THE SOVEREIGN'S MASK

Performed Irrationality, Coercive Bargaining, and the Ethics of Silence in the Theory of Sovereign Power

A Philosophical and Jurimetric Inquiry into the Madman Doctrine from Machiavelli to Schelling

By

ISAAC CHRISTOPHER LUBOGO

THE SOVEREIGN'S MASK

Performed Irrationality, Coercive Bargaining, and the Ethics of Silence in the Theory of Sovereign Power

A Philosophical and Jurimetric Inquiry into the Madman Doctrine from Machiavelli to Schelling

By
Isaac Christopher Lubogo

Suigeneris Publishers
Kampala

A Note on This Book

This book unfolds in six parts. Part One states the problem of performed unpredictability as a bargaining phenomenon. Part Two traces its classical foundations in Machiavelli and Hobbes. Part Three develops its rigorous modern statement in Thomas Schelling's theory of coercive bargaining. Part Four examines the doctrine's special relevance to succession and inheritance. Part Five turns to the ethics and phenomenology of silence as a response to it, in dialogue with both Stoic and Ubuntu philosophical resources. Part Six closes with a general synthesis and its jurisprudential implications.

The argument throughout is structural and historical. It is not, and is not intended as, a psychological account of any living person.

Table of Contents

Preface

This book begins from a discomfort. Anyone who has watched a political actor — a minister, a general, a claimant to a throne — say something so excessive that it cannot possibly be calculated, has likely had two thoughts in immediate succession: first, that the man has lost control of himself; second, a colder thought arriving a moment later, that perhaps he has not.

The two readings cannot both be the most charitable account of the same act. Either the speaker has lost his grip on the instrument he commands, or he has found in that instrument something the disciplined exercise of power cannot supply: the credibility that only an unconstrained will can purchase. This book takes the second reading seriously — not as an apology for the first, but as a long-neglected branch of political philosophy, running from Machiavelli’s Florence through Hobbes’s Leviathan to Thomas Schelling’s Cold War game theory, that treats the appearance of irrationality as among the oldest instruments of statecraft.

It is also a book about what is owed, ethically, to those who live downstream of that instrument. A sovereign’s performed unpredictability is not costless to those who must absorb it. The volume to follow this one will ask what we owe each other, and what kind of resistance preserves the resister’s dignity, when the bargaining strategies of the powerful make ordinary deliberation difficult.

This is, finally, a work of jurimetric and philosophical synthesis rather than a chronicle of any single episode. Its cases are historical, its argument structural; the doctrine it describes recurs across centuries and regimes precisely because it is not the property of any one man.

Kampala, 2026

PART ONE: THE PROBLEM OF THE UNPREDICTABLE SOVEREIGN

Chapter 1: Introduction — Why Irrationality Can Be Rational

1.1 The Inverted Premise

Classical political theory begins, almost without exception, from the premise that the sovereign’s strength is measured by command of himself before command of others. Aristotle’s account of the statesman, the Stoic ideal of the ruler governed by reason, and the mirror-for-princes literature of the medieval period converge on a single image: the good ruler is the self-possessed ruler. Disorder in the prince’s soul, on this view, precedes and produces disorder in the state.

This book is about the doctrine that inverts that premise. It argues — following a line of thought that surfaces explicitly in Niccolò Machiavelli, finds its starkest theoretical statement in Thomas Hobbes, and receives its rigorous modern formulation in Thomas Schelling’s bargaining theory — that under specific structural conditions, the appearance of a ruler’s unreason is not a failure of statecraft but among its most efficient instruments. Call this the madman doctrine: the strategic cultivation of perceived unpredictability as a means of extracting compliance that calm, calculated communication cannot extract at the same price.

The madman doctrine is not the claim that irrational rulers sometimes prevail. That claim is trivially true and theoretically uninteresting: chance favours the unpredictable as it favours anyone, some of the time. The doctrine this book examines is narrower and more unsettling — that a rational actor, reasoning correctly about their interests, can conclude that appearing irrational is the optimal strategy, and that audiences are structurally incapable of fully discounting that appearance, because the cost of being wrong about it is asymmetric.

1.2 The Argument in Brief

The book's central argument unfolds in six movements, corresponding to its six parts.

First, we must understand the problem in its own terms before reaching for solutions: what does it mean for unpredictability to function as a bargaining resource rather than a personal failing (Part One).

Second, we trace the doctrine's classical foundations. Machiavelli's prince must sometimes appear other than virtuous because the appearance of constraint by ordinary moral expectation is itself a constraint his rivals can exploit. Hobbes's sovereign, instituted to escape a state of war defined by the unpredictability of others' wills, paradoxically inherits a will that no covenant can bind — and that unboundedness is part of what makes the Hobbesian sovereign's threats credible in a way a constitutionally hedged ruler's threats are not (Part Two).

Third, we move to the twentieth century, where Thomas Schelling and the bargaining theorists of nuclear deterrence gave the doctrine its first fully rigorous treatment: credibility, commitment, and brinkmanship are not psychological accidents but the products of structures that make backing down costly (Part Three).

Fourth, we examine the doctrine's special relevance to succession — to heirs, claimants, and those whose legitimacy is borrowed rather than earned, for whom the ordinary instruments of authority are foreclosed and provocation becomes one of few remaining levers (Part Four, forthcoming).

Fifth, we turn from explanation to ethics: what is owed by, and to, those who live under a sovereign deploying this doctrine, and what

principled non-reaction actually requires of a person, as distinct from what it superficially resembles (Part Five, forthcoming).

Sixth, we attempt synthesis: a general statement of the doctrine's logic, its structural limits, and the jurisprudential constraints that exist, imperfectly, to contain it (Part Six, forthcoming).

1.3 What This Book Is Not

This book is not a psychological case study of any living person, and deliberately so. The madman doctrine is a structural account of incentives facing a category of actor — the heir, the claimant, the figure whose authority is unconsolidated — not a diagnosis of any individual's mental state. Psychiatric language applied to public figures from a distance is neither reliable nor philosophically serious; the entire argument of this book is that the relevant question is never "is this person actually irrational?" but "does it benefit a rational actor in this position to be perceived as irrational, and what follows from that incentive structure regardless of what is true of any particular mind?" The book's evidentiary base is accordingly historical and theoretical: documented successions, codified bargaining theory, and the textual record of political philosophy's oldest tradition.

Nor is this a manual. The volume to follow, on the ethics of silence, is philosophical rather than operational; it asks what dignity and resistance mean under asymmetric power, not how to evade detection by any particular apparatus. That omission is a feature of the project's seriousness rather than a deficiency.

Chapter 2: The Paradox Stated — Discipline, Office, and Performance

2.1 Office Without Discipline

Every theory of legitimate authority known to the Western and African traditions alike ties the office to a corresponding discipline. The judge is bound by procedure; the soldier by command structure and rules of engagement; the priest by liturgy and vow. Office without discipline is, on the classical view, not authority at all but its counterfeit — what Aristotle would call tyranny: the exercise of rule for the ruler’s private advantage rather than the common good, unconstrained by the law that distinguishes king from tyrant.

The paradox this book investigates arises precisely where office and discipline visibly diverge in the public performance of a power-holder, and where that divergence persists without producing the institutional correction classical theory predicts. The office continues to be occupied; the discipline appears, publicly and repeatedly, to be discarded; and yet the arrangement does not collapse on contact with reality in the way Aristotelian theory would lead us to expect.

2.2 Two Competing Explanations

There are, broadly, two ways to explain a sustained gap between the discipline an office demands and the performance its holder gives.

The deficiency explanation holds that the gap reflects an actual failure of self-governance — the holder cannot meet the discipline the office requires, and the persistence of the arrangement reflects institutional failure to correct it.

The strategic explanation holds that the gap is performed rather than suffered — that the holder retains full capacity to meet the office’s

discipline and chooses, in specific public registers, not to, because the performance itself does work that disciplined conduct cannot do.

These explanations are not mutually exclusive in any given case, and this book does not claim that the strategic explanation is always, or even usually, correct. Its claim is narrower: that the strategic explanation is coherent, has a long and well-theorized pedigree, and deserves to be taken seriously as a hypothesis before observers default to the deficiency explanation merely because it is more intuitive.

2.3 The Asymmetry of Audiences

The strategic explanation gains its force from a structural feature of political communication that Schelling would later formalise: an audience cannot costlessly verify which explanation is true. Distinguishing a man who has lost control from a man who has chosen to appear to lose control requires information the audience typically does not have. This informational asymmetry is not a bug in the system; for the strategic actor, it is the entire mechanism.

This generates what this book will call the discount problem: a rational rival, recognising that performed irrationality is possible, should in principle discount some fraction of any apparently irrational threat as performance rather than genuine intent. But the discount can never reach certainty without inviting the actor to prove otherwise — which is precisely the dynamic that gives the threat its residual force. We return to this formally in Chapter 8.

2.4 Why the Paradox Recurs

If the paradox were idiosyncratic to a single ruler or regime, it would belong to biography rather than philosophy. Its recurrence — across the Italian city-states of the Quattrocento, the dynastic courts of early

modern Europe, and the bargaining rooms of the nuclear age — is what licenses treating it as a structural phenomenon. Wherever an actor's legitimacy is contested, bottlenecked, or borrowed rather than fully settled, and wherever ordinary disciplined communication has failed to resolve that contest, the incentive to perform unpredictability reappears.

Chapter 3: Defining Terms — Bargaining, Credibility, and the Threat That Works Because It Might Be Carried Out

3.1 Bargaining as the Relevant Frame

Political philosophy has tools other than bargaining theory for analysing power — theories of legitimacy, justice, and the common good chief among them. This book deliberately adopts the bargaining frame for its central analytical chapters because the madman doctrine is, at its core, a claim about credibility under uncertainty, and credibility under uncertainty is the proper subject matter of bargaining theory rather than of normative political philosophy as classically practised. We return to legitimacy and ethics in the volume to follow.

3.2 Credibility

A threat or commitment is credible to the extent that the party it is addressed to believes it will be carried out if its conditions are met. Credibility is not the same as sincerity. A perfectly sincere threat issued by an actor with no capacity or willingness to follow through is not credible; a threat the issuer privately regards as bluff can be entirely credible if the audience cannot tell the difference and structural incentives make following through plausible regardless of private intent.

3.3 Commitment

A commitment is a costly signal that narrows an actor's own future options in order to make a stated intention credible.[1]

Performed irrationality functions as a species of commitment device — by behaving in ways that would be costly or humiliating to reverse, the

[11]Schelling, The Strategy of Conflict, ch. 5–6 on the strategic value of irreversible commitment ("burning bridges" as a commitment device).

actor forecloses the safe, calculated retreat that a rival might otherwise expect, and thereby makes the threat to continue behaving unpredictably more credible than a calm assurance of the same intent could ever be.

3.4 Brinkmanship

Brinkmanship denotes the deliberate manipulation of risk — pushing a confrontation toward a threshold of mutual harm not because crossing it is desired, but because the willingness to approach it conveys information about resolve that safer signals cannot convey.[2]

Brinkmanship and the madman doctrine are closely related but not identical: brinkmanship requires the threat of a costly outcome conditional on the rival’s response; the madman doctrine requires only that the actor’s own future conduct be perceived as not fully governed by ordinary cost-benefit reasoning, which raises the rival’s estimate of how likely the actor is to follow through on a brink in the first place.

3.5 The Threat That Works Because It Might Be Carried Out

The chapters that follow return repeatedly to a single formal idea: the threat that works because it might be carried out. An ordinary threat works, when it works, because the threatened party calculates that compliance costs less than the threatened harm, discounted by the probability the harm will actually be inflicted. Where that probability is very low — because the threat is plainly costly to the threatener, or plainly irrational on its face — ordinary game-theoretic reasoning predicts the threat should fail.

The madman doctrine is the strategic exploitation of exactly this prediction’s fragility. By raising the rival’s estimate of the probability that

[22]Schelling, Arms and Influence, ch. 3, “The Manipulation of Risk,” on brinkmanship and the deliberate generation of a probability, rather than a certainty, of mutually disastrous outcomes.

an apparently self-destructive threat will be carried out, the actor can make threats credible that a transparently rational, fully self-interested actor could never make credible. The price of this strategy, examined at length in Part Three, is that it must continually re-purchase its own credibility through conduct that confirms the perception — which is what makes the doctrine self-consuming over time.

3.6 A Note on Method

Each subsequent part of this book pairs a theoretical chapter with a historically or doctrinally grounded illustration. The illustrations are drawn from settled historical record — episodes with the benefit of historical distance and documentary consensus — rather than from contested contemporary events, for the methodological reason given in Chapter 1.3: the argument concerns structure, not biography, and structure is best demonstrated where the biographical noise has had time to settle.

PART TWO: THE CLASSICAL FOUNDATIONS

Chapter 4: Machiavelli and the Uses of Feared Unpredictability

4.1 The Misreading of Machiavelli

No political theorist's name is invoked more carelessly in defence of the madman doctrine than Niccolò Machiavelli's, and few are more carelessly read. The popular shorthand — "Machiavelli said rulers should be unpredictable and cruel" — flattens an argument considerably more conditional, more concerned with perception management than with cruelty as such, and more anxious about the ruler's actual security than the caricature allows.

The textual anchor is the famous discussion in The Prince, Chapter XVII, of whether it is better for a ruler to be loved or feared. Machiavelli's answer is not "feared," full stop. It is that love is unreliable because it depends on the will of the governed, which the ruler does not control, whereas fear depends on the will of the ruler, which he does control — and that the prudent ruler will therefore prefer to rely on the instrument within his own power.[3]

Crucially, Machiavelli immediately qualifies this: the ruler should be feared in such a way as to avoid hatred, because hatred, unlike fear, invites conspiracy.[4]

The doctrine is not "maximise terror"; it is "maximise the credibility of consequences while minimising the personal animus that makes consequences worth risking everything to prevent."

4.2 The Fox and the Lion

---

[3]Niccolò Machiavelli, The Prince, ch. XVII, "On Cruelty and Clemency, and Whether It Is Better to Be Loved Than Feared" (1532; Mansfield trans., 2nd ed., University of Chicago Press, 1998).

[4]Machiavelli, The Prince, ch. XVII.

The chapter most directly relevant to performed unpredictability is Chapter XVIII, on the keeping of faith. Machiavelli argues that a ruler who wishes to maintain his position cannot always keep faith, because circumstances change and rigid commitment to past promises can become a liability; the prudent ruler must know how to be, when necessary, a fox to recognise traps and a lion to frighten wolves.[5]

The often-quoted image is frequently mistaken for an endorsement of dishonesty as such. Its actual argument is narrower: Machiavelli describes a ruler who must appear consistently virtuous while retaining, privately, the flexibility to act otherwise when the situation demands it — and who must be skilled enough in the performance that the gap between appearance and capacity is not visible to those who would exploit it.

This is the first clear statement, in the canon this book traces, of managed perception as a tool of rule distinct from the underlying disposition of the ruler. Machiavelli's prince is not asked to become unpredictable in his own deliberation — Machiavelli's entire treatise is, if anything, a manual for extremely careful, calculated deliberation. He is asked to be perceived as someone whose conduct cannot be reliably modelled by ordinary expectations of virtue, because that unpredictability of perception is what keeps rivals from planning confidently against him.

4.3 Severity Concentrated, Not Diffused

A further and frequently overlooked Machiavellian principle bears directly on the madman doctrine's self-limiting character, examined at length in Chapter 10. In the Discourses on Livy, Machiavelli praises rulers

[5]Machiavelli, The Prince, ch. XVIII, "In What Mode Faith Should Be Kept by Princes."

who, when cruelty is necessary, commit it once, decisively, and early, rather than in a continuous drip — because injuries done all at once are felt less, over time, than injuries repeated, which compound resentment and produce exactly the conspiratorial hatred Chapter XVII warned against.[6]

This is a theory of dosage. The performance of severity or unpredictability, on Machiavelli's own account, has a half-life; deployed continuously rather than in calculated bursts, it does not sustain fear indefinitely — it converts fear into the contempt and hatred that destroy rulers.

4.4 Legitimacy and the New Prince

Machiavelli's analysis is calibrated specifically to what he calls the "new prince" — the ruler whose authority is not inherited through long-settled custom but newly acquired or newly contested.[7]

The hereditary prince who succeeds to an old and settled state, Machiavelli observes, has an easier task: custom itself disposes subjects toward obedience, and only gross misrule will dislodge him.[8]

The new prince enjoys no such presumption. He must construct legitimacy that custom has not yet supplied, and Machiavelli's treatise can be read as a catalogue of substitute instruments available to him — instruments the settled hereditary ruler has comparatively little need of. Performed unpredictability belongs, in Machiavelli's own framework, principally to this second category.

---

[6]Machiavelli, Discourses on Livy, bk. III, ch. 6 and surrounding discussion of conspiracies and the concentration of harsh measures (Mansfield & Tarcov trans., University of Chicago Press, 1996).
[7]Machiavelli, The Prince, ch. I–III, distinguishing hereditary from new principalities, and ch. VI–VII on princes who acquire states by their own arms and ability.
[8]Machiavelli, The Prince, ch. II.

4.5 Machiavelli's Limiting Condition

It is worth stating plainly what Machiavelli does not license, because the popular caricature elides it. Machiavelli explicitly warns against rulers becoming objects of contempt through actions seen as frivolous, irresolute, effeminate, or fainthearted — instability in the popular perception of competence, as opposed to calculated unpredictability in specific tactical registers, is treated by Machiavelli as a vulnerability to be avoided, not a strategy to be pursued.[9]

The Machiavellian doctrine, properly read, is a doctrine of controlled and purposive departures from expected conduct — not a license for indiscriminate volatility, which Machiavelli treats as a symptom of weakness rather than an instrument of strength.

[9]Machiavelli, The Prince, ch. XIX, "That One Should Avoid Being Despised and Hated."

Chapter 5: Hobbes, Sovereign Will, and the Logic of Unanswerable Power

5.1 The State of Nature as the Problem Unpredictability Solves

Thomas Hobbes's Leviathan (1651) begins from a condition the madman doctrine, paradoxically, both escapes and reproduces. The state of nature, on Hobbes's account, is a war of every man against every man not because men are uniformly vicious but because, absent a common power to keep them in awe, no one can ever be confident enough of another's restraint to lay down his own arms first.[10]

Hobbes's entire argument for the necessity of sovereignty is an argument that only an artificial, overwhelming, and credible threat of punishment can convert mutual unpredictability into stable expectation.[11]

5.2 The Sovereign as the Last Unconstrained Will

Hobbes's solution carries a built-in irony directly relevant to this book. To solve the problem of universal unpredictability, the covenant by which the commonwealth is formed concentrates an unconstrained will in a single person or assembly — the sovereign, who is party to no covenant himself and is therefore bound by none.[12]

Hobbes is explicit that the sovereign cannot, properly speaking, break covenant with his subjects, because he never made one with them — he is the author of the laws that bind everyone else, and authors are not bound by what they author in the way the bound are.[13]

This produces a structural feature of the Hobbesian sovereign highly

---

[10]Thomas Hobbes, Leviathan, Part I, ch. XIII, "Of the Natural Condition of Mankind" (1651).
[11]Hobbes, Leviathan, Part II, ch. XVII–XVIII, "Of the Causes, Generation, and Definition of a Commonwealth" and "Of the Rights of Sovereigns by Institution."
[12]Hobbes, Leviathan, Part II, ch. XVIII.
[13]Hobbes, Leviathan, Part II, ch. XVIII, on the sovereign as author rather than party to the covenant.

relevant to the madman doctrine: the sovereign is, definitionally, the one actor in the commonwealth whose conduct cannot be predicted by reference to enforceable covenant, because there is no higher power to enforce any covenant against him.

5.3 Credibility Without External Constraint

This generates what we will call the Hobbesian credibility surplus. An ordinary actor's threats are discounted by rivals' awareness of the constraints — legal, social, reputational — that bind the threatener. The Hobbesian sovereign, by design, faces no such external constraint; the only check on the sovereign's conduct, on Hobbes's account, is the sovereign's own prudential calculation of what preserves his power, and prudence is not the same thing as a binding rule a rival can rely upon in advance.[14]

A rival reasoning about whether to test the Hobbesian sovereign's threats cannot reduce his estimate of their credibility by appeal to any external check, because Hobbes's system has constructed the sovereign specifically to have none.

5.4 The Limiting Principle Hobbes Builds In

Hobbes, like Machiavelli, builds a limiting principle into the very logic that seems to license unconstrained power. The entire purpose of instituting sovereignty, for Hobbes, is the subjects' own peace and security — the sovereign's power is instrumental to an end, not an end constituted by the sovereign's appetite.[15]

Hobbes further holds that the obligation of subjects to obey lasts only so long as the sovereign is able to protect them; where the sovereign's

[14]Hobbes, Leviathan, Part II, ch. XVIII and XXI, "Of the Liberty of Subjects."
[15]Hobbes, Leviathan, Part II, ch. XVII, on the end ("finis") of the institution of sovereignty being the subjects' peace and common defence.

power decays to the point that it can no longer perform the protective function for which it was instituted, the obligation itself lapses, because the covenant's entire rational basis was the exchange of obedience for protection.[16]

A sovereign whose performed unpredictability becomes so destabilising that it undermines, rather than secures, the peace it was instituted to provide is, on Hobbes's own terms, sawing through the branch on which his authority sits.

5.5 Hobbes and the Problem of the Heir

Hobbes's discussion of succession, particularly in Leviathan, Part II, Chapter XIX, is concerned almost entirely with stability — with how a commonwealth avoids relapsing into the state of nature at the moment of transition between sovereigns.[17]

An heir whose claim is contested, or who must establish that he commands the same unconstrained credibility his predecessor accumulated over a lifetime, faces — within Hobbes's own framework — a distinctive incentive to demonstrate that unconstrained will quickly and visibly, since the Hobbesian credibility surplus does not transfer automatically with the title; it must, on this reading, be re-established by the heir's own conduct.

5.6 Synthesis: What Machiavelli and Hobbes Together Establish

Read together, Machiavelli and Hobbes establish the two load-bearing premises on which the modern bargaining-theoretic treatment in Part Three builds a formal apparatus. Machiavelli establishes that perception

[16]Hobbes, Leviathan, Part II, ch. XXI, on the duration of the obligation of subjects being coextensive with the sovereign's power to protect them.
[17]Hobbes, Leviathan, Part II, ch. XIX, "Of the Several Kinds of Commonwealth by Institution, and of Succession to the Sovereign Power."

of a ruler’s disposition, rather than the ruler’s actual private disposition, is the operative variable in political bargaining. Hobbes establishes that the absence of enforceable external constraint is itself a source of bargaining credibility, independent of anything about the constrained party’s actual psychology. Neither philosopher, properly read, endorses indiscriminate volatility as a strategy; both build explicit limiting principles that anticipate the doctrine’s self-limiting character.

Chapter 6: The Renaissance Court as Laboratory — Performance, Spectacle, and Survival

6.1 Why the Court, Specifically

The Italian princely courts of the fifteenth and sixteenth centuries — the world Machiavelli wrote within and about — offer an unusually well-documented setting for the dynamics described in Chapters 4 and 5, because the court was a structurally small, densely observed environment in which a ruler's every gesture was simultaneously a private act and a public signal read by rivals, subjects, and foreign emissaries alike. The historical study of Renaissance diplomacy has long emphasised how closely resident ambassadors observed and reported on the comportment of the princes to whom they were posted, precisely because that comportment carried strategic information their home governments needed.

6.2 The Asymmetry of the Observed Court

The court setting illustrates with unusual clarity the discount problem introduced in Chapter 2.3. A foreign observer, weighing a prince's volatile conduct, faced real consequences for misjudging which explanation was correct: discount a genuine threat as performance, and disaster might follow; treat performance as genuine and capitulate unnecessarily, and one had needlessly surrendered advantage. This structural bind — in which the cost of misjudging is high and asymmetric in both directions — is precisely the condition under which the madman doctrine is most likely to be rationally adopted by the party able to exploit it.

6.3 Spectacle as Distinct from Loss of Control

It is worth distinguishing, within this historical setting, between two phenomena frequently conflated by later popular history: the deliberately

staged spectacle of princely severity — choreographed and calibrated to a specific audience and effect — and the genuinely uncontrolled episode. The existence of both categories in the historical record is itself instructive: contemporaries were capable of, and did, distinguish performance from genuine breakdown when sufficient information was available to them, which is exactly why public, low-information audiences remain considerably more vulnerable to the doctrine's effects than a small circle of continuous observers.

6.4 What the Laboratory Teaches

The Renaissance court does not, on the available record, support a simple story in which performed unpredictability invariably succeeds or invariably fails. What it supports is the more specific claim that motivates this book's structure: that the rulers remembered by history as having wielded feared unpredictability effectively are reliably distinguished, in retrospect, by exactly the limiting principles Machiavelli and Hobbes articulate — concentrated rather than continuous displays, calibrated to specific strategic ends, and discontinued once the desired recalibration of rival expectations had been achieved. Where the performance instead became continuous and untethered from specific strategic objectives, the historical verdict is considerably harsher, and considerably more frequently fatal to the ruler concerned.

PART THREE: THE MODERN BARGAINING DOCTRINE

Chapter 7: Schelling and the Strategy of Conflict — Commitment, Credibility, Brinkmanship

7.1 From Philosophy to Formal Theory

Thomas Schelling's The Strategy of Conflict (1960) did not set out to theorise sovereigns or princes. It set out to understand a narrower and more urgent twentieth-century problem: how could two nuclear-armed powers, each capable of annihilating the other, communicate threats and commitments credibly enough to avoid war, when the threats both sides needed to make — to use weapons that would also devastate the threatener — were precisely the threats that rational self-interest, taken at face value, made nearly impossible to believe? Schelling's achievement, recognised by the 2005 Nobel Memorial Prize in Economic Sciences, was to show that certain strategies can make seemingly irrational threats credible, not by making the threatener irrational, but by altering the structure of the situation so that carrying out the threat becomes the threatener's genuinely rational choice once a certain point is reached.

This formal apparatus, built for a problem far removed from princely courts, supplies exactly the missing rigour that Machiavelli and Hobbes gesture toward but never formalise. We import it here not because Schelling was thinking about heirs and provocateurs, but because the structure he identified is general, and the madman doctrine is one of its applications.

7.2 The Rationality of Apparent Irrationality

Schelling's central insight, developed across several chapters of The Strategy of Conflict and extended in Arms and Influence (1966), is that bargaining power often derives not from strength or resources but from

the capacity to limit one's own future choices in a way the other party can verify.[18]

An actor who can credibly demonstrate that they have foreclosed their own option to back down has made a threat that calculation alone could not support.[19]

The connection to perceived irrationality is direct and was made explicit by Schelling himself: a reputation for being somewhat irrational, somewhat insensitive to ordinary costs, or somewhat beyond the reach of reasoned persuasion can function as exactly this kind of commitment device.[20]

If a rival believes that the threatener might carry out a mutually costly threat regardless of whether ordinary cost-benefit reasoning would recommend it, the rival's own incentive to test that threat is reduced — not because the threat has become more attractive to carry out, but because the rival can no longer comfortably assume it will not be carried out.

7.3 The Vocabulary of Commitment

A threat is conditional: it promises a cost contingent on the target's future behaviour, and its purpose is to influence behaviour that has not yet occurred.[21]

A commitment is the structural move that makes a threat believed: a

---

[18]Thomas C. Schelling, The Strategy of Conflict (Harvard University Press, 1960), esp. ch. 2, "An Essay on Bargaining," and ch. 5–6 on commitment and credibility. Argument paraphrased; Schelling holds copyright in the original text.

[1919]Schelling, The Strategy of Conflict, ch. 5–6 on the strategic value of irreversible commitment ("burning bridges" as a commitment device).

[20]Schelling, The Strategy of Conflict, ch. 2 and ch. 7, on the deliberate cultivation of a reputation for irrationality or insensitivity to cost as a bargaining device; see also Thomas C. Schelling, Arms and Influence (Yale University Press, 1966), ch. 1–2.

[21]Schelling, The Strategy of Conflict, ch. 2, on the conditional and forward-looking structure of a threat as distinct from a simple statement of hostile intent.

costly, often irreversible action that narrows the threatener's own future options, so that the threat's fulfilment becomes more likely not because the threatener wants to fulfil it but because alternatives have been foreclosed.[22]

Brinkmanship is the deliberate generation of a probability, rather than a certainty, of mutual disaster as the actual instrument of coercion.[23]

Schelling's insight here is genuinely counterintuitive: a threat that guarantees catastrophe if defied is sometimes less credible than a strategy that merely raises the risk of catastrophe, because the threatener retains enough residual control to make the threat believable as something issued by an actor still recognisably pursuing self-interest, while the target's uncertainty about exactly how much risk has been generated does the actual coercive work.

7.4 The "Rationality of Irrationality" Stated Formally

Schelling's most direct treatment of the phenomenon this book is concerned with concerns the proposition that a fully rational actor, anticipating that a rival will exploit any known points of rational restraint, may find it optimal to eliminate predictable rational restraint from their own observed behaviour, thereby making threats credible that a transparently rational actor's interlocutor would otherwise discount.[24]

Crucially, on this formalisation, what is actually irrational is never the underlying strategic choice to cultivate the reputation — that choice is

---

[22]Schelling, The Strategy of Conflict, ch. 5, on commitment as a means of narrowing one's own future choice set to render a stated intention credible.

[2323]Schelling, Arms and Influence, ch. 3, "The Manipulation of Risk," on brinkmanship and the deliberate generation of a probability, rather than a certainty, of mutually disastrous outcomes.

[24]This formulation draws on Schelling's discussion in The Strategy of Conflict, ch. 2 and ch. 7, of the strategic advantages of apparent irrationality; the specific label "rationality of irrationality" is used descriptively in this book to denote the mechanism Schelling identifies and is not attributed to Schelling as his own coined phrase.

the product of careful calculation about audience perception and remains fully rational at the level of strategy selection. What is performed, and only performed, is the appearance of irrationality in the specific tactical conduct that constitutes the reputation. This is the formal vindication, three and a half centuries later, of the Machiavellian fox who must appear as something other than what he calculatedly is.

7.5 The Limits Schelling Himself Identified

It would misrepresent Schelling badly to present his theory as an unqualified endorsement of cultivated unpredictability. Schelling was deeply concerned with the problem of loss of control: a strategy that depends on convincing a rival that events may unfold independently of rational restraint creates a genuine risk that events will unfold independently of rational restraint, because the actor has deliberately weakened the very mechanisms — clear communication, verifiable signals, calibrated graduated response — that would otherwise allow a crisis to be managed back down from the brink.[25]

The doctrine, on Schelling's own account, is therefore inherently and irreducibly risky to the practitioner, not merely to the target: it is a strategy for purchasing coercive power at the price of accepting a standing risk of catastrophic miscalculation — a price that compounds the longer and more frequently the strategy is deployed. This is the modern, formalised counterpart to the Machiavellian dosage principle and the Hobbesian conditioning of obligation on continued protective capacity, examined in Chapters 4.3 and 5.4 respectively: three independent theoretical traditions, separated by centuries, converging on the same structural conclusion.

---

[25]Schelling, Arms and Influence, ch. 1–2, on the risks of genuine loss of control inherent in strategies that deliberately weaken calibrated, rational signalling.

Chapter 8: The Rationality of Apparent Madness — Game-Theoretic Foundations

8.1 A Simple Formal Model

To make the discount problem of Chapter 2.3 precise, consider a minimal model in the spirit of Schelling's own informal game-theoretic reasoning, without claiming the false precision of a fully specified mathematical proof. An actor, A, issues a threat against a target, B: if B does not comply with a demand, A will impose a cost on B, at some cost also to A. B's optimal response depends entirely on B's estimate of the probability, p, that A will actually carry out the threat if B refuses. Where p is sufficiently low — because B believes A to be a calculating, self-interested actor for whom carrying out the threat is, on its face, costly and pointless — standard rational-actor reasoning predicts B should refuse, calling what amounts to a bluff.

The madman doctrine operates entirely on the value of p. By cultivating a reputation such that B cannot confidently assign a low value to p — because B cannot be sure A's conduct is governed by the cost-benefit calculation B would otherwise rely upon to predict it — A raises the expected cost of refusal in B's own calculation, without A needing to change anything about what A would actually prefer to do if the threat were tested.

8.2 Why the Discount Can Never Reach Zero — or Unity

Chapter 2.3 introduced the discount problem informally; the model in 8.1 lets us state its structure precisely. A fully rational B, aware that A might be performing rather than genuinely unconstrained, faces a second-order uncertainty: not merely "will A carry out the threat" but "is A's apparent unpredictability itself real or staged." If B could verify the

answer to the second question with certainty, B's problem would collapse back into an ordinary, fully discountable threat — once an audience is certain a threat is theatre, the theatre stops working, exactly as Schelling's reputational logic predicts. But verifying the second question with certainty is, by the doctrine's own design, exactly what the strategy is structured to prevent.

8.3 Reputation as an Asset That Depreciates Through Both Use and Disuse

A further refinement, owed to the broader deterrence-theoretic literature that builds on Schelling's foundational work, is that the reputation for unpredictability behaves asymmetrically with respect to its use. A reputation entirely untested can decay through sheer staleness, as rivals' priors drift back toward assuming ordinary rational restraint in the absence of confirming evidence. But a reputation tested too frequently, or in registers where the costs of follow-through become visibly ruinous to the actor's own position, depreciates by a different and faster mechanism: rivals update toward the deficiency explanation of Chapter 2.2, concluding not that the actor is performing strategically but that the actor's conduct reflects a settled pattern institutions, allies, and rivals alike can plan confidently around — precisely the condition the doctrine exists to prevent. The reputation occupies a narrow band: confirmed often enough to remain credible, and rarely enough to remain genuinely uncertain.

8.4 The Asymmetry Between State and Personal Application

It is necessary to flag a limit on the analogy this chapter has been building. The deterrence literature from which it draws was developed principally to describe relations between sovereign states, where the actor deploying the strategy is an institution with internal checks and

collective decision-making distributed across many hands, even when a single head of state is its public face. The application of the same logic to an individual power-holder's personal, public conduct introduces a complication the classical deterrence literature does not fully anticipate: an individual's reputation for unpredictability is harder to calibrate precisely because it is not the product of a deliberative process with internal moderating voices, but the product of one person's choices, made in real time, often in public, with no equivalent internal check on dosage of the kind Machiavelli's discipline or a state's bureaucratic process would otherwise supply. This is a genuine theoretical gap rather than a refutation of the analogy, and the volume to follow engages with it directly in the specific context of the heir or claimant whose position structurally resembles the individual rather than the institutional case.

Chapter 9: Deterrence, Compellence, and the Economy of Threats

9.1 Two Directions of Coercion

Schelling's later work, particularly Arms and Influence, draws a distinction essential to evaluating the madman doctrine's actual field of application: the distinction between deterrence and compellence.[26]

Deterrence seeks to prevent an action that has not yet occurred — it is a standing threat conditioned on the target's restraint, and its success is invisible, since a successful deterrent threat produces nothing observable beyond the absence of the deterred action. Compellence, by contrast, seeks to produce an action — to make a target do something, or stop doing something already underway — and its success or failure is visible, often immediately, which makes compellent threats considerably harder to sustain credibly than deterrent ones.

9.2 Why the Madman Doctrine Suits Deterrence Better Than Compellence

This distinction matters directly for evaluating claims that performed unpredictability secures a power-holder's position. The doctrine functions relatively well in a deterrent register — discouraging rivals from testing the actor's resolve, without ever requiring the actor to demonstrate anything beyond a standing, untested reputation.[27]

It functions considerably worse in a compellent register, where the actor makes specific, visible demands, because compellent demands invite an immediate test: either the demand is met, which is observable, or it is refused, which forces the actor either to escalate, at real cost and

[26]Schelling, Arms and Influence, ch. 2, "The Diplomacy of Violence," on the distinction between deterrence and compellence.
[27]Schelling, Arms and Influence, ch. 2, on the comparative difficulty of sustaining compellent as against deterrent threats, owing to the visibility of compliance or defiance.

real risk, or to back down, which inflicts exactly the reputational damage the strategy was designed to avoid.

9.3 The Economy of Threats

Schelling’s broader contribution, beyond the specific mechanisms above, was to insist that threats are not free goods to be issued without limit, but a scarce resource subject to genuine economic constraints: each threat issued draws on the same finite stock of credibility, and a threat issued and not enforced does not merely fail on its own terms but depletes the credibility available for every subsequent threat, including threats on entirely unrelated matters.[28]

This “economy of threats” framing supplies the conceptual bridge to the historical material the volume to follow will treat at length: an actor who issues threats across an unusually wide range of registers is not multiplying coercive leverage but drawing down a single shared account at an accelerating rate — an account that depreciates through both overuse and the eventual, near-statistically-inevitable instance of a threat tested and not enforced.

9.4 Synthesis Closing Part Three

The three chapters of Part Three together establish the formal architecture this book needed in order to move from suggestive historical parallel to rigorous structural claim. Performed unpredictability is, on this account, a specific and theoretically well-understood instrument of coercive bargaining: it manipulates a target’s probability estimate of follow-through, it is asymmetrically suited to deterrent rather than compellent applications, and it draws on a depletable, shared stock

[28]Schelling, The Strategy of Conflict, ch. 2, on threats and commitments as drawing upon a finite, shared stock of credibility rather than constituting a costless resource.

of credibility rather than generating costless leverage — all while remaining, per Schelling's own explicit warnings, inherently and irreducibly risky to its own practitioner through the standing possibility of genuine loss of control.

Part Four now turns to the specific structural position — the heir, the claimant, the new prince in Machiavelli's own original sense — for whom this entire apparatus has historically proven most tempting to deploy, before Part Five turns to the ethics of those who must live within reach of it, and Part Six draws the whole inquiry to a close.

PART FOUR: SUCCESSION, INHERITANCE, AND THE PERFORMANCE OF POWER

Chapter 10: Heirs, Thrones, and the Problem of Borrowed Legitimacy

10.1 The Structural Position of the Heir

Every theoretical thread assembled in Parts Two and Three converges on a single structural figure: the actor whose claim to authority is not yet settled by the ordinary sources of legitimacy available to an established ruler. Machiavelli's "new prince" (Ch. 4.4), Hobbes's anxiety about the moment of succession (Ch. 5.5), and the bargaining theorist's observation that reputation must be actively built rather than inherited (Ch. 8.3) all describe, from different angles, the same predicament: a position from which the ordinary instruments of authority — custom, settled expectation, an already-proven record of consequences delivered — are largely unavailable, precisely because they take time and continuity to accumulate, and the heir's claim, by definition, has not yet had that time.

This chapter examines why this structural position recurrently produces the incentive toward performed unpredictability described in Parts Two and Three, drawing the threads together into a single account of why the doctrine clusters where it clusters.

10.2 Legitimacy as a Stock, Not a Flow

Borrowing the economic vocabulary introduced in Chapter 9.3, legitimacy is best modelled as a stock — an accumulated reserve built over time through consistent, predictable governance that meets subjects' settled expectations — rather than a flow that can be generated instantaneously by any single act, however dramatic. A hereditary ruler who has occupied an office for decades draws on a stock of accumulated legitimacy that no single crisis can fully deplete; a new

claimant, by contrast, has no such stock, and faces the structural temptation to substitute intensity for duration — to attempt, through dramatic and attention-commanding acts, what only time would otherwise supply.

This substitution is exactly the dynamic the madman doctrine offers. Where the hereditary ruler's threats are credible because rivals have observed, across years, a consistent pattern of follow-through, the new claimant cannot point to such a pattern. Performed unpredictability is one of very few instruments capable of generating some credibility in compressed time, because — as established in Chapter 8.1 — it works by raising a rival's uncertainty rather than by demonstrating a track record, and uncertainty, unlike a track record, can be manufactured quickly.

10.3 The Inheritance Problem Proper

A further complication, anticipated in Chapter 5.5's discussion of Hobbes, is what this book terms the inheritance problem: where an heir succeeds to, or stands poised to succeed to, a position previously occupied by a ruler who did successfully accumulate a stock of feared unpredictability over a long tenure, the heir faces pressure to demonstrate continuity of that same credibility — but cannot simply inherit it as a title is inherited, because the reputation, as established in Chapter 8.3, is a property of demonstrated conduct, not of office. The heir who wishes to be taken as seriously as the predecessor must, on this account, re-earn through personal conduct a reputation the predecessor spent a career building, and the temptation to compress that re-earning into a short, intense burst of conspicuous unpredictability is structurally built into the position itself, independent of anything particular to the individual occupying it.

10.4 The Asymmetric Stakes of the Heir's Experiment

The heir's position carries a final structural feature worth isolating. An established ruler experimenting with performed unpredictability risks a stock of legitimacy built over a career; an heir experimenting with the same instrument risks comparatively little accumulated legitimacy, because there is little yet accumulated to lose — but risks disproportionately more of the only asset available to a contested claimant: the perception, among rivals and institutions alike, that the succession itself is settled and need not be relitigated. Chapter 7.5's discussion of Schelling's warning about loss of control applies here with particular force: an heir's miscalculated brinkmanship does not merely risk a single crisis, as it might for an established ruler with reserves of legitimacy to draw on, but risks reopening the succession question itself — inviting exactly the recalculation among institutional actors and rival claimants that the performance was designed to foreclose.

Chapter 11: Historical Successions as Case Studies

11.1 Method and Selection

Consistent with the methodological commitment stated in Chapter 1.3 and Chapter 3.6, this chapter examines successions with the benefit of historical distance — episodes on which the documentary record has settled and which do not require speculation about any living person's intentions or mental state. Two episodes are examined for their structural fit to the inheritance problem described in Chapter 10.3, not for any claim that they are uniquely representative; they are illustrations of a recurring structure, not an exhaustive survey.

11.2 The Roman Imperial Succession and the Problem of the Purple

The early Roman Empire institutionalised, more starkly than perhaps any other system in the Western historical record, the inheritance problem described in Chapter 10.3. Imperial authority depended on a triangulated legitimacy among the Senate, the Praetorian Guard, and the legions — none of which was bound by anything resembling a fixed rule of hereditary succession in the way later European monarchies would develop.[29]

A new emperor inheriting the position, even when descended from or adopted by a popular predecessor, could not simply assume that the legions' loyalty transferred automatically with the title; that loyalty, where it existed, had typically been earned by the predecessor through donatives, demonstrated favour, and reputation built over a command

[29]On the institutionalised triangulation of Senate, Praetorian Guard, and legions as the basis of early Roman imperial legitimacy, see generally Fergus Millar, The Emperor in the Roman World, 31 BC–AD 337, 2nd ed. (London: Duckworth, 1992).

career the heir frequently had not yet had time to replicate.

The historical pattern visible across multiple early imperial successions — without attributing uniform motive to every actor involved, which the documentary record does not support — is that incoming emperors frequently moved with conspicuous speed and visible decisiveness in the earliest period of a new reign, including in registers calibrated less toward administrative necessity than toward rapid, visible signal-sending to exactly the audiences — Senate, Guard, legions — whose continued acquiescence the new emperor could not yet take for granted. This is the inheritance problem in its starkest institutional form: legitimacy contested simultaneously among three audiences with different and partially competing criteria for what would satisfy them.

11.3 Renaissance Succession and the Afterlife of a Feared Predecessor

The Italian city-states examined in Chapter 6 furnish a complementary illustration. Where a ruling family’s founder had built, over a long tenure, a reputation for unpredictable severity that secured the family’s hold on a contested city, successors faced the dilemma described in Chapter 10.3 directly: contemporary diplomatic correspondence from such transitions repeatedly registers observers’ explicit uncertainty about whether a successor’s early conduct — often markedly more severe or theatrically assertive than circumstances obviously required — reflected the successor’s own temperament or a calculated attempt to signal, to a watching court and a watching populace accustomed to the founder’s reputation, that the family’s capacity for consequence had not lapsed with the change of occupant.

The documentary record does not permit confident adjudication, in any single case, between the deficiency and strategic explanations

introduced in Chapter 2.2 — which is itself the point: the very uncertainty that made the founder's original reputation effective is what the successor is straining, with uneven success across recorded cases, to reproduce on a compressed timeline.

11.4 What the Comparison Yields

Read together, these illustrations support the chapter's narrow claim: the inheritance problem is not a quirk of any single political culture or century but a structural feature of succession itself, wherever authority depends even partly on perceived unpredictability or feared capacity for consequence rather than purely on settled procedural legitimacy. They also illustrate the self-limiting dynamics established in Chapters 4.3, 5.4, and 7.5: the historical record's harshest judgments, across both settings, fall not on successors who demonstrated resolve through calibrated, eventually-moderated displays, but on those for whom the performance became continuous, untethered from specific strategic objectives, and indistinguishable — even to close observers — from the deficiency explanation. Chapter 12 turns to the institutional consequence of exactly this failure mode.

Chapter 12: Institutions Under Strain — When the Apparatus Resents the Performer

12.1 The Institution as a Distinct Interest

Parts Two and Three have treated the bargaining relationship largely as a dyad: an actor, and the rival or audience whose behaviour the actor seeks to influence. Real political authority, however, is rarely exercised by an individual alone; it depends on an apparatus — bureaucratic, military, administrative — whose continued cooperation the position-holder requires, and whose own professional norms, institutional culture, and accumulated reputational stake in the institution’s standing constitute an interest distinct from the position-holder’s own, even where the apparatus is nominally and legally subordinate to the position-holder’s command.

12.2 The Professional Discipline as a Counter-Reputation

Bureaucratic and military institutions, in particular, typically cultivate their own internal reputational economy, built around professional competence, institutional discipline, and international standing among peer institutions. This institutional reputation is, in the terms of Chapter 9.3, a separate stock of credibility from the position-holder’s personal reputation for unpredictability, and the two stocks are not merely independent but can be placed in direct tension: conduct that builds the position-holder’s personal bargaining credibility along the lines described in Part Three can simultaneously deplete the institution’s own professional reputational stock, particularly where that conduct is, or is widely perceived within the institution to be, inconsistent with the institution’s own professional self-conception.

12.3 The Resentment Mechanism

This generates a distinctive risk, alluded to at several points in Parts Two and Three but not yet examined directly: that the very apparatus on which a position-holder’s authority practically depends may come to regard the position-holder’s performed unpredictability not as effective strategy but as a depletion of an asset — institutional standing, professional reputation, external relationships — that the apparatus’s own members have themselves invested careers in building and have reason to wish preserved. Where this perception becomes sufficiently widespread within an institution’s senior ranks, the institution’s continued cooperation becomes itself a variable rather than a constant — and a position-holder who has spent a reputational stock built on personal unpredictability may discover that stock cannot be exchanged for the cooperation of an apparatus that has concluded the exchange no longer serves its own institutional interest.

12.4 A Structural Observation, Not a Prediction

It is important to be precise about what this chapter does and does not claim. It does not claim that institutional resentment of this kind reliably or predictably produces any particular institutional response in any given case; institutions vary enormously in internal cohesion, in the channels available for internal dissent, and in the costs and risks attached to any form of internal correction, and the historical record offers no general law governing when accumulated institutional resentment translates into institutional action as opposed to continued, if reluctant, acquiescence. The chapter’s claim is narrower: that the tension described in 12.2–12.3 is a real and theorisable feature of the relationship between a position-holder deploying the madman doctrine and the apparatus that doctrine’s continued effectiveness ultimately depends upon, and that this tension constitutes a further, institution-

level instance of the self-limiting dynamic this book has traced at the level of individual reputation (Ch. 8.3), classical dosage (Ch. 4.3, 5.4), and bargaining risk (Ch. 7.5).

PART FIVE: THE ETHICS AND PHENOMENOLOGY OF SILENCE

Chapter 13: Silence as Strategy Versus Silence as Complicity

13.1 The Question This Part Addresses

Parts One through Four have been almost entirely explanatory: an account of why performed unpredictability functions as it does, for whom, and under what conditions it succeeds or fails on its own terms. This part changes register. It asks a normative question Parts One through Four deliberately bracketed: what, if anything, is owed — ethically, by a person to themselves and to others — when that person lives within reach of an actor deploying the doctrine examined above? And in particular, what does not reacting to provocation actually mean, ethically, as opposed to what it superficially resembles?

13.2 Two Silences

The philosophical literature on political quietism and complicity has long recognised that "silence" is not a single act but a name covering at least two structurally distinct postures, and the distinction matters enormously for what follows.

The first is what we will call silence as withdrawal: the abandonment of judgment itself, a retreat from evaluating what is happening at all, whether from exhaustion, denial, or the simple cognitive economy of not wanting to know. This is the silence classical accounts of complicity are right to treat with suspicion, because it is not really a response to the situation but an evasion of having a response.

The second is what we will call silence as discipline: a fully engaged, clear-eyed judgment about what is happening, combined with a deliberate choice not to externalise that judgment in the particular register the provocateur is soliciting. This is not an absence of response;

it is a specific, chosen response, selected because the alternative — the reaction being solicited — would itself serve the provocateur’s strategic interest, for reasons traceable directly to the bargaining-theoretic analysis of Parts Three and Four.

13.3 Why the Distinction Matters Ethically

Edmund Burke’s frequently invoked dictum — that the triumph of evil requires only that good people do nothing — is correct as a description of silence-as-withdrawal and a poor description of silence-as-discipline, and collapsing the two is the most common ethical error this part aims to correct. Burke’s warning is properly addressed to the abandonment of judgment, not to every instance of declining to react in a particular visible register. A person who has formed a clear, considered judgment that a provocation is calculated precisely to elicit a reaction that serves the provocateur’s bargaining position, and who declines to supply that reaction while continuing to act, in other registers, on the judgment they have formed, has not done “nothing” in Burke’s sense. They have made a strategic and, this chapter argues, potentially ethically serious choice about which register of response actually serves the values Burke’s warning is meant to protect.

13.4 The Risk of Self-Deception

None of this licenses an uncritical conflation in the opposite direction. The distinction between silence-as-discipline and silence-as-withdrawal is analytically clear; the lived experience of choosing between them is not, and a person practising what they take to be principled non-reaction is in a poor position to be fully certain, from the inside, that they are not actually practising withdrawal and supplying after the fact a more flattering description of it. This chapter offers a single diagnostic

question worth taking seriously precisely because it resists comfortable self-flattery: does the silence in question coexist with continued, active judgment and continued action in other registers, or has the judgment itself gone quiet along with the public reaction? Silence-as-discipline is compatible with — indeed depends upon — vigorous continued thought, continued private and trusted-circle conversation, and continued action through other channels; silence-as-withdrawal is typically accompanied by the same quieting across all registers, public and private alike.

Chapter 14: Burke Revisited — The Limits of “Good Men Doing Nothing”

14.1 Situating Burke’s Claim

Burke’s dictum, properly read in the context of his broader political writing, was addressed to a specific historical anxiety about revolutionary upheaval and the erosion of settled institutions through cumulative inaction by those positioned to defend them. It was not formulated with the bargaining-theoretic structure examined in Part Three in view, for the simple reason that Burke wrote more than a century and a half before Schelling. Applying Burke’s dictum to the specific situation this book examines — a situation in which the “doing something” being solicited may itself be the precise mechanism by which a provocateur’s strategy succeeds — requires an extension Burke’s own text does not supply, and this chapter attempts that extension carefully rather than simply assuming the dictum transfers without remainder.

14.2 The Compellence Problem, Revisited at the Level of the Individual Citizen

Chapter 9.1’s distinction between deterrence and compellence, developed there at the level of state-to-state bargaining, has a precise analogue at the level of the individual citizen’s response to a provocateur’s conduct. A provocation calculated to elicit a specific, visible, public reaction — the angry public statement, the participation in a particular form of visible protest, the conduct that can be photographed, dated, and later cited — functions, from the citizen’s position, exactly as a compellent demand functions from a state’s position in Chapter 9.1: its success is observable, and the citizen’s compliance with the solicited reaction is precisely the data point the

provocateur's broader strategy may be structured to harvest.

Where this is so — and only an examination of the specific situation can establish whether it is so in any given case — declining to supply the solicited reaction is not equivalent to declining to resist the underlying conduct being provoked against. It is closer to a recognition that resistance, like any other strategic interaction examined in this book, has its own bargaining structure, and that supplying exactly the move an opponent's strategy is built to harvest is not obviously equivalent to resisting that strategy, whatever its superficial appearance of defiance.

14.3 What Burke's Dictum Still Correctly Demands

None of the foregoing licenses treating Burke's warning as obsolete. What it demands — sustained, clear-eyed moral attention to what is actually happening, a refusal to let injustice become invisible through habituation, and continued action through whatever channels remain genuinely effective — remains exactly as binding under the analysis of this chapter as under Burke's own. What changes is only the identification of which specific behaviours constitute the "doing something" Burke's warning is actually concerned to demand. Chapter 13.4's diagnostic question is the bridge between Burke's enduring demand and this book's bargaining-theoretic refinement of it: the test of whether a given silence honours or betrays Burke's warning is whether moral attention and continued action through other channels persist alongside it, not whether the specific, solicited, public register of reaction is supplied.

Chapter 15: A Phenomenology of Dignified Non-Reaction

15.1 What "Dignity" Adds to "Strategy"

The analysis so far in Part Five has been framed primarily in terms of strategic efficacy — what kind of response serves, or fails to serve, the underlying values a person holds, given the bargaining structure described in Parts Three and Four. This chapter asks a related but distinct question: what is the lived experience, the phenomenology, of choosing silence-as-discipline well, as opposed to merely choosing it correctly in some external, third-person sense? A choice can be strategically correct and still corrosive to the chooser if its felt experience is one of suppression, humiliation, or self-betrayal rather than considered restraint.

15.2 The Stoic Resource, Used Carefully

The Hellenistic Stoic tradition — Epictetus's discipline of distinguishing what is and is not within one's control, and the broader Stoic practice of withholding assent from impressions rather than suppressing the impressions themselves — offers a phenomenological resource directly relevant here, used with appropriate care rather than imported wholesale.[30]

The Stoic distinction between an event and one's assent to a particular interpretation of that event maps with unusual precision onto this book's distinction between silence-as-discipline and silence-as-withdrawal: the Stoic practitioner does not pretend the provocation has not occurred, nor suppress the judgment that it is wrong; the practitioner

[30]Epictetus, Enchiridion, ch. 1, on the distinction between what is and is not within one's control, and Discourses, bk. I, ch. 1, on the discipline of withholding assent from impressions (trans. Elizabeth Carter, public domain; see also W. A. Oldfather's Loeb Classical Library translation for comparative reference).

withholds, specifically, the further assent that would convert the judgment into the precise emotional and behavioural reaction the provocateur's strategy, per Chapter 9.2, is calibrated to harvest.

This is not, on a careful reading of the Stoic sources, a counsel of passivity. The Stoic distinction between what is and is not within one's control is explicitly paired with vigorous, undiminished action with respect to everything that is within one's control[31]

— which, applied to this book's argument, includes precisely the continued judgment, continued private conversation, and continued action through other channels that Chapter 13.4 identifies as the marker distinguishing discipline from withdrawal.

15.3 Dignity as Self-Authorship

A further phenomenological resource, drawn more loosely from existentialist accounts of authentic choice, supplies what the Stoic framework alone does not fully capture: the felt difference between a silence experienced as imposed — a capitulation extracted by fear, indistinguishable from the inside from simple submission — and a silence experienced as chosen, an act the person can recognise, even in the act of performing it, as continuous with their own values rather than a betrayal of them. The practical difference this book is concerned to draw out is not primarily about outward behaviour, which may look identical in both cases to an external observer; it is about whether the person performing the silence can give themselves, honestly, the account given in Chapter 13.3 — that this is a chosen strategic and ethical response, not an evasion — and have that account survive their own scrutiny.

---

[31]Epictetus, Discourses, bk. I, ch. 1–4, on vigorous, undiminished engagement with what is within one's control as the necessary complement to non-engagement with what is not.

15.4 The Danger of Performing Dignity for Others

A final caution belongs in this chapter, in the same spirit as the limiting principles this book has insisted upon throughout (Ch. 4.5, 7.5, 12.4). The framework offered here can itself be misused — turned into a performance of stoic dignity for an audience, which reintroduces, at the level of the individual's own conduct, exactly the gap between performance and genuine disposition this book has spent its earlier parts analysing in rulers. A silence cultivated primarily to be admired as dignified by onlookers has drifted from silence-as-discipline back toward a different kind of performance, and the diagnostic offered in Chapter 13.4 — does judgment and action in other channels continue? — remains the most reliable available check against this drift as well.

Chapter 16: Ubuntu, Obuntu Bulamu, and Communal Resilience as Counter-Doctrine

16.1 Why a Communal Framework Belongs Here

The phenomenological resources examined in Chapter 15 — Stoic self-governance, existentialist self-authorship — share a common limitation worth naming directly: both are framed primarily at the level of the individual, examining what an individual owes themselves and how an individual might experience their own restraint well. The philosophical and ethical traditions of Ubuntu, and the Basoga articulation of communal mutual flourishing captured in the concept of obuntu bulamu, supply a resource this book has not yet drawn on and that corrects this individual-level framing directly: an account of resilience and dignity constituted between persons rather than achieved by a single person in isolation.[32]

16.2 The Communal Reframing of the Discount Problem

Ubuntu philosophy's central proposition — that personhood is constituted through relationship with others rather than achieved by a self-sufficient individual prior to relationship — offers a genuinely distinct resource for thinking about the bargaining dynamics examined in Parts Three and Four, rather than merely a different vocabulary for the same individual-level analysis.[33]

Where Chapter 8's formal model treats the discount problem as a calculation an individual rational actor performs alone, a community

[32]Mogobe B. Ramose, African Philosophy Through Ubuntu, rev. ed. (Harare: Mond Books, 2002), on personhood as constituted through relationship rather than achieved by a self-sufficient individual prior to relationship.
[33]Desmond Tutu, No Future Without Forgiveness (London: Rider, 1999), on Ubuntu as the recognition that a person's humanity is bound up in, and cannot be fulfilled apart from, the humanity of others.

organised around Ubuntu's relational premises is positioned to perform exactly the kind of distributed, cross-checked judgment that Chapter 6.3 identified as the historical condition under which the strategic and deficiency explanations actually become distinguishable: a single isolated observer, as Chapter 2.3 established, faces an information asymmetry that resists cheap resolution, but a community sharing observation, memory, and judgment across many members is structurally better positioned to triangulate toward an accurate account than any single member could manage alone.

16.3 Obuntu Bulamu as a Resource Against the Isolation the Doctrine Exploits

A further and more direct point follows from the analysis in Chapters 9 and 12. The madman doctrine's coercive power depends, as Chapter 8 established formally, on the target's uncertainty and, frequently, on the target's isolation — a population atomised into individual calculations of individual risk is considerably more susceptible to the doctrine's effects than a population whose members hold and continually cross-check judgment collectively, in the manner obuntu bulamu's ethic of mutual obligation and shared welfare specifically cultivates.[34]

This is not a claim that communal solidarity is a sufficient defence against the doctrine examined in this book, which would overstate what any single ethical or social resource can deliver; it is the narrower and more defensible claim that the specific vulnerability the doctrine exploits — isolated, individually-borne uncertainty — is precisely the vulnerability a genuinely lived Ubuntu ethic, as distinct from a merely rhetorical

[34]Isaac Christopher Lubogo, Ubuntu Constitutionalism in Africa: The Lubogo–Ubuntu Constitutionalism Index (LUCI) and the Ubuntu Legal Index (ULI): A Comprehensive Jurimetric Framework for African Constitutional Jurisprudence (Kampala: Suigeneris Publishers, 2026), on obuntu bulamu as a Basoga articulation of communal mutual flourishing and its jurimetric application to African constitutional jurisprudence.

invocation of it, is structured to reduce.

16.4 Synthesis Closing Part Five

The four chapters of Part Five together supply the ethical and phenomenological apparatus the book's explanatory chapters (Parts One through Four) deliberately withheld. Silence, properly disaggregated (Ch. 13), is neither uniformly complicit nor uniformly virtuous; Burke's enduring demand survives the disaggregation once correctly relocated (Ch. 14); the lived experience of choosing well, rather than merely choosing correctly, has its own resources and its own characteristic risk of self-deception (Ch. 15); and the individual-level framework of Chapters 13 through 15 finds its necessary communal complement in Ubuntu's relational account of personhood and obligation (Ch. 16), which corrects the isolation the doctrine's own logic depends upon. Part Six now turns to drawing these six parts together into a single general statement, and to the question of what jurisprudential and constitutional resources exist, imperfectly, to constrain the doctrine examined throughout.

PART SIX: SYNTHESIS AND JURISPRUDENTIAL IMPLICATIONS

Chapter 17: Toward a General Theory of the Performed Threat

17.1 Restating the Doctrine in Full

The argument of this book can now be stated as a single, integrated theory, each element of which has been established and qualified across the preceding sixteen chapters.

Performed unpredictability is a bargaining instrument available, in principle, to any actor whose position depends on extracting compliance or deterring challenge under conditions of contested or unconsolidated legitimacy (Ch. 1–2). It works by manipulating a target's probability estimate of follow-through on a threat that ordinary rational-actor reasoning would otherwise discount (Ch. 3, 8), exploiting a discount problem that cannot be fully resolved by the target because the actor has structural incentive and structural capacity to keep the staged and the genuine indistinguishable (Ch. 8.2). Both Machiavelli and Hobbes identify versions of this mechanism centuries before its formal statement, and both build, into their own accounts, explicit limiting principles — concentrated rather than continuous deployment, and the conditioning of the sovereign's authority on the protective function that authority was instituted to serve — that anticipate the doctrine's self-limiting character (Ch. 4–5). Schelling's twentieth-century bargaining theory supplies the rigorous formal apparatus that vindicates this classical intuition while specifying its precise structure, its asymmetric suitability to deterrent rather than compellent applications, and its character as a depleting rather than a costless resource (Ch. 7–9).

The doctrine clusters disproportionately around the structural position of the heir or contested claimant, for whom the ordinary, time-

dependent sources of legitimacy are foreclosed and for whom performed unpredictability offers one of few available instruments for generating, on a compressed timeline, a credibility that an established ruler accumulates over a career (Ch. 10–11). Its deployment carries costs not only to external rivals and to the wider public but to the very institutional apparatus the deploying actor's authority depends upon, where that apparatus comes to perceive the deploying actor's personal reputational strategy as a depletion of the apparatus's own distinct institutional stock of professional standing (Ch. 12).

Finally, the doctrine's targets are not without resources of their own. A disciplined, ethically serious silence — properly distinguished from mere withdrawal, and properly understood as compatible with continued judgment and continued action through other channels — denies the doctrine some portion of the very reaction its bargaining structure depends upon harvesting, and finds support both in the individual-level resources of the Western philosophical tradition and in the communal, relational resources of Ubuntu philosophy, which directly addresses the isolation the doctrine's coercive mechanism exploits (Ch. 13–16).

17.2 The Doctrine's Internal Contradiction, Stated Generally

Across every theoretical register this book has examined — Machiavellian dosage, Hobbesian conditional obligation, Schellingian reputational depreciation, and the institutional resentment mechanism of Chapter 12 — a single structural contradiction recurs, and this book's most general theoretical claim is that this recurrence is not coincidental. The doctrine purchases credibility precisely by making the actor's future conduct appear less governed by ordinary cost-benefit reasoning; but the actor's actual interest in retaining power, authority, or position remains, throughout, an entirely ordinary cost-benefit interest, however

well-disguised. The doctrine therefore requires the actor to act, with some consistency, as if unconstrained by the very calculation that motivates the entire strategy — and the gap between the performance and the motivating calculation is exactly what eventually produces the loss-of-control risk Schelling identified (Ch. 7.5), the institutional resentment Chapter 12 examined, and the historical pattern, visible across the case studies of Chapter 11, in which the doctrine's most severe failures attach to its most continuous and least disciplined practitioners.

17.3 A General Statement

The theory developed across this book can be compressed into a single general statement: performed unpredictability is a rational strategy for the rapid manufacture of bargaining credibility under conditions of contested legitimacy, available at a price that compounds with the frequency and duration of its use, structurally better suited to deterrence than to compellence, corrosive to the institutional apparatus the strategy's practitioner depends upon, and resistible, though not without cost, by a target capable of the specific ethical and communal discipline examined in Part Five. Nothing in this statement is original to any single chapter of this book; its originality, if any, lies in the integration of classical political philosophy, modern bargaining theory, and Ubuntu ethics into a single account, which is the project this book set out, in its Preface, to attempt.

Chapter 18: Constitutional and Legal Constraints on the Madman Doctrine

18.1 Why Law Belongs in This Book’s Final Movement

A book that has so far proceeded almost entirely through political philosophy and bargaining theory owes its reader, in closing, some account of how the constitutional and legal traditions familiar to this author’s own jurisdiction, and to comparable jurisdictions generally, attempt to constrain the doctrine examined throughout — not because law offers a complete solution to a phenomenon this book has shown to be deeply rooted in the structure of contested authority itself, but because law is one of the few institutional resources deliberately designed, across multiple constitutional traditions, with exactly this kind of unconstrained personal authority in view.

18.2 Constitutionalism as Institutionalised Distrust of the Unconstrained Will

Modern constitutionalism, in its general theoretical justification across many traditions, can be read as a sustained institutional response to precisely the Hobbesian problem examined in Chapter 5: if an unconstrained sovereign will is, as Hobbes argued, the price of escaping the state of nature, constitutionalism is the considered judgment of later political thought that the price is too high to pay without remainder, and that mechanisms — separation of powers, entrenched rights, independent judicial review, fixed and procedurally regular succession — can supply much of Hobbes’s stability without requiring acceptance of Hobbes’s unconstrained sovereign.[35]

[35]On the general theoretical justification for separation of powers as an institutional response to the risks of unconstrained sovereign authority, see Montesquieu, The Spirit of the Laws, bk. XI, ch. 6 (1748).

Every constitutional mechanism of this kind functions, among its other purposes, as a structural attempt to narrow the very gap between office and discipline examined in Chapter 2.1 — to make it institutionally costly, rather than merely reputationally costly, for office to be exercised without its corresponding discipline.

18.3 The Limits of Codified Constraint

This book's analysis throughout, and particularly Chapter 12's examination of institutional strain, suggests why codified constitutional constraint, however well-designed, cannot by itself fully resolve the phenomenon examined here. Constitutional mechanisms depend, for their actual operation, on institutional actors — courts, legislatures, military command structures — willing to invoke and enforce them, and Chapter 12 established that an apparatus's willingness to act against a position-holder's conduct is itself a variable shaped by exactly the reputational and bargaining dynamics this book has examined, not a fixed institutional given that operates independently of them. A constitutional system's formal mechanisms for constraining unconstrained authority are, in this sense, only as effective as the institutional actors charged with operating them remain willing to invoke them — which returns the analysis, in its final movement, to the same bargaining-theoretic structure with which it began: even the law's own enforcement is a strategic interaction subject to the dynamics of credibility, reputation, and institutional interest examined throughout this book.

18.4 Customary and Communal Law as a Complementary Resource

Consistent with Chapter 16's argument for the distinct contribution of Ubuntu philosophy, this book's final substantive observation is that

customary and communal legal traditions — which typically root authority's legitimacy directly in continuous relational accountability to a community, rather than solely in formal procedural compliance — supply a complementary rather than redundant resource alongside codified constitutionalism.[36]

Where codified constitutional constraint depends on formal institutional actors' willingness to act, customary accountability of the kind Ubuntu legal philosophy theorises depends on the continuous, distributed, communal judgment examined in Chapter 16.2 — a resource less easily captured or neutralised by exactly the institutional-capture dynamics that can compromise formal constitutional mechanisms, precisely because it does not depend on any single institutional choke point.

[36]Lubogo, Ubuntu Constitutionalism in Africa, on customary and communal legal accountability as rooted in continuous relational obligation to community rather than solely in formal procedural compliance.

Chapter 19: Conclusion — The Sovereign’s Mask and Its Limits

19.1 What This Book Has Argued

This book opened with a discomfort: the difficulty of distinguishing, from the outside, a ruler who has lost control from a ruler who has chosen to appear to. It has not resolved that discomfort, and could not, because the discomfort is not a failure of analysis correctable by better theory; it is, as Chapter 8.2 established formally, a structural feature of the situation itself, built into the doctrine by the very actors who benefit from its persistence. What this book has offered instead is a way of thinking clearly about the discomfort — a vocabulary, drawn from Machiavelli, Hobbes, and Schelling, for naming what is actually at stake in the uncertainty; a historical record, examined through settled and undisputed cases, for seeing the doctrine’s recurrence and its characteristic failure modes; and an ethical and communal framework, drawn from both Western and Ubuntu traditions, for those who must decide how to live, and resist, within reach of it.

19.2 The Mask Metaphor, Examined

The title of this book invokes a mask deliberately, and it is worth closing by examining the metaphor’s own limits rather than letting it stand unexamined. A mask, in ordinary usage, implies a true face beneath it — a settled, recoverable reality the mask merely obscures. This book’s argument has been more unsettling than that image suggests: the entire force of the doctrine examined throughout depends on the impossibility, from the target’s position, of confidently locating any such recoverable true face at all (Ch. 2.3, 8.2). The mask, on this book’s account, is not merely worn over a face; for the doctrine to

function as effectively as Parts Two and Three describe, the wearing must itself become, to outside observation, indistinguishable from the face — which is precisely why this book has insisted, from its Preface onward, that the relevant question is never what lies beneath the mask, but what it costs, and to whom, to keep the question permanently unanswerable.

19.3 A Final Word

This book has deliberately declined, throughout, to adjudicate whether any specific living actor’s conduct should be read through the strategic or the deficiency explanation introduced in Chapter 2.2, for the methodological and ethical reasons given in Chapter 1.3. That restraint has not been a limitation on the argument but a condition of its seriousness: a theory of structure is falsified by nothing so quickly as its author’s eagerness to apply it to whichever contemporary case the author finds most rhetorically convenient. What the theory offers, to a reader who encounters the doctrine in their own time and place, is not a verdict on any individual but a discipline of judgment — the capacity to ask, calmly and with the vocabulary this book has supplied, what incentive structure a given position actually creates, what the historical record suggests about that structure’s characteristic costs and limits, and what a dignified, communally rooted response to it might actually require. That discipline of judgment, exercised by the reader rather than supplied by this book in the reader’s place, is the only conclusion this book has ever intended to offer.

Bibliography

Aristotle. The Politics. Translated by Carnes Lord. 2nd ed. Chicago: University of Chicago Press, 2013.

Epictetus. Discourses and Enchiridion. Translated by Elizabeth Carter. London, 1758 (public domain; see also W. A. Oldfather, trans., Loeb Classical Library, Harvard University Press, 1925–28, for comparative reference).

Hobbes, Thomas. Leviathan. 1651. Edited by Richard Tuck, revised student ed. Cambridge: Cambridge University Press, 1996.

Lubogo, Isaac Christopher. Ubuntu Constitutionalism in Africa: The Lubogo–Ubuntu Constitutionalism Index (LUCI) and the Ubuntu Legal Index (ULI): A Comprehensive Jurimetric Framework for African Constitutional Jurisprudence. Kampala: Suigeneris Publishers, 2026.

Machiavelli, Niccolò. The Prince. 1532. Translated by Harvey C. Mansfield. 2nd ed. Chicago: University of Chicago Press, 1998.

Machiavelli, Niccolò. Discourses on Livy. Translated by Harvey C. Mansfield and Nathan Tarcov. Chicago: University of Chicago Press, 1996.

Millar, Fergus. The Emperor in the Roman World, 31 BC–AD 337. 2nd ed. London: Duckworth, 1992.

Montesquieu, Charles de Secondat, Baron de. The Spirit of the Laws. 1748. Translated by Anne M. Cohler, Basia C. Miller, and Harold S. Stone. Cambridge: Cambridge University Press, 1989.

Ramose, Mogobe B. African Philosophy Through Ubuntu. Rev. ed. Harare: Mond Books, 2002.

Schelling, Thomas C. The Strategy of Conflict. Cambridge, MA: Harvard University Press, 1960.

Schelling, Thomas C. Arms and Influence. New Haven: Yale University Press, 1966.

Tutu, Desmond. No Future Without Forgiveness. London: Rider, 1999.

Index

An index of key terms follows, keyed to chapter and section rather than to page number, consistent with this book’s internal cross-referencing throughout.

POWER NEEDS MORE THAN FORCE. SOMETIMES, IT NEEDS A MASK.

From ancient courts to modern statecraft, sovereigns have often walked the fine line between reason and madness. This book explores one of the most enigmatic doctrines in political and legal theory—the so-called Madman Doctrine—through a bold interdisciplinary lens.

The Sovereign's Mask traces the evolution of the idea that a ruler's feigned or real irrationality can become a strategic instrument of power. Drawing from philosophy, political theory, legal doctrine, and jurimetrics, Isaac Christopher Lubogo examines how performed irrationality can neutralize opposition, justify coercion, and reshape the moral architecture of governance.

A rigorous inquiry from Machiavelli to Schelling, this book challenges contemporary assumptions about leadership, legitimacy, and ethics in the exercise of sovereign power.

ISAAC CHRISTOPHER LUBOGO

LL.B, PGDLP, LL.M

Advocate | Doctor of Laws Fellow | Legal Scholar | Author | Researcher | Legal Innovator

Isaac Christopher Lubogo is a legal scholar, advocate, and researcher whose work explores the intersections of law, power, ethics, and political philosophy. He is committed to advancing justice, accountability, and constitutionalism through rigorous scholarship and public engagement.

SUIGENERIS PUBLISHERS

KAMPALA, UGANDA

ISBN 978-9913-9762-1-3

9 789913 976213