is to be construed in itself and educationally, is. Democracy is understood here not as one ideology among others, not as a set of norms merely conducive to the fostering of autonomous acts, but as the meta-ideology: the formal framework that secures the very possibility of public reason, of ideological contestation, and of the free exercise of practical rationality itself — not itself one partisan preference among others, but the institutional structure within which preferences of all kinds may be articulated, examined, and subjected to rational scrutiny; not a value to be freely contested among others, but the formal condition of that free contestation as such. So understood, the democratic norms operative in education — civility in disagreement, the submission of contested claims to reason as ultimate arbiter, the illegitimacy of *ad hominem* argument, the right to dissent, the equal standing

of each interlocutor to speak and reply — are not something whose presence renders autonomous moral reasoning easier: they are that without which the relevant acts of autonomous reasoning are not merely impeded but logically impossible, the *sine qua non* of the practice whose conditions the preceding phases have shown to be quasi-transcendental.

Within this understanding, the central dilemma of Phase 3 may be stated in its essential form. A student, in the course of a seminar on ethics, gives voice to views whose denial of the rational standing of certain others operates at the level of illocutionary force rather than propositional content alone: it is not that the student *argues for* a position — which would be protected speech, however repugnant — but that the speech *enacts*, in the *hic et nunc* of the seminar, the denial of the equal standing of those it targets as rational interlocutors, and does so with a virulence and performative intent that forecloses rather than invites response. The student talks over those who would reply, directs targeted provocations — in speech and in gesture — at those whose standing as rational interlocutors they contest, deploys intimidating assertion whose purpose is not argument but the coercion of silence; the speech is, in its very mode of delivery, contemptuous of the right to reply, treating the anticipated rational response as already delegitimised, already beneath the threshold of what need be heard. Such speech is demeaning in a precise sense: it reduces those it targets to a standing beneath that of rational interlocutors, achieving by its illocutionary force the silence it would, if unchallenged, secure coercively. The specific content of such denial is variable across historical and social contexts, and the examples are easily multiplied; what is invariant is this logical and performative structure: the exploitation of the freedoms of open discourse precisely in order to foreclose the equal standing those freedoms guarantee. The effect is immediate and specific: not a general hostility to certain voices but a determinate act that forecloses, *here and now*, the capacity of particular students to participate as co-legislators in the shared rational practice the classroom is — at the precise moment their need to assert their right to express themselves autonomously is most acute.

The teacher must decide whether to intervene — to silence that student, in the name of protecting the autonomy and equal right of self-expression of the other students — or to abstain, and thereby ostensibly honour that student's exercise of autonomous free speech. The logic of the dilemma is structural and not historical; that a political climate *du jour* might embolden students

and authorise them to voice their denial of equal standing to others could serve to contextualise the scenario; but it serves only as context rather than as a necessary premise for our argument.

Two maxims, individually submitted to the universalisation test, appear coherent: a world of teachers committed to intervening when a student's speech would silence others, and a world of teachers committed to strict formal neutrality before that same situation, can each be conceived without immediate self-contradiction. The self-defeat emerges not from the maxim considered as a principle of individual action but only under the universalisation test — a single broken promise will not destroy the institution of promising; it is because the universalised adoption of the maxim of promise-breaking would in principle undermine the entire institution of promising that the singular act of promise-breaking is a CIC and therefore immoral. The same structure governs both horns of our dilemma.

The maxim of abstention runs: 'When a student's speech would deny others the equal standing to speak and reply, I will not silence them.' Universalised, this becomes the practical law of every teacher in every such case. Each student in the classroom is a rational agent capable of self-legislation — capable of giving voice to their own rational assessments, of contesting the views of others, of participating as a co-legislator in the shared discursive practice of moral reasoning. It is this capacity that moral education presupposes, is directed toward, and has as its very aim to cultivate. The student whose speech silences others — forecloses their responses, denies their equal standing to reply — enacts, within the space of the classroom, the annihilation of this capacity in those it targets. The teacher who refrains from intervening withholds institutional authority from the defence of those students whose capacity to speak and reply is being systematically denied; and if every teacher so refrains in every such case, the result is not a world of scrupulous formal neutrality but a world in which the capacity of the student as rational self-legislator is rendered inoperative for any student who might be subject to such encroachment. The universalised maxim of abstention thus contradicts itself in conception: it presupposes and purports to serve the cultivation of rational self-legislative agency while willing, in its universalisation, the destruction of the conditions under which that agency can be exercised. The structure is identical to the CIC of false promising, which draws upon and presupposes the institution of promising while willing, in its universalisation, the annihilation of the trust without which that institution cannot stand.

To be clear about the logical type of both violations — and here a potential misreading must be headed off at the outset — the wrong done to the silenced students is not a welfare wrong, not a failure of beneficence that might be compensated for by future goods or corrected by subsequent solicitude. One might suppose, following Donagan's reading of apparent duty-conflicts (*The Theory of Morality*, ch. 5), that the duty to intervene is grounded in the imperfect, positive duty of beneficence — the duty to promote the wellbeing or flourishing of those whose participation is impaired — and that, as an imperfect duty, it must yield mechanically to the perfect duty to respect the speaking student's autonomy, dissolving the dilemma asymmetrically. But this misdescribes the ground of the duty to intervene. The wrong done to the silenced students is the denial of their standing as rational self-legislative agents — the annihilation, in the *hic et nunc* of this discursive act, of their equal standing to participate as co-legislators in the shared practice of moral reasoning. This is a violation of autonomy, not of welfare; it is the universalisation test that identifies it as a CIC, not a CIW, placing both horns of the dilemma on the same logical footing. O'Neill's analysis of the parasitic structure of non-universalisable maxims makes precisely this vivid: the wrong is structural and logical, not empirical or compensable. Wood's strong reading of the Formula of Humanity (*Kant's Ethical Thought*, ch. 3) already grounds a duty to maintain conditions of others' rational agency — a reading sympathetic to the present argument — but does not itself generate the collision of two perfect duties that the scenario discloses.⁴

Yet the teacher who acts — who intervenes, silences the student, insists that certain speech is impermissible within the classroom — confronts the CIC from the opposite direction. Here the relevant violation is immediate and specific: it is the violation of that student's autonomy, concretely and individually. The teacher who silences a student determines for them — by authority rather than by argument — which views may be expressed and which must be suppressed. At this point, a response by reasoned argument — engaging the student's speech discursively rather than prohibiting it — might appear to constitute a third option, dissolving the binary structure of the dilemma. But such a response is precisely what the maxim of abstention would do: it allows the speech to remain in the discursive space, treating it as one contribution among others rather than as an act that forecloses the equal participation of those it targets; it withholds the institutional authority that alone is capable of restoring to those students the standing to reply that has been denied them. Reasoned engagement is, in the relevant sense, a form of non-intervention; the binary stands.

To silence the student is to treat them as incapable of rational self-legislation and as a means to a pedagogical or civic end. The maxim of intervention, universalised, becomes the law of all teachers in all such cases: every teacher silences every student whose views are judged inimical to the autonomy of others. What this universalisation destroys is the concept of the student as a rational agent capable of *Selbstdenken* — the very concept the activity of moral education presupposes and has as its end to cultivate. For if the teacher reserves the right to determine by authority which exercises of reason are admissible and which are to be suppressed, the classroom ceases to be a space of moral formation and becomes its negation: a domain of heteronomous determination in which the teacher's judgement displaces the autonomous reasoning it was the purpose of education to foster. Popper's paradox of tolerance here assumes its Kantian form: just as the tolerant society must, to survive, reserve the right to be intolerant of intolerance, the teacher who would preserve autonomy is, in this situation, compelled to violate it — the autonomy of the student who would destroy it — in order to preserve the conditions of autonomy for all. The compulsion is real, and so is the violation. Neither dissolves the other.

What has just been demonstrated at the level of the logical argument repeats, at the level of content, what Phase 2 showed at the level of form: the emboldened student's speech — its denial of others' equal right to contest — reproduces the structure of the *a-legal* founding act whose authority the norms could never derive from the democratic order they were in the process of founding. The founding violence that the norms' ordinary operation keeps invisible is here rendered visible; the originary exclusion that made the democratic framework possible resurfaces as the content of the act that contests it.

It might be supposed — following Herman's account of "deliberative priority" — that the educator's experience of conflict here reflects the appropriate phenomenology of an agent navigating the bounded deliberative space that the perfect duty to the speaker's autonomy delineates, rather than evidence of a collision within the moral law itself.⁵ On Herman's account, the perfect duty to respect the speaking student's autonomy forecloses certain responses and thereby constitutes a bounded deliberative space; within what remains, the duty to foster the discursive conditions of autonomous practice is exercised, and the tension the educator experiences is the phenomenology of navigating this constrained space — not a logical contradiction between two obligations at the same level. But to apply Herman's account to the present case would be to

commit what we might call, following Husserl, a psychologistic error: it would translate a logical, structural fact about the universalisation test into a description of the agent's deliberative phenomenology, as though the former could be read off from or established by the latter. The dilemma is constituted not by the subjective experience of the person undergoing it — subjective experience being neither a sufficient nor a necessary condition of the dilemma's reality — but by the nature of the duties' perfect and unconditional status, such that they simultaneously impose themselves upon the educator and may, under determinate conditions, be experienced as mutually excluding. Husserl's point in the *Prolegomena to the Logical Investigations* (§§17–19) is that logical relations — the necessity obtaining between premises and conclusion in a valid inference — are not psychological relations and cannot be established or disconfirmed by attending to the experience of an agent thinking through the inference. A person may think through a valid syllogism hesitatingly, confusedly, or with vivid phenomenological tension; none of this alters the logical relations inhering between the syllogism's elements. The analogy holds here: what is being demonstrated is not the phenomenology of the teacher's deliberative experience but the logical structure disclosed when the universalisation test is applied to each maxim. The demonstration is that the universalised maxim of abstention generates a CIC — it destroys the concept of the self-legislating student that *Bildung* requires; and that the universalised maxim of intervention equally generates a CIC — it destroys the concept of the student as a rational agent capable of *Selbstdenken*. Both violations are CICs. Herman's account describes how agents experience and manage competing moral considerations; it cannot show that one of the two universalised maxims fails to generate a CIC. To defeat the present argument, it would be necessary to show precisely this; Herman does not show it.

It should be noted, against a further potential misreading, that neither horn of the dilemma has its source in inclination, institutional command, or heteronomous pressure — neither in the phenomenology of a pathologically affected will, as Kant uses that term, torn between duty and institutional anxiety (*CPrR*, 5:20). Each obligation arises from the universalisation test applied to a determinate situation by a rational agent seeking to act from duty. The conflict is internal to pure practical reason itself.

We arrive, then, at the *Pflichtenkollision* Kant declared inconceivable. It is the collision of two perfect duties, each generating a CIC upon universalisation, each grounded in the same

concept of the student as a rational agent capable of moral self-legislation, and each incapable of being honoured without transgressing the other. Both horns of the dilemma proceed from a violation of autonomy — each from autonomy itself, as it appears under two descriptions the situation renders mutually exclusive: the autonomy of the student silenced by intervention, and the autonomy of those students whose equal standing to speak and reply is destroyed by abstention.

The link between Phase 3 and Phase 2 is decisive: the abstention horn is precisely the scenario of non-enforcement that Phase 2 identified as the condition under which the founding imposition would retroactively transform from a CIW into a CIC. What Phase 3 adds is not merely an intensification of that scenario but its full clarification: the democratic norms of the classroom always were the *sine qua non* of autonomous educational practice; Phase 2's characterisation of them as merely conducive was insufficiently adequate to what the CIC analysis itself requires. The moral law, which was to speak with a single categorical voice, here issues contradictory commands; and practical reason, which was designed to dissolve all apparent *Pflichtenkollisionen* by exposing them as confusions between duties of unequal standing, discovers within itself the very antinomy it was constructed to preclude. The argument proceeds, throughout, entirely from within the Kantian system: it requires no appeal to consequences, no utilitarian supplementation, no external critique.

The two duties in collision are, in Kant's terminology, internal rather than external perfect duties: they are not juridical obligations susceptible to external lawgiving or institutional sanction (*MM* 6:224), but duties that live and are decided in conscience alone. This is philosophically significant in a further respect developed in the Conclusion: what Kant treats as a legal conflict when it arises between institutions becomes an irresolvable ethical dilemma when it arises within a single rational agent's conscience. The pedagogical case shows precisely this: the teacher does not face a collision between two institutional rights but between two internal perfect duties — duties of conscience, non-coercible from outside, living in the will alone. By recasting the faculty's dilemma as a legal conflict rather than an ethical collision, Kant deploys the internal/external perfect duty distinction as a further mechanism of containment: the philosophical faculty's duty to reason freely is reframed as an external right subject to legal qualification, rather than an internal perfect duty subject only to reason's own demands. This reframing is itself an expression of the

founding exclusion: the conditions of autonomy's exercise are assigned to the domain of external right and institutional negotiation rather than to the domain of ethics.

It is worth registering, before proceeding, the scope of the claim. The pedagogical dilemma is not anomalous but exemplary: it is one instance of a principle that the structural fault line in Kant's system generates wherever the categorical imperative is applied to situations in which the conditions of autonomy's exercise are themselves at stake. The teacher's dilemma brings this structure to light in its purest and most philosophically transparent form; but the principle, once excavated, is reproducible across any domain in which a practice has autonomy both as its presupposition and as its end, and in which the conditions of that autonomy's exercise are intersubjective and institutional. The judge in a democratic legal system faces a structurally identical collision when the law they are bound to apply categorically violates the dignity of the person before them. The journalist's duty to report truth conflicts with the perfect duty to protect the life of a source whose exposure would follow from that report. The legislator faces a genuine collision when the law demanded by their constituency would systematically exclude the rational standing of a minority. Each of these dilemmas has the same logical structure: two perfect duties, each generating a CIC upon universalisation, each grounded in the same concept of rational agency, rendered mutually exclusive by a determinate situation. The Kantian preclusion of *Pflichtenkollision* does not merely fail here but fails structurally, wherever reason is forced to reckon with the conditions of its own realisation.

4. Conclusion

I. Dilemma and the Collapse of Distinctions

That Kant should declare a conflict of duties inconceivable is not an incidental dogma: it is the keystone of the entire architectonic of pure practical reason. To admit genuine dilemmas would be to allow that reason can bind the will to contradictory necessities — to concede that the moral law, which was to be the form of reason's self-consistency, is capable of self-contradiction. The purity of the moral law depends upon the purity of the distinctions that support it, and these distinctions form a system: between form and matter; autonomy and heteronomy; law and condition; perfect and imperfect duty; contradiction-in-conception (CIC) and contradiction-in-will

(CIW); conditional and unconditional; categorical and hypothetical imperative. These distinctions do not merely organise the moral domain; they constitute it.

What must now be made explicit is that these distinctions do not merely stand together in solidarity — they each perform the same exclusion. The form/matter distinction secures the purity of the moral law by excluding from its determining ground all material conditions, all heteronomous ends. The autonomy/heteronomy distinction secures the will's self-legislation by excluding the intersubjective and institutional conditions within which self-legislation is always exercised. The perfect/imperfect duty distinction secures the unconditional force of categorical commands by excluding the conditions of their realisation from the domain of perfect obligation. The CIC/CIW distinction secures the logical absoluteness of the prohibition by restricting CIC-status to those maxims that destroy their conditions in the absence of any material, empirical mediation — thereby excluding from the domain of the strictly impossible any maxim whose self-defeat is mediated through institutional or social conditions. The categorical/hypothetical distinction secures the independence of the moral law from empirical ends — by excluding those ends, along with the conditions that answer to them, from the law's own domain. Each distinction performs a variant of the same founding exclusion: the exclusion of the quasi-transcendental conditions of autonomy's realisation from the domain of pure practical reason. This shared structure of exclusion gives the system its characteristic unity. And it is this shared structure that explains why the demonstration of impurity in one propagates through all: to demonstrate the permeability of the one is to lay bare the illegitimacy of the exclusion as such — an exclusion performed, repeatedly and with different tools, at every level of the moral architectonic. The system holds together or falls together not because of some accidental interconnection but because each distinction is a re-inscription of the same exclusion.

And fall together they do, once the pedagogical dilemma asserts itself. The duty to autonomy, which was to be purely formal, discovers that it cannot be satisfied without attending to the conditions of autonomy's exercise: the distinction between form and matter collapses. The perfect duty, which was to command unconditionally, discovers its status as perfect or imperfect to be conditional upon the ongoing fulfilment of further duties — a conditionality of the unconditional that Kant's system cannot countenance: the distinction between the conditional and the unconditional collapses. The CIC bifurcates: the maxim of abstention, universalised, logically

destroys the institutional conditions presupposed by the very practice it concerns — precisely as universalised false promising destroys the institution of promising — while the maxim of intervention, universalised, destroys the concept of the student as a rational agent capable of self-legislation: the distinction between CIC and CIW collapses. And autonomy itself proves, in the intersubjective and institutional domain in which it is always already exercised, the ground of two contradictory categorical demands: the distinction between autonomy and heteronomy collapses. Each distinction implicates and sustains the others; when one fails, all fail. But the failure is not simply collapse: it is contamination. The excluded, subordinated element does not merely re-enter the system as an equal; it comes to contaminate the purity of the hierarchical term it was excluded to protect. The external end becomes the necessary condition of possibility of the act the moral law was designed to govern; the imperfect duty imposes unconditionally, because the conditions of possibility it concerns are not simply empirical; the CIW becomes a CIC, because the conditions are not merely useful but constitutive. What was excluded as matter irrupts into form; what was conditional binds categorically; what admitted latitude admits none.

The existence of the dilemma does not render these distinctions meaningless; it renders them impure. It reveals them as semipermeable rather than impermeable: what they exclude continues, as Derrida would say, to seep through as the supplement — at once secondary and originary, at once excluded and indispensable. The heteronomous conditions of autonomy are the supplement of the moral law's formal purity: secondary, in that they derive from the empirical domain the law transcends; originary, in that without them the law has no world in which to appear. To suppress them, as Kant's system requires, is to found the law upon the repression of its own conditions of possibility. The dilemma is therefore not an anomaly that disrupts the Kantian edifice from without; it is the fissure that runs through it from within — the trace of its constitutive impurity, the proof that the system can be maintained only by excluding what it cannot do without.

The philosophical literature on moral dilemmas provides a taxonomy that illuminates the structure of what has been demonstrated. Walter Sinnott-Armstrong (*Moral Dilemmas*, Blackwell, 1988) distinguishes three types. A symmetrical dilemma is one in which neither obligation overrides the other; both leave a moral residue — Bernard Williams's term for the remainder of the overridden obligation that persists as grounds for regret, apology, or reparation. Sartre's famous case of the student torn between loyalty to his ailing mother and the call of the Resistance

approaches this type: neither claim obviously overrides, neither disappears when the other is acted upon. An asymmetrical dilemma is one in which one obligation does override but the other nonetheless leaves a residue. The doctor who breaks medical confidentiality to prevent serious harm acts rightly — the duty to prevent harm overrides — but remains in the position of having violated a genuine obligation; the confidentiality that was breached was not simply voided by the override, and the doctor has reasons for apology and explanation that would be unintelligible if the overridden duty had simply ceased to exist. An absolute dilemma is one in which neither obligation can be satisfied and neither leaves the agent without moral remainder — the most extreme form, in which there is no 'right answer' that leaves nothing morally unresolved. The paradigm case is Sophie's choice: whatever choice Sophie makes, she has not merely made a difficult decision but has incurred a genuine moral violation, and no retrospective rationalisation of the outcome as 'the best available' dissolves the moral horror of the act. The teacher's dilemma in Phase 3 approaches this absolute type: whichever choice the teacher makes — intervene or abstain — a genuine violation of autonomy has occurred, a genuine CIC has been enacted, and the moral residue cannot be discharged by the rightness of the choice made. Ruth Barcan Marcus (*Moral Dilemmas and Consistency*, *Journal of Philosophy*, 77, 1980) provides the deontic-logical framework within which all three types can be accommodated without conceding inconsistency in the moral system: a moral code is consistent if its principles do not logically contradict one another; a dilemma arises not from inconsistency in the code but from the fact that circumstances can make it impossible for all obligations to be simultaneously satisfied. The Kantian precondition of the argument — that a genuine dilemma entails a contradiction in the moral law itself — is precisely what Marcus's analysis denies; and the argument of this paper shows that this denial is vindicated from within the Kantian system itself.

II. Legal Conflicts and the Threshold of Ethical Dilemmas

That this structural impurity was not entirely invisible to Kant himself becomes apparent in *The Conflict of the Faculties*, where he concedes that the pursuit of truth within the university may give rise to legitimate legal conflicts — conflicts in which both parties act under a form of right, neither reducible to the other. The distinction between legal and illegal conflicts in that text is precise: a legal conflict is one in which each party acts within the scope of its legitimate authority, even as those authorities collide; an illegal conflict is one in which a party exceeds the

limits of its licensed domain, as when the philosophical faculty carries its corrective judgements upon the higher faculties into the public sphere — an act which transgresses the terms under which its autonomy was conferred. Kant knew something of this from experience: the rescript of King Friedrich Wilhelm II (1794) censuring him for his theological writings in *Religion within the Limits of Reason Alone* furnished a real-world instance of exactly the conflict he theorised.

These legal conflicts are not ethical dilemmas in the Kantian sense: they do not generate two equally binding duties that logically exclude one another within the same rational agent's conscience. They are conflicts between institutional rights and powers — external, juridical, coercible — that Kant can contain by assigning them to the domain of right rather than morality, and by drawing a legal threshold (the academy wall) beyond which one party's claim becomes illegitimate. Kant almost allows the ethical dilemma here — the structural parallel is unmistakable — but does not quite, because the legal/ethical distinction does the work of containment: what would be an irresolvable collision of duties in the domain of pure practical reason is recast as an institutional conflict between rights and powers, admitting of a juridical resolution. What Kant treats as a legal conflict when it arises between institutions becomes an irresolvable ethical dilemma when it arises within a single rational agent's conscience. The pedagogical case shows precisely this: the teacher does not face a collision between two institutional rights but between two internal perfect duties — duties of conscience, non-coercible from outside, living in the will alone. By recasting the faculty's dilemma as a legal conflict rather than an ethical collision, Kant deploys the internal/external perfect duty distinction as a further mechanism of containment: the philosophical faculty's duty to reason freely is reframed as an external right subject to legal qualification, rather than an internal perfect duty subject only to reason's own demands. This reframing is itself an expression of the founding exclusion: the conditions of autonomy's exercise are assigned to the domain of external right and institutional negotiation rather than to the domain of ethics. The structural inevitability and recurrence of these legal conflicts — their non-anomalous character, their inherence in the very arrangement Kant theorises — mirrors a reason in which conflict is not accidental but constitutive.

Derrida's reading renders this symptom legible. The philosophical faculty, which was to stand for the autonomy of reason, is constituted by the very heteronomy it disavows: its freedom of inquiry is both bestowed and constrained by the state that lies beyond it, an invaginated pocket

in which the law of truth and the law of power — and here the double genitive is deliberate: the power that the law exercises (*potestas legis*) and the law that power imposes upon reason's claims — coincide and clash. The founding act by which the university's autonomy is instituted cannot itself be authorised by the order of legality it founds; it is, as Derrida argues in *Mochlos, a-legal* — neither simply legal nor simply illegal, but that which precedes and makes possible the very distinction between the two. What Kant acknowledges in political philosophy — that the conditions upon which the autonomous exercise of reason depends are heteronomous in part — is precisely what his ethics suppresses.

Kant's exclusion of legal conflicts from the ethical domain rests on the same assumption as the *Pflichtenkollision* preclusion: that the ethical domain is self-sufficient and internally consistent, immune to the contingencies that generate collisions in the juridico-political sphere. But this assumption is precisely what the present paper contests. If two perfect duties — both deriving from the single concept of autonomy — can collide under determinate conditions, the ethical domain is not self-sufficient, and the principled distinction between ethical and legal conflict loses its philosophical ground. Kant's classification of a conflict as legal rather than ethical is made from the perspective of the architectonic; it does not alter the objective moral structure of the situation the agent confronts. The teacher who must decide whether to silence a student whose speech would destroy the discursive conditions on which others' autonomy depends faces two claims, each grounded in a perfect duty, each deriving from the concept of autonomy, and mutually exclusive under the conditions that obtain. Whether the architectonic assigns this conflict to the domain of right or the domain of ethics does not alter this structure: the conflict is ethical not because it is experienced as such but because both its terms are generated by the moral law's own requirements.

That the educator experiences this structure as an irresolvable demand — that the resolution in favour of one duty leaves intact a sense of genuine obligation to the other, a remainder that is not confusion but moral attentiveness to what could not be honoured — is not what constitutes the dilemma but what the dilemma, objectively constituted, appears as within the consciousness of the practical reasoner, this appearance having its own phenomenological descriptive character, as it confronts what it cannot adjudicate. To invoke this experiential dimension is not to make the psychologistic error of grounding the dilemma's reality in subjective feeling; it is to enter

phenomenology's proper opening: the articulation of experience as responsive to, and disclosive of, a logical structure that exceeds and precedes it. The legal/ethical distinction, on this reading, is not the containment of the problem but one of its symptoms: another site at which the *CPrR*'s claim to the ethical domain's self-sufficiency reproduces the suppression the *Pflichtenkollision* preclusion enacts.

III. From the Highest Good to the Antinomy Within Duty

Kant was not unaware of conflict within the moral domain; he simply located it elsewhere. In the *Critique of Practical Reason*, he identifies practical reason's antinomy as the disparity between virtue and happiness — between the moral worth of the will and the contingent configuration of a world in which happiness is not proportioned to merit. The *summum bonum*, the highest good (*das höchste Gut*), is the complete object of pure practical reason: virtue as its supreme condition, happiness as its necessary complement (5:110–111). The highest good is crucially the object of a pure will, not its determining ground: as Kant explicitly insists, 'though the highest good may be the whole object of a pure practical reason, that is, of a pure will, it is not on that account to be taken as its determining ground, and the moral law alone must be viewed as the ground for making the highest good and its realization or promotion the object' (*CPrR*, 5:109–110). The realisation of the highest good is therefore not named as a duty. The conditions of its realisation are assigned instead to the postulates of God and immortality, which function as guarantees of rational faith (*Vernunftglaube*) rather than as demands of practical reason.

Reath's 'Two Conceptions of the Highest Good in Kant' (*Journal of the History of Philosophy*, 26, 1988) identifies two competing readings: the comprehensive conception, which treats the highest good as the ideal total state in which virtue and proportioned happiness are jointly realised — a regulative idea projected by reason; and the achievement conception, which treats the highest good as the object of a duty to bring it about in the world. Kant tends toward the comprehensive conception: the highest good is the ideal toward which reason orients the will, but the conditions of its realisation are not themselves named as duties. This tendency is not incidental but systematic — it is the same founding gesture, performed in theological register, as the preclusion of *Pflichtenkollision*. In both cases, the quasi-transcendental conditions of autonomy's realisation are excluded from the domain of perfect duty by being assigned to a secondary domain

— rational faith, theology, imperfect duty — where they cannot generate genuine collision within the categorical structure. The postulates of God and immortality do for the highest good what the hierarchy of perfect over imperfect duties does for the pedagogical dilemma: they defer the conditions of realisation into a domain where they cannot press back against the formal purity of the moral law.

The highest good, the conferred autonomy of the philosophical faculty, and the preclusion of *Pflichtenkollision* are three expressions of the same founding exclusion — not merely parallel instances but repeated performances of the same move: securing the formal purity of the moral law by assigning its conditions of realisation to a domain where they cannot generate binding obligations within the categorical structure. And since each of the distinctions within that system — form/matter, autonomy/heteronomy, perfect/imperfect, CIC/CIW, conditional/unconditional, categorical/hypothetical — is a re-inscription of this same exclusion, the three expressions are also one: variants of the single structural gesture by which reason preserves its purity at the cost of what it cannot do without. The teacher's dilemma is what happens when this exclusion fails — when the conditions that were to be contained as facts of institutional life or postulated objects of rational faith return, insistently and immediately, as the grounds of two simultaneously binding and mutually exclusive perfect duties.

The true conflict within practical reason lies not between virtue and the recalcitrant world but within duty itself — within the categorical demand of a moral law that, as soon as it enters the world in which it must be exercised, encounters its own shadow: the necessity of choosing between equally binding imperatives. Kant's formalism avoids this internal conflict by purifying duty of all empirical conditions, by insisting that the moral law legislate for a will abstracted from its worldly situation. But the teacher's dilemma demonstrates that no such abstraction is available: the formal law, in the very moment of its application, is implicated in material, institutional, and political conditions that are not given by reason but within which reason's practical demands find their only possible enactment. The antinomy is not between reason and the world but within reason itself, once it is forced to reckon with the conditions of its own realisation.

What the doctrine of the highest good reveals, and what the pedagogical dilemma confirms, is that pure practical reason cannot remain content with the mere form of law: it must also will the actuality of a world in which the law can be obeyed. The demand that the highest good be realisable

is, as Kant's own argument in the second *Critique* concedes, internal to reason's demand upon itself. And yet to will the realisability of the good is already to introduce what the law was designed to exclude. The very thing the categorical imperative sought to banish — the empirical end, the longing that what ought to be come to be — returns, within the doctrine of the highest good, as reason's own unrelinquishable aspiration. Kant assigns this aspiration not to duty proper but to rational hope (*Hoffnung*) — to the postulates of God and immortality, whose function is precisely to guarantee what duty demands but cannot, from within itself, secure. The formal purity of the moral law is found to depend, for its very coherence, upon a demand for worldly realisation it cannot acknowledge as a duty without conceding the possibility of the dilemmas it has denied.

IV. Return to the Initial Dilemma

We return, then, to the paradox from which this inquiry began. The teacher who would refrain from imparting specific values at the very moment they are most needed; who would withdraw from effective pedagogical action in the name of an autonomy whose conditions are being dismantled around them; who would confine themselves to a formally neutral presentation of moral theory while the social and political infrastructure within which those theories can be freely contested is being systematically eroded — this teacher inhabits, in microcosm, the predicament of practical reason itself. The demand that the moral law be effective in the world and the demand that the will legislate for itself independently of the world cannot, in the moment of their pedagogical realisation, be simultaneously honoured.

The initial dilemma is thus neither a contingent pedagogical difficulty nor an anomaly in Kantian ethics but the site at which a structural tension internal to the *Critique of Practical Reason* becomes demonstrable. In the *Critique of Pure Reason*, what in theoretical reason was identified — candidly and with unerring self-awareness — as its constitutive limitation, in need of being named and curtailed by critique: when pure theoretical reason attempts to apply its principles beyond the conditions of possible experience, it generates Antinomies — genuine contradictions, each side equally provable, resolvable not by refuting either but by showing that both presuppose a transcendental illusion. The *Critique of Practical Reason* faces an analogous problem — whether practical reason, in applying its principles to action in the world, can generate contradictions — and resolves it by fiat: the *Pflichtenkollision* preclusion declares collision impossible, or illusory

— an illusion to be rectified by the correct specification of the perfect and imperfect duties involved — rather than diagnosing its source. The asymmetry between the two *Critiques* is not an achievement but a symptom: what in theoretical reason was identified as its constitutive limitation, practical reason suppresses, distributing its consequences into the doctrine of the postulates and the highest good — where, far from resolving the problem, Kant has recourse to the guarantees of God and immortality in such a way that the very question of *Pflichtenkollision* is made to fade from view, rather than faced as what it is: the practical antinomy pure practical reason necessarily generates when it applies its legislative principle to a conditioned world of finite agents (*CPrR*, AA 5:110–114). No postulate reconstitutes, from outside the situation, the conditions of autonomous reasoning that are being destroyed within it; and no prospect of future proportionality between virtue and happiness resolves the dilemma of the educator who must choose, in this classroom, this semester, whether to silence the student whose speech forecloses the autonomy of others, or to abstain and thereby perhaps abandon those conditions to their erosion.

A practical reason that wills the moral law must will the conditions of its realisation; to will those conditions is already to be subject to the collision of duties their quasi-transcendental status generates. Critique, on this reading, acquires a new function in the practical domain: not only the deduction of the moral law's authority, which Kant accomplishes, but the mapping of the dilemmas that authority generates when it encounters the quasi-transcendental conditions of its own exercise. The *CPrR* cannot acknowledge this without conceding that pure practical reason is not self-sufficient — that it requires, for its exercise in the world, conditions it cannot itself guarantee and duties it cannot itself adjudicate. This is the practical antinomy the *Pflichtenkollision* preclusion suppresses and the pedagogical dilemma discloses.

The ethics classroom is the modest but philosophically precise site of this disclosure. It is the educator — or anyone who finds themselves at the fault line between the good and the conditions of its realisation, between the *summum bonum* as rational ideal and the highest good as something that must be achieved in and through a world that reason cannot fully command — who inhabits this dilemma most thoroughly and most authentically, who discovers in its unresolvable pressure not merely a pedagogical predicament but the limit at which the *Critique of Practical Reason* must itself be re-read: not as the demonstration of pure practical reason's self-sufficiency but as the involuntary disclosure of its constitutive exposure to the conditions of a world it cannot

command. The educator's double bind — to respect the formal autonomy of the student and to preserve the quasi-transcendental conditions without which that autonomy cannot be exercised in any actual world — is not a failure of practical reasoning but its moment of genuine self-disclosure: the point at which practical reason, in order to be consistent with itself, must will not only the moral law's formal validity but the conditions of possibility in which alone that law can be realised, and discovers in those conditions the dilemmas it had declared inconceivable. The law's claim to govern conduct in the world means it cannot be wholly insulated from the conditions under which such law-adhering and law-expressive conduct in that world is possible.⁷

1 The same structure can be formulated as a double-standard or self-exemption: the agent who makes a false promise relies on a norm of trustworthiness that their own act, if universalised, would destroy — treating their own conduct by one rule and others' conduct by another. Cf. Marcus G. Singer, *Generalization in Ethics* (Knopf, 1961); Christine Korsgaard, 'Kant's Formula of Universal Law,' *Pacific Philosophical Quarterly* 66 (1985), pp. 24–47.

2 See Jacques Derrida, 'Mochlos, or the *Conflict of the Faculties*,' in Richard Rand, ed., *Logomachia: The Conflict of the Faculties* (University of Nebraska Press, 1992), pp. 1–34; and 'The University in the Eyes of its Pupils,' *Diacritics* 13 (1983), pp. 3–20. On iterability and the impossibility of containing the addressee: Derrida, 'Signature Event Context,' in *Limited Inc* (Northwestern University Press, 1988), pp. 1–24.

3 Cf. *The Conflict of the Faculties*, AA 7:18–27. The 1794 rescript from Friedrich Wilhelm II to Kant provides the historical correlate: a legal conflict brought to the threshold of illegality by the public circulation of philosophical judgement.

4 Allen W. Wood, *Kant's Ethical Thought* (Cambridge University Press, 1999), ch. 3. Wood's reading of the Formula of Humanity generates duties to maintain conditions of others' rational agency; it is, in this respect, sympathetic to the present argument. But Wood does not draw the conclusion that these conditions can generate a collision of two perfect duties.

5 Barbara Herman, 'Obligation and Performance: A Kantian Account of Moral Conflict,' in Owen Flanagan and Amélie Oksenberg Rorty, eds., *Identity, Character and Morality* (MIT Press, 1990), pp. 311–338. Herman's concept of deliberative priority holds that perfect duties bound the space within which imperfect duties operate, such that experienced moral tension reflects the agent's navigation of this bounded space rather than a logical collision within the law. Cf. Edmund Husserl, *Logical Investigations*, trans. J. N. Findlay (Routledge, 2001), Prolegomena to Pure Logic, §§17–19.

⁶ On Hegel's empty formalism objection and the deposit example, see G. W. F. Hegel, *Philosophy of Right*, trans. T. M. Knox (Oxford University Press, 1952), §§129–141, and §135 specifically. The question of whether *Sittlichkeit* dissolves or relocates the practical antinomy identified here — and the further question of what the *CPrR* suppresses in distributing the conditions of practical reason's realisation across the postulates and the doctrine of the highest good — is developed in [Author], '[Title]' (companion essay). The proximity of the present argument to Hegel's diagnosis should not obscure the difference in conclusion: Hegel treats the worldly dependence as grounds for abandoning Kantian formalism; the present paper treats it as grounds for demonstrating an irresolvable collision within it. On the *Metaphysics of Morals* doctrine of right and the civil condition, see MM 6:230–256 (Gregor trans.).

⁷ The structure identified here — a formal principle constitutively dependent on the conditions it formally excludes — has structural analogues in Derrida's analyses of the gift, forgiveness, and unconditional hospitality. These analyses are best understood as horizon-intentionality analyses in the precise phenomenological sense: they attend to the full *noetico-noematic* structure of these acts including their institutional horizon, and disclose that the act's intelligibility as the act it is depends on conditions the act cannot itself generate or command. In each case, what Derrida's phenomenological method discloses is the same quasi-transcendental logic: the act defined by its independence from a circuit of exchange or worldly condition that proves constitutive of the act's very possibility.

See Jacques Derrida, *Given Time: I. Counterfeit Money*, trans. Peggy Kamuf (University of Chicago Press, 1992), pp. 7–30; *The Gift of Death*, trans. David Wills (University of Chicago Press, 1995); *On Forgiveness*, in *On Cosmopolitanism and Forgiveness*, trans. Mark Dooley and Michael Hughes (Routledge, 2001); *Of Hospitality*, trans. Rachel Bowlby (Stanford University Press, 2000). That Derrida’s phenomenological method, in contrast to Kant’s formalist and rationalist derivation of obligation — which operates at a level of abstraction at which worldly conditions structurally cannot appear — is what discloses the quasi-transcendental conditions Kant’s own procedure uses but cannot acknowledge as such, is argued in the companion essay cited in fn.⁶.

⁸ The failure of the opening gambit begins, it should be noted, earlier than the CIC demonstration requires: at the level of the will’s own intentional constitution, prior to any universalisation test. The *Groundwork*’s stripping procedure — progressively abstracting from inclination, from empirical content, from all material determination — operates at the level of the concept of rational willing. But willing is intentionally structured as such: every act of willing has a *noetico-noematic* structure, a mode of intending and a content intended, and the formal act of willing is already a directedness-toward something. One cannot arrive, by progressive abstraction, at a willing evacuated of its intentional arc: at that point one has not a purer will but no will — a will with no willing. The noematic content of a specific promise (what is promised, to whom, under what normative structure of accountability) and the noetic quality of the act (its character as binding, future-directed, addressed-to-another) are not separable from the horizon within which promising constitutes a distinct act-type. The quasi-transcendental conditions — the institution of promising, the convention of trust, the intersubjective structure of address and accountability — are not empirical facts merely accompanying the act from outside, but horizon-constitutive of its *noetico-noematic* structure: the background of sense within which both the mode of intending and the specific content acquire their promissory character rather than the character of prediction, assertion, or aspiration. They are simultaneously worldly and institutionally operative (existing as intersubjective, contingently maintained structures in the lifeworld) and constitutive of the act’s own intentional fabric; these are complementary descriptions at different levels, not competing ones. A related point bears on the false-promising case specifically. One might suppose that the false promiser intends in the absence of the horizon-intentionality of genuine promising — performing a mere simulacrum, a hollow form. The opposite is the case. The false promiser’s act depends no less than the genuine promiser’s on the entire apparatus of promising: the convention of trust, the address structure, the addressee’s constitution as one whose expectations will be bound, the institution of accountability within which the false promise secures its effect. Without the horizon-intentionality of promising operative in both parties — in the false promiser’s own intention and in the one to whom the promise is made, these intersecting within the shared institutional lifeworld — the money could not be secured. The lack of sincerity does not rob the act of its horizon-intentionality; it requires and exploits it. The parasitism is horizon-intentional through and through, which is precisely why its universalisation destroys what it necessarily presupposes. This horizon-constitutive structure belongs to what Husserl calls the lifeworld (*Lebenswelt*) — the always-already-operative background of sense that precedes and envelops individual acts of willing and theorising alike. We find ourselves always already within such horizon-frameworks; their repeatability and intersubjective availability precede and exceed any particular will’s act. The stripping procedure is not a return to a prior horizon-free state of the will; it is an abstraction performed from within an already horizon-structured practical field, producing not the will prior to conditions but the will with its conditions suppressed at the level of analysis. The conditions are not behind the will; they are beneath the level at which the analysis operates. The iterability of the horizon is inseparable from its availability as horizon: for a horizon to function as a horizon, the sense it makes available must be capable of being taken up across different contexts, by different subjects, in different temporal situations. The horizon must be, in this precise sense, repeatable. The repeatable structure of the horizon is the institution, and the institution is the iterable structure of the horizon. It is this structure that connects the present argument to Derrida’s account of the sign in ‘Signature Event Context’: for a sign to function as a sign — and for a promise to function as a promise — it must be capable of operating in the absence of any particular sender, recipient, or originating context. Its availability to any given act of intending is a function of its in-principle iterability across contexts; its iterability a function of its perennial intersubjective availability. This is the institution of the sign, and it is the same structure as the institutional horizon of promising. On the Husserlian texts: the horizon-structure is developed most fully in *Cartesian Meditations*, trans. Dorothea Cairns (Martinus Nijhoff, 1960), §§19–22 (inner and outer horizon) and §§43–62 (intersubjectivity and the constitution of a shared world); and in *The Crisis of European Sciences and Transcendental Phenomenology*, trans. David Carr (Northwestern University Press, 1970), §§28–37 (the lifeworld as pre-given horizon of all theoretical activity). The question of the Kantian will’s status — whether it is one intentional act among others (as Husserl’s taxonomy of intentional acts would suggest) or the transcendental condition of any morally significant act (as Kant holds) — is itself philosophically symptomatic: Kant’s move to the will as transcendental condition rather than intentional act is the founding gesture that places the will above the level at which its intentional structure, and the

horizon-constitutive conditions that structure belongs to, can be analysed. The critique applies not to the will as abstract formal principle but to its exercise in concrete acts of willing, and it is there that the horizon-intentional structure makes its claim.

Keywords

Pflichtenkollision; moral dilemma; Kantian ethics; perfect duty; contradiction-in-conception; categorical imperative; moral pedagogy; practical reason; immanent critique; quasi-transcendental

Bibliography

- Baron, Marcia. *Kantian Ethics Almost Without Apology*. Ithaca: Cornell University Press, 1995.
- Derrida, Jacques. 'Force of Law: The Mystical Foundation of Authority.' In *Acts of Religion*, edited by Gil Anidjar, 228–298. London: Routledge, 2002.
- Derrida, Jacques. *Introduction to Husserl's Origin of Geometry*. Translated by John P. Leavey Jr. Lincoln: University of Nebraska Press, 1989.
- Derrida, Jacques. 'Mochlos, or the *Conflict of the Faculties*.' In *Logomachia: The Conflict of the Faculties*, edited by Richard Rand, 1–34. Lincoln: University of Nebraska Press, 1992.
- Derrida, Jacques. 'Signature Event Context.' In *Limited Inc*, 1–24. Evanston: Northwestern University Press, 1988.
- Donagan, Alan. *The Theory of Morality*. Chicago: University of Chicago Press, 1977.
- Gowans, Christopher, ed. *Moral Dilemmas*. Oxford: Oxford University Press, 1987.
- Herman, Barbara. 'Obligation and Performance: A Kantian Account of Moral Conflict.' In *Identity, Character and Morality*, edited by Owen Flanagan and Amélie Oksenberg Rorty, 311–338. Cambridge, MA: MIT Press, 1990.
- Husserl, Edmund. *Cartesian Meditations: An Introduction to Phenomenology*. Translated by Dorothea Cairns. The Hague: Martinus Nijhoff, 1960.
- Husserl, Edmund. *The Crisis of European Sciences and Transcendental Phenomenology*. Translated by David Carr. Evanston: Northwestern University Press, 1970.
- Husserl, Edmund. *Logical Investigations*. Translated by J. N. Findlay. 2 vols. London: Routledge, 2001.
- Kant, Immanuel. *Critique of Practical Reason*. Translated by Mary Gregor. In *Practical Philosophy*, 133–272. Cambridge: Cambridge University Press, 1996.
- Kant, Immanuel. *Groundwork of the Metaphysics of Morals*. Translated by Mary Gregor. In *Practical Philosophy*, 37–108. Cambridge: Cambridge University Press, 1996.
- Kant, Immanuel. *The Metaphysics of Morals*. Translated by Mary Gregor. Cambridge: Cambridge University Press, 1996.
- Kant, Immanuel. *The Conflict of the Faculties*. Translated by Mary J. Gregor and Robert Anchor. In *Religion and Rational Theology*, edited by Allen W. Wood and George di Giovanni, 233–328. Cambridge: Cambridge University Press, 1996.
- Korsgaard, Christine M. *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press, 1996.

- Korsgaard, Christine M. 'Kant's Formula of Universal Law.' *Pacific Philosophical Quarterly* 66 (1985): 24–47.
- Marcus, Ruth Barcan. 'Moral Dilemmas and Consistency.' *Journal of Philosophy* 77, no. 3 (1980): 121–136.
- O'Neill, Onora. *Constructions of Reason: Explorations of Kant's Practical Philosophy*. Cambridge: Cambridge University Press, 1989.
- O'Neill, Onora. 'The Public Use of Reason.' *Political Theory* 14, no. 4 (1986): [pp. —].
- Reath, Andrews. 'Two Conceptions of the Highest Good in Kant.' *Journal of the History of Philosophy* 26, no. 4 (1988): 593–619.
- Singer, Marcus G. *Generalization in Ethics*. New York: Knopf, 1961.
- Sinnott-Armstrong, Walter. *Moral Dilemmas*. Oxford: Blackwell, 1988.
- Williams, Bernard. 'Ethical Consistency.' In *Problems of the Self*, 166–186. Cambridge: Cambridge University Press, 1973.
- Wood, Allen W. *Kant's Ethical Thought*. Cambridge: Cambridge University Press, 1999.