Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Beyond Obligation: Prospective Responsibility and the Structure of Moral Conflict

Morgan Jackson

mjac3670@alumni.sydney.edu.au

Freie Universität Berlin

Research Article

Keywords: moral responsibility, prospective responsibility, deontic logic, STIT, moral agency

Posted Date: November 24th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11178137/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Beyond Obligation: Prospective Responsibility and the Structure of Moral Conflict

Abstract

This paper argues that prospective responsibility is structurally distinct from moral obligation and that the distinction matters for formal moral reasoning. Whereas obligation concerns what morality requires of an agent, prospective responsibility concerns the outcomes for which an agent stands in a practically answerable relation. I identify four conditions of prospective responsibility—*autonomous agency*, *epistemic competence*, *practical feasibility*, and *temporal-contextual adaptability*—and show how they determine when an outcome falls within an agent's responsibility. The account also explains how multiple prospective responsibilities should guide practical reasoning when they cannot all be maximally realized. Rather than treating such conflicts solely as problems of ranking competing requirements, the paper conceives them as problems of responsible approximation: agents must seek achievable ways of preserving competing responsibilities within the constraints of their circumstances.

Keywords: moral responsibility; prospective responsibility; deontic logic; STIT; moral agency.

1. *Introduction*

When we judge that someone acted responsibly in navigating a moral conflict, we mean more than that they fulfilled whichever obligation ranked highest. We mean that they exercised practical judgment: they weighed competing commitments, considered their particular capacities and circumstances and found a course of action that honored those demands as far as the situation permitted. The formal resources analytic philosophy has developed for moral reasoning are, for the most part, obligation-centered. Deontic logic since von Wright (1951), its refinement through Horty's (2001) agency-relative framework, and Makinson and van der Torre's (2000) input/output logic all formalize what agents ought to do. Recent developments in deontic logic complicate any simple contrast between responsibility and obligation. Action-based deontic frameworks attach normative status directly to actions; STIT (seeing to it that) frameworks represent what agents bring about; epistemic extensions represent what agents know about their actions; and branching-time semantics supplies temporal structure. The question is therefore not whether existing logic can represent any one of the conditions that matter to responsibility. Rather the question is what is being represented when we say that an agent is responsible for an outcome, and how that relation behaves when several outcomes compete for the agent's attention. Although in this paper I develop a particular account of prospective responsibility, I am not offering a general theory of responsibility. Prospective responsibility, in the way I define it, concerns an agent's standing toward an outcome that remains to be realized, maintained, or managed.

The central claim is that prospective responsibility is not simply a more carefully restricted form of obligation. Obligation and responsibility answer different questions. An obligation specifies what a norm requires of an agent. Prospective responsibility specifies the outcomes for which an agent is practically answerable through their own agency. An agent may remain obligated when they lack the means to comply, when they do not understand the requirement, or when the effective opportunity

for compliance has passed. In none of these cases does the agent stand in a practical relation to the outcome on which their agency can bear. By contrast, the present account treats agency, epistemic competence, feasibility, and temporal position as constitutive conditions on prospective responsibility.

This claim should not be conflated with a claim of formal indefinability. I make no claim that an appropriately enriched deontic language could not represent the conditions developed below. Rather, an operator whose application is constitutively conditioned by these four dimensions represents a different relation from obligation, even if the operator is implemented within an otherwise deontic formalism. The issue is therefore one of conceptual architecture and extension rather than the expressive power of a particular notation.

The paper advances two interconnected claims. First, prospective responsibility is a normative relation concerning an agent's practical standing toward an outcome, rather than the binding force of an independently specified norm. This difference is constituted by four conditions: autonomous agency, epistemic competence, practical feasibility, and temporal-contextual adaptability. The claim is not that these conditions could not be represented within an enriched deontic system. It is that, once these conditions are constitutive of whether an agent stands in the relevant relation to an outcome, the object being represented is different from an obligation understood as the binding force of a norm. Second, I show that this distinction has structural consequences for practical reasoning. In particular, it changes how responsibility is constrained by agency and feasibility, how responsibilities persist and interact over time, and how an agent should reason when responsibilities are practically incompatible.

The central contribution is therefore not the introduction of a formally inexpressible new modality. It is the identification of a distinctive object of normative assessment: an agent's standing toward an as-yet-unsettled outcome under conditions that make responsible practical engagement

possible. The formal notation is used to make the structure of that relation explicit, while leaving open the question of how a complete semantics should be constructed.

The paper proceeds as follows. Section 2 establishes the distinction and introduces the responsibility operator. Section 3 develops the four conditions and their structural consequences. Section 4 compares the framework with existing approaches; Section 5 applies it to an intelligence officer's dilemma; and Section 6 identifies remaining formal limitations.

2. Obligation

Obligation, in its core philosophical usage, refers to what morality requires of an agent, where the authority of the requirement can be specified independently of the agent's particular deliberative situation.[1] The linguistic and logical structure of obligations can be expressed through deontic operators. O(φ) encodes a relation between an agent and a norm; where what matters is that the requirement binds a particular agent, I write O_A(φ). O_A(φ) says that A is required to see to φ. It says nothing about whether A can, whether A understands what compliance involves, or whether the opportunity for compliance remains open. Obligation is therefore norm-relative even when indexed to an agent. I return to this at §4.4, where the objection I address is directed at O_A rather than O.

First, deontic operators are norm-relative: they express what norms require without encoding the agent's relationship to those requirements, their understanding, their practical situation, or their deliberative engagement. Second, they are structurally static: O(φ) holds or fails to hold, without resources for modelling how obligations interact with an agent's evolving circumstances over time. Third, they generate well-known conflict paradoxes (Chrisholm 1963; Prior 1954; Hansen 2001). Two qualifications are important here, since they shape the comparison developed in Section 4. The first is that these three features characterize standard deontic logic and are not evenly distributed across the agentive frameworks that have developed from it. Horty's (2001) dominance-based deontic STIT indexes obligation to an agent's available choices, so it is not norm-relative in the first sense. Moreover, it addresses conflict, though by classification rather than resolution. Shi's (2024) account of agentive permissions makes actions and their deontic status objects of the

[1] Scanlon (1998: 248–294; 2008: 122–214) distinguishes substantive responsibility from attributive responsibility. Watson (1996) draws a related distinction between the attributability and accountability faces of responsibility. Shoemaker (2024) develops a pluralistic account that further subdivides these categories. The present distinction between obligation and responsibility crosscuts these taxonomies: it concerns the structural properties of the moral concepts rather than the conditions under which they apply.

semantics directly. The second qualification is that even the frameworks that address some of these features do so within an obligation-centered paradigm that leaves the agent's autonomous engagement with competing demands unrepresented.

2.1 Responsibility and the Responsibility Operator

Prospective responsibility, as I use the term, denotes an agent's standing in relation to an outcome that remains within the practical field of their agency. To say that an agent is responsible for φ is not merely to say that a norm requires the agent to bring φ about. It is to say that φ falls within the domain of outcomes for which the agent has undertaken, or has otherwise acquired, a practical responsibility and for which the agent can be expected to exercise judgment in light of the conditions that govern their agency. I represent this relation by R_A(φ). R_A is a unary sentential operator. It takes a sentence and yields a sentence; it binds no variables and is therefore not a quantifier.

An agent may be responsible for pursuing an outcome even though no action available to the agent guarantees that the outcome will occur. The agent may instead have an action that keeps the outcome open, advances it, or preserves the possibility of its realization. On this point, Shi (2024) distinguishes actions that ensure an outcome from actions that merely admit of it, and Shi and Naumov (2025) elucidate why the distinction becomes especially important in multi-step decision schemes. The present account adopts that distinction rather than treating responsibility as requiring unilateral guarantee. The operator therefore does not itself mean "A ought to bring about φ". It records that φ falls within A's prospective responsibility. Questions about what A ought to do in light of R_A(φ), including what A should do when several responsibilities compete, are further questions. The distinction becomes clearer when we consider two cases in which one concept applies without the other. First, obligation can survive the loss of practical capacity. For example, an impoverished debtor may remain obligated to repay a debt even when they currently lack the means

to do so. The obligation does not disappear merely because the opportunity for compliance has narrowed. The present account would not infer from that fact alone that the debtor bears prospective responsibility for ensuring repayment. Second, responsibility can arise from an undertaking or role before a particular obligation concerning the outcome is independently specified. An employee who voluntarily assumes responsibility for maintaining the integrity of a laboratory's records may acquire a prospective responsibility for that outcome even though no independently stated moral norm specifies that this particular employee must maintain it. The undertaking may subsequently generate obligations, but the responsibility relation is not identical to those derivative obligations. Neither case is novel in the tradition. Kant's doctrine of imputation already distinguishes what morality requires from what can be ascribed to an agent as their deed.[2] The present paper does not claim that Kant anticipated the four-condition framework, but the distinction shows that the normative status of a requirement and the conditions under which conduct is attributable to an agent need not be treated as identical.

The logical point is therefore not that responsibility excludes obligation. Frequently both will obtain: $R_A(\varphi) \wedge O_A(\varphi)$. Nor is the claim that responsibility is never normatively demanding. It is that the two modalities are not coextensive. In schematic form, the paper rejects both $R_A(\varphi) \leftrightarrow O_A(\varphi)$ and any simple reduction of R_A to a conjunction of deontic claims. Several notational conventions will be observed throughout. Actions(A) denotes the set of actions available to agent A. I write $x \rightsquigarrow \varphi$ for the relation in which action x ensures φ, so that φ obtains whatever the other agents do, and $x \dashrightarrow \varphi$ for the relation in which x admits of φ, leaving open the possibility that φ obtains. The arrow → is reserved for material implication. Ensuring entails admitting, and the two coincide in single-agent deterministic settings (Shi 2024: 3558). Two control notions follow, indexed

[2] Kant's doctrine of *Zurechnung*, developed across the *Metaphysics of Morals* (6:227–228), distinguishes between what the moral law requires (*Pflicht*) and what can be ascribed to an agent as their deed (*imputatio*).

by the relation from which each is built: $C^{\rightsquigarrow}_A(\varphi)$ holds when some action available to A ensures φ, and $C^{\dashrightarrow}_A(\varphi)$ when some action available to A admits of φ. $C^{\rightsquigarrow}_A(\varphi)$ implies $C^{\dashrightarrow}_A(\varphi)$, and the converse fails. The framework does not treat either as the control relation.

2.2 Prospective Responsibility in the Existing Literature

The distinction developed here belongs to an established family of forward-looking conceptions of responsibility. Cane (2002) distinguishes prospective responsibility from historic responsibility, associating the former with roles, tasks, and future-directed requirements, while Goodin (1995) characterizes task responsibility in terms of seeing to it that a specified outcome is brought about. Van de Poel (2011) likewise distinguishes forward-looking from backward-looking responsibility, treating the former as directed toward what an agent is responsible for bringing about rather than toward the appraisal of what has already occurred. These accounts establish an important point on which the present proposal builds: responsibility need not be exhausted by retrospective questions of blame, accountability, or causal contribution.

The present account nevertheless differs in emphasis. The question here is not simply which future-directed tasks or outcomes are assigned to an agent, but what it is for an agent to stand in a prospective responsibility relation to an outcome. That question is sensitive to the conditions under which the agent can responsibly exercise practical agency. A role or task may identify what an agent is responsible for in a broad forward-looking sense without by itself determining whether the agent presently possesses the agency, epistemic competence, practical means, and temporal opportunity required to stand in that relation at a particular time. The four conditions developed below are intended to supply that missing structure. This also places the present proposal alongside Vincent's (2011) taxonomy of responsibility concepts. Vincent's project demonstrates the importance of distinguishing different uses of 'responsibility' rather than treating the concept as univocal. The

present paper pursues a more specific question: given a future-directed outcome, what conditions make it appropriate to say that an agent presently bears responsibility for pursuing, maintaining, or managing that outcome? The answer proposed here is not a further taxonomic label. It is a structural account of the conditions under which the prospective relation obtains and of the practical consequences that follow once it does.

2.3 Positioning: Scanlon's taxonomy

Scanlon (1998) distinguishes responsibility as attributability from substantive responsibility. According to Scanlon, attributability concerns "whether some action can be attributed to an agent in the way that is required in order for it to be a basis for moral appraisal" (1998: 248). That is to say that a person is responsible in this sense "is only to say that it is appropriate to take it as a basis of moral appraisal of that person" (Scanlon 1998: 248). Substantive responsibility expresses "substantive claims about what people are required (or, in this case, not required) to do for each other" (Scanlon 1998: 248), claims that are sensitive to the choices the agent has made and the opportunities they have had.

R_A(φ) runs laterally to Scanlon's taxonomy rather than mapping onto one side of it. It incorporates conditions he associates with attributability — the agent must possess the kind of self-governance that makes moral appraisal appropriate, and must act from judgment-sensitive attitudes rather than from compulsion or manipulation — while also making what Scanlon would recognize as a substantive claim about what the agent is committed to pursuing. When we say that an intelligence officer bears responsibility for both operational security and legal compliance, we are saying two things at once. Their subsequent actions can properly be appraised as reflecting their practical judgment. And they bear a particular practical relationship to those outcomes, one that depends on their capacities, their knowledge, and the choices available to them. The four conditions developed in the next section are a specification of those conditions. The deontic operator records

whether an independently specified norm binds the agent; $R_A(\phi)$records whether the agent has governed themselves in a manner adequate to the moral demands they face, where adequacy is specified by the four conditions

2.4 Summary of the structural differences

R_A(φ) is indexed to the agent's practical capacities and circumstances in a way that O_A(φ) is not: it is relativized to what this agent, with their particular capacities and circumstances, can ensure or admit of. It is epistemically conditioned: it presupposes that the agent understands what φ requires and can reason about its moral implications. It is temporally dynamic: the agent's responsibility for φ may obtain during some intervals and not others, depending on the evolution of the agent's practical situation. And it carries internal resources for conflict resolution: when two responsibilities prove practically incompatible, the framework generates a structured account of what responsible practical reasoning demands. These differences generate structural consequences of their own.

3. Four Conditions and Their Structural Consequences

This section argues for four conditions that responsibility imposes on moral agents and traces the structural consequences that follow from each. The conditions are defended rather than stipulated: for each, I argue that it marks a genuine difference between what responsibility requires of an agent and what obligation requires. What the conditions do is motivate and constrain: each consequence states, in surveyable form, something that must hold if the condition it answers to is genuine. Section 3.5 sets out the mapping.

3.1 The Condition of Autonomous Agency

A bears responsibility for φ only if A maintains influence over φ's potential outcomes, capacity for reflective deliberation about φ, and ability to modify actions regarding φ based on such deliberation.

Why does responsibility require this and obligation not? Because obligations can bind agents who lack control, deliberative capacity, or the ability to modify their behavior accordingly. An agent can be obligated to repay a debt even when they presently lack the means to do so; the obligation persists, generating an "ought" that may be unfulfillable in the agent's current circumstances. We would not say that the agent is responsible for repaying the debt, in the sense of standing in a practical relation to repayment on which their agency can bear. The concept of responsibility builds the agent's practical situation into the conditions of its correct application, and the concept of obligation does not. The responsibility literature likewise connects responsible agency with guidance, reflective endorsement, and higher-order practical reasoning (Frankfurt 1969; Fischer and Ravizza 1998; Watson 1975; Bratman 1987, 2007).

These accounts converge on the insistence that responsibility demands more of the agent's deliberative involvement than the mere existence of a norm that applies to them. Scanlon's analysis develops this point. Moral criticism, according to Scanlon, is appropriate only when an action reflects the agent's judgment-sensitive attitudes — attitudes that are "under the control of reason" and that agent can "in principle be called on to defend with reasons and to modify them if an appropriate defense cannot be provided" (1998: 272). Scanlon's discussion of cases such as hypnosis and sleepwalking likewise shows that moral appraisal can be suspended when the connection between action and judgment-sensitive attitudes is severed (1998: 277). The *Autonomous Agency* condition requires that the agent's relationship to φ be mediated by judgment-sensitive attitudes and reflective deliberation, rather than by causal processes that happen to issue in φ's obtaining. This is a richer conception of agency than causal efficacy, and it is the conception that underwrites attributions of moral responsibility as distinct from causal responsibility. Two structural consequences follow:

Control Requirement. $R_A(\varphi) \rightarrow C^{\wedge}\dashrightarrow_A(\varphi)$, where $C^{\wedge}\dashrightarrow_A(\varphi) =_{df} \exists x \in \text{Actions}(A) : x \dashrightarrow \varphi$.

If an agent bears responsibility for φ, then some action available to the agent leaves the realization of φ open. The weaker relation is the right one here, and not as a concession to hard cases. An agent embedded in an institution, whose outcomes depend on the responses of colleagues, superiors, and oversight bodies will rarely have an action that guarantees any morally significant outcome. Stated with C^⇝_A, the *Control Requirement* would leave such an agent bearing almost no prospective responsibilities, which is the result §2.1 rejects. C^⇝_A(φ) is the closer analogue of the STIT operator [α stit: φ] (Belnap, Perloff, and Xu 2001), which registers the agent's constraining the future course of events to a definite subset of histories. The fact that responsibility requires only C^⇢_A therefore marks a difference from STIT agency, over and above the further conditions that STIT leaves unspecified. *Consistency* ¬(R_A(φ) ∧ R_A(¬φ)). An agent exercising reflective deliberation cannot coherently commit to contradictory ends.

Standard deontic logic rules out O(φ) ∧ O(¬φ) by axiom: the D schema, O(φ) → ¬O(¬φ), is part of the system. For responsibility, consistency is a consequence of what responsibility is. An agent who reflectively endorses a commitment to φ while reflectively endorsing a commitment to ¬φ has not encountered a moral conflict; they have failed a condition on holding either commitment. *Autonomous agency* and *epistemic competence* do the work jointly: deliberation adequate to the moral situation, informed by adequate understanding of what the commitments require, cannot terminate in a pair of commitments that no action could serve at any level of approximation. This is why the framework needs a separate treatment of practical incompatibility, which I develop in §3.3. The cases that matter morally are those in which φ and ψ are jointly unrealizable by this agent without being contradictory.

3.2 The Condition of Epistemic Competence

For any φ for which A bears responsibility, A must possess an adequate epistemic relationship to φ: A must have sufficient understanding of φ's potential consequences, sufficient cognitive capacity to

reason about its moral implications, and sufficient awareness of how φ relates to the broader reasons and principles relevant to the situation. Epistemic Competence therefore concerns not whether φ is achievable, but whether the agent can understand what pursuing φ involves and can be expected to recognize the morally relevant relations among the considerations that bear on it. Fischer and Ravizza (1998: 41–82) develop the point through their analysis of moderate reasons-responsiveness, which requires that the mechanism issuing in action exhibit "regular receptivity to reasons," an "understandable pattern" of reasons-recognition that is "minimally grounded in reality" (1998: 71, 73). The agent must recognize how reasons fit together, grasp why one reason is stronger than another, and understand how accepting one reason as sufficient implies that a stronger reason must also be sufficient. This is a structured cognitive relationship to reasons rather than an accidental sensitivity to one of them. How demanding this relationship must be is the subject of a productive dispute among Rosen (2003), FitzPatrick (2008), and Zimmerman (1997), one that bears directly on the condition's content. Rosen defends what he calls a "Parity Thesis": whenever an agent acts from ignorance, whether factual or moral, the agent is culpable for the action only if culpable for the ignorance from which they act (2003: 64).

Rosen's position allows blameless moral ignorance to excuse when the agent is not culpable of the ignorance. If the slaveholder's moral ignorance is genuinely blameless, blame is unfair. The underlying principle is one of reasonable expectation: "It is unfair to blame someone for doing something if he blamelessly believed that there is no compelling moral reason not to do it" (2003: 75). Zimmerman (1997:415-6) argues that culpability for ignorance is indirect, since ignorance must be remedied through further action. FitzPatrick (2008) resists the collapse of the epistemic condition into skepticism. He accepts that ignorance can excuse, and takes the slaveholder case to establish it, while arguing that Rosen and Zimmerman are too permissive about when ignorance qualifies as blameless. FitzPatrick's key move is to identify a category of culpable normative ignorance that does

not require akrasia at any point: ignorance resulting from the agent's failure to fulfil epistemic obligations they had the "personal capabilities" and "social opportunities" to meet (2008: 590). FitzPatrick therefore accepts that ignorance can excuse but argues that normative ignorance may nevertheless be culpable where agents fail to use epistemic capacities and opportunities available to them.

All three authors agree on the structural point that responsibility requires a substantive epistemic relationship between agent and action. Where they disagree is about how demanding the relationship must be. The *epistemic competence* condition is formulated at a level of generality that accommodates all three positions — requiring "adequate understanding" and "sufficient cognitive capacity," and leaving the specification of adequacy and sufficiency as a substantive question the framework structures without foreclosing. What the condition does settle is that epistemic considerations are constitutive conditions on whether responsibility obtains, rather than external mitigating factors to be weighed once responsibility has been established. One structural consequence follows. *Restricted entailment*. Responsibility extends from φ to ψ when every action by which the agent could realize φ also realizes ψ. Therefore, ψ is not settled independently of what the agent does, and the agent can be expected to recognize that every action by which they could realize φ also realizes ψ:

$$R_A(\varphi) \land C^{\wedge}\!\!\rightsquigarrow\!_A(\varphi) \land [\forall x \in \text{Actions}(A) : (x \rightsquigarrow \varphi) \rightarrow (x \rightsquigarrow \psi)] \land [\exists y \in \text{Actions}(A) : y \not\rightsquigarrow \psi] \land K_A(\varphi \rhd_A \psi) \rightarrow R_A(\psi)$$

where $\varphi \rhd_A \psi$ abbreviates the universally quantified conjunct and K_A marks the requirement that A can be expected to recognize it. This consequence takes the ensuring relation where the *control requirement* takes the admitting relation, and the asymmetry is deliberate. Practical inheritance depends on ψ arriving with φ along every route by which the agent could bring φ about. An action that merely admits of φ carries no corresponding guarantee about ψ, since whether either outcome

eventuates then turns on what other agents do, and responsibility for φ would extend to outcomes that might never accompany it. The quantifier therefore ranges over the agent's ensuring actions. The second conjunct requires that this range be nonempty. Where the agent has no action that ensures φ, the universally quantified conjunct is vacuously satisfied, and responsibility would extend to any outcome meeting the remaining conjuncts. C^⇝_A(φ) blocks that. This restriction is substantive: practical inheritance is available only when the agent can guarantee the outcome that grounds the inherited responsibility. This is a narrower class of cases than those in which the agent bears a prospective responsibility at all. Where the agent has only actions that admit of φ, nothing is inherited.

The universally quantified conjunct replaces logical consequence with practical inheritance. What matters is that ψ comes with φ along every route by which this agent could bring φ about, rather than that φ entails ψ. A formulation using material implication would be far too permissive, since φ → ψ is true whenever ψ is true: any true outcome within the agent's influence would then inherit responsibility from any responsibility the agent already bears. Strengthening the implication to necessity would remove that failure and leave another, since □(φ → ψ) holds whenever ψ is a necessary truth, so responsibility would still extend to tautologies. Quantifying over the agent's available actions avoids both, because it makes the extension of responsibility depend on the structure of what the agent can do rather than on relations among propositions. The existential conjunct blocks what remains. An outcome that obtains whatever the agent does is not one the agent's commitment can bear upon, and responsibility does not extend to it. A condition of the same shape appears in a different setting: the deliberative STIT carries a negative condition requiring that the outcome not be settled independently of the agent's choice (Horty and Belnap 1995). The convergence is not accidental. Both conditions register that agency is a matter of making a difference. The epistemic conjunct puts the epistemic condition to work. If responsibility extends to

ψ only through the agent's practical relationship to φ, and if responsibility requires the epistemic relationship argued for above, then the extension must be one the agent can recognize.

Without this conjunct, responsibility would extend to outcomes that accompany φ unavoidably and invisibly. I leave K_A deliberately weak: the connection must be one the agent had the capacity and opportunity to recognize, which is FitzPatrick's standard rather than Zimmerman's, though nothing in the structural claim turns on settling that dispute. Where the agent has an action that ensures φ, the universally quantified conjunct says that this action ensures ψ as well.

3.3 The Condition of Practical Feasibility

For any φ for which A bears responsibility, φ must be achievable within A's context, there must exist concrete mechanisms through which A can realize φ, and A must have genuine capacity to influence φ's outcomes. This condition may appear to overlap with Autonomous Agency, though it makes a distinct point. Autonomous Agency concerns the agent's deliberative and causal relationship to φ. Practical Feasibility concerns whether φ itself is the kind of thing that can be brought about in the relevant circumstances. An agent may possess full deliberative capacity and causal control in the relevant sense while the specific outcome they are committed to remains practically unachievable, whether through institutional constraint, resource limitation, or the structure of the situation. The familiar principle that ought implies can provides a useful point of comparison, but the present argument does not depend on treating it as a general constraint on obligation. The claim here is narrower: if prospective responsibility concerns an agent's active practical engagement with an outcome, then the outcome must fall within the range of what the agent can in the relevant circumstances bring about or influence. Responsibility, understood as active moral engagement, cannot coherently attach to outcomes that are practically impossible.

Two structural consequences follow:

Possibility: $R_A(\varphi) \rightarrow \Diamond\varphi$. If an agent bears responsibility for φ, then φ is achievable in at least

one accessible possible world.

Conjunction: (R_A(φ) ∧ R_A(ψ) ∧ ¬⊥^⇢_A(φ, ψ)) → R_A(φ ∧ ψ). When an agent bears responsibility for two states of affairs and some available action admits of both, the agent bears responsibility for their conjunction.

The weaker relation is what the Control Requirement demands of the conjunction, and requiring an action that ensures both would ask more of the conjunction than the framework asks of either conjunct.

The framework employs two relations of practical incompatibility, each of which must be distinguished from logical contradiction:

Definition. ⊥^⇝_A(φ, ψ) =df ¬∃x ∈ Actions(A) : ((x ⇝ φ) ∧ (x ⇝ ψ)).

Definition. ⊥^⇢_A(φ, ψ) =df ¬∃x ∈ Actions(A) : ((x ⇢ φ) ∧ (x ⇢ ψ)).

φ and ψ are ensuring-incompatible for A when no action available to A ensures both, and admitting-incompatible when no action available to A admits of both. Since ensuring entails admitting, ⊥^⇢_A(φ, ψ) entails ⊥^⇝_A(φ, ψ), and the converse fails. Admitting-incompatibility is the more demanding condition on the agent's action-set, and it is also the condition under which the agent has fewer resources: an agent for whom φ and ψ are ensuring-incompatible may still have an action that admits of both, while an agent for whom they are admitting-incompatible has none. The gap between the two is where responsible practical reasoning under conflict does its work.

Propositions may be logically compatible, in that there exist possible worlds in which both obtain, while being practically incompatible in either sense for a given agent whose available actions contain no element realizing both. These definitions ground the framework's treatment of moral conflict in the realities of agency rather than in abstract logical relations, and they capture a distinction that obligation-centered frameworks working at the level of propositions do not draw:

the difference between norms that conflict because they are logically inconsistent and demands that conflict because no available action satisfies both.

3.4 The Condition of Temporal-Contextual Adaptability

Agent A's responsibility for φ requires the capacity to reason about φ across relevant time intervals and to maintain a principled practical orientation toward φ as circumstances change. Temporal-contextual adaptability concerns the agent's capacity to revise the means by which φ is pursued in response to changes in information, available options, and competing demands, without thereby abandoning the underlying commitment that grounds responsibility for φ. It does not require that every conflict be successfully resolved. Rather, it provides the temporal and practical capacity within which conflicts among responsibilities can be recognized, managed, and, where possible, resolved. Consider a physician's responsibility for a patient's welfare. This responsibility does not consist in a single obligation holding timelessly. It involves ongoing assessment of the patient's condition, adaptation of treatment plans as new information emerges, management of competing demands, and maintenance of a principled commitment to the patient's interests through all of these changes. Three structural consequences follow:

$$\textit{Temporal Bounds}\text{: } R_A(\varphi)@t \rightarrow \exists[t_1,t_2] : t \in [t_1,t_2] \wedge C^{\wedge}\dashrightarrow_A(\varphi)[t_1,t_2].$$

Time-indexed responsibility for φ at t requires an interval containing t across which the agent maintains practical influence over φ. This ensures that responsibility is not ascribed when the window for effective action has closed, nor when the agent's capacity to influence outcomes has not yet obtained.

The conditional runs in one direction only, and deliberately so. Having influence over φ across an interval is not sufficient for bearing responsibility for φ during it; the other three conditions must hold as well. Reading the consequence as a biconditional would make control sufficient for responsibility, which would collapse the distinction between moral and causal responsibility argued

for in §3.1 and would undercut the comparison with STIT developed in §4.1. Stability. Call a change in an agent's circumstances condition-preserving with respect to φ when the agent's situation after the change still satisfies all conditions on R_A(φ):

1. the agent retains influence over φ, in that C^⇢_A(φ) continues to hold;
2. the agent retains the capacity to deliberate about it;
3. the agent retains adequate understanding;
4. φ remains achievable;
5. the window for effective action has not closed.

Writing w ≈_φ w' for the relation that holds when w' differs from w only by condition-preserving changes, the consequence is this:

if R_A(φ) at w and w ≈_φ w', then R_A(φ) at w'

As stated, the consequence is close to a definitional truth, and the philosophical work lies in the classification rather than in the statement. The substantive claim is that a specified range of changes are condition-preserving: increases in the social cost of discharging a responsibility, reputational exposure, and inconvenience. None of these alter whether the agent can act, understands what is at stake, or has time to act; they alter what acting costs. Obligation-centered frameworks have no analogous distinction available, since O(φ) does not encode the conditions whose preservation is at issue. This is a stability claim rather than a necessity claim: a necessity claim, R_A(φ) → □R_A(φ), would render responsibility an all-or-nothing property whenever it obtains, collapsing the modal distinctions a dynamic account requires. Pressure can become severe enough to remove options rather than to price them. An agent threatened with professional ruin for disclosure may still be able to disclose at great cost; the change is condition preserving and the responsibility stands. An agent placed under surveillance that makes disclosure impossible has lost practical influence: Practical Feasibility fails, and the responsibility lapses. What matters is that the framework locates the

question in the right place: whether the change defeats a condition, rather than whether it makes compliance harder.[3]

Conflict Resolution: (R_A(φ) ∧ R_A(ψ) ∧ ⊥^⇝_A(φ, ψ) ∧ ¬⊥^⇢_A(φ, ψ)) → R_A(Resolve(φ, ψ)). Two conditions trigger the consequence. Ensuring-incompatibility establishes that the agent cannot fully discharge both commitments. The negated admitting-incompatibility establishes that some available action admits of both, which is what makes resolution by approximation possible at all. The gap between the two relations, identified in §3.3, is therefore the condition under which the consequence fires. The side condition is required rather than optional. Without it, an agent for whom φ and ψ are admitting-incompatible would acquire R_A(Resolve(φ, ψ)) while no available action met the characterization below, and the framework would generate a responsibility that no exercise of the agent's agency could discharge. Possibility rules such a responsibility out, and the objection to obligation-centred treatments of dilemma turns on obligation's willingness to retain demands of exactly that kind. Obligation-centred frameworks have no corresponding restraint, since an obligation survives the closing of the opportunity to comply, which is why a dilemma there remains a dilemma however far the agent's options contract.

The Resolve function is not a decision procedure that mechanically generates unique solutions. When obligations conflict, the standard move is to rank them (Ross 1930), apply a threshold principle, or classify actions as optimal or suboptimal through dominance reasoning (Horty 2001). In each case the resolution comes from outside the obligation framework: supplementary principles are imported to do work the deontic operator cannot do on its own. The Resolve characterization draws on the same four conditions that define responsibility in the first place. *Resolve*(φ, ψ) denotes the

[3] The distinction between stability and necessity bears on debates about the persistence of moral obligations under changing circumstances. Wallace (2019) argues that moral obligations exhibit a relational structure sensitive to context; the stability consequence offers an account of how responsibility persists through contextual change while remaining defeasible under genuine shifts in the conditions that ground it.

proposition that A performs an action α satisfying the conditions below, so that R_A(Resolve(φ, ψ)) remains a sentence and R_A remains the unary sentential operator of §2.1. The action α is a responsible resolution of a conflict between φ and ψ when:

1. α ∈ Actions(A) (Autonomous Agency);
2. the agent possesses adequate understanding of α's consequences (Epistemic Competence);
3. α is practically achievable (Practical Feasibility); and
4. α best approximates both φ and ψ, *where best approximates means that there are nearest jointly achievable versions* of φ and ψ, call them φ' and ψ', such that α ⇢ φ' and α ⇢ ψ'.

An action that ensures both approximations satisfies this *a fortiori*; admitting is used because conflict may preclude guaranteeing even the approximations. The characterization does not guarantee uniqueness, since 'nearest jointly achievable' versions are not yet metrically specified. *Conjunction* and *conflict resolution* may apply simultaneously to the same pair of responsibilities, and each then yields a distinct consequence. Suppose R_A(φ) and R_A(ψ) hold, some available action admits of both, and none ensures both. *Conjunction* yields R_A(φ ∧ ψ), recording the standing; *conflict resolution* yields R_A(Resolve(φ, ψ)), recording what discharging it requires.

Where φ and ψ are admitting-incompatible, neither consequence fires. Conjunction's side condition fails, and Conflict Resolution's fails with it, since the two share the requirement that some available action admit of both. Such a case is not one in which responsible reasoning has an answer the agent has failed to find; no approximation of both is available, and the framework issues no further responsibility. The shared side condition also keeps both consequences clear of *consistency*: φ

and ¬φ are admitting-incompatible for any agent, so neither R_A(φ ∧ ¬φ) nor R_A(Resolve(φ, ¬φ)) is ever generated, and Possibility would exclude both independently.

3.5 The Architecture: Conditions and Consequences

Each consequence is traceable to one or more conditions. *Practical feasibility* motivates *possibility*: responsibility requires practical achievability. *Autonomous agency* and *epistemic competence* jointly motivate *consistency*: reflective deliberation with adequate understanding precludes contradictory commitments. *Autonomous agency* and *practical feasibility* jointly motivate *conjunction*: influence plus feasibility yields conjunctive responsibility. *Conjunctio*n and *conflict resolution* address the same conflicts from two directions, the first recording the standing and the second what discharging it requires. *Autonomous agency* and *epistemic competence* jointly motivate *restricted entailment*: responsibility extends only to outcomes that come with φ along every available route and that the agent can recognize as doing so. *Autonomous Agency* motivates the *control requirement*. *Temporal-contextual adaptability* motivates *conflict resolution* and *temporal bounds*, and, in conjunction with the other three, *stability*.

4. What Responsibility-Centered Analysis Adds

The structural consequences identified above would be of merely theoretical interest if existing frameworks already accommodated them. This section argues that they do not, engaging in turn with agentive deontic frameworks (§4.1), including recent action-based deontic frameworks (§4.1.1) and epistemic extensions of STIT (§4.1.2), Fischer and Ravizza's guidance-control account (§4.2), and the formal literature on retrospective and counterfactual responsibility (§4.3), before addressing directly the question whether an obligation-centered framework could absorb the four conditions (§4.4).

4.1 Agentive Deontic Frameworks

Horty's framework represents the most rigorous attempt to formalize agent-relative moral reasoning within an obligation-centered paradigm. By embedding the STIT operator, which captures an agent's causal role in guaranteeing an outcome, within a dominance-based deontic logic, Horty provides a powerful apparatus for analyzing what agents ought to see to. To that end, Horty's treatment of agency abstracts from the agent's epistemic state, and he is explicit about this: "we idealize in three ways: by ignoring any intentional components that might be involved in the concept of action, by ignoring vagueness and probability, and by treating actions as momentary in duration" (2001: 12). On his account, [α cstit: A] holds at an index m/h just in case Choice^m_α(h) ⊆ |A|_m (2001: 15–16): the agent's action guarantees the truth of A. No condition is imposed on whether the agent understands that their action carries this guarantee, is aware of its consequences, or possesses the cognitive resources to evaluate its moral significance. An agent who unknowingly guarantees the truth of A through their choice satisfies [α cstit: A], and we do not hold them responsible in the same way as an agent who acts with full understanding. Epistemic Competence and Restricted Entailment address this gap. §4.1.2 revisits the claim in light of the epistemic extensions of STIT that have developed since.

Horty's framework treats actions as momentary. The branching-time semantics on which STIT is built provides rich temporal structure, moments ordered into trees, histories representing possible temporal evolutions (2001: 6–7), while the agency operator evaluates what an agent does at a single moment. There is no mechanism for representing how responsibility evolves across a temporal interval, how the window for effective action opens and closes, or how the agent should adapt their commitments as circumstances change. Horty's deontic operator ⊙ [α cstit: A] inherits this momentary character: it specifies what an agent ought to see to at a moment, evaluated against the dominance ordering of actions available at that moment (2001: 75). Temporal Bounds and

Stability address a dimension of moral responsibility that momentary evaluation does not accommodate. Conflict classification without resolution. Horty's deontic logic handles moral conflict by classification. When an agent faces competing demands, the framework classifies actions as optimal (undominated) or suboptimal (dominated) through a state-by-state comparison of their results (2001: 68–72). The principle $D^*_{\alpha\odot}$ guarantees that there can be no conflict in what a single agent ought to do (2001: 79). The guarantee is achieved at a cost. When two actions are incomparable with respect to dominance, because each leads to better results in some states, both are classified as right, and the framework falls silent on how the agent should choose between them. Horty acknowledges this: "different actions might be incomparable with respect even to weak dominance" (2001: 69). In the situations that motivate the present account, where an agent bears responsibility for competing commitments that cannot both be fully realized, this incomparability is the norm rather than the exception. *Conflict resolution* addresses the gap: where Horty's framework terminates in the classification of competing actions as jointly optimal, responsibility-centered analysis characterizes the practical reasoning that must follow, drawing on criteria internal to the framework. Causal agency without moral endorsement. The STIT operator registers causal agency — the agent's action constraining the future course of events to lie within a definite subset of possible histories (2001: 12–13) — and does not register moral endorsement. An agent who constrains histories such that A obtains sees to it that A whether or not they reflectively endorse A as a moral commitment, have deliberated about its significance, or merely produce A as a side-effect of pursuing some other goal. *Autonomous agency* addresses this dimension.

4.1.1 Action-based deontic constraints

A framing correction is needed before the comparison can be completed. It is sometimes said that deontic logic concerns states of affairs while responsibility concerns actions, and that this marks the basic difference between them. The contrast does not survive contact with recent work. Shi's (2024)

account of agentive permissions makes actions and their deontic status formal objects of the semantics. Shi's transition systems assign to each agent, at each state, a set of available actions together with a nonempty subset of permitted ones, and deontic status attaches to actions in the model rather than being derived from a preference ordering over outcomes by way of dominance reasoning, which is Horty's route (2024: 3560). Whatever separates responsibility from obligation, it is not that one concerns actions and the other does not. Shi's framework therefore supplies the ensuring/admitting distinction adopted here. An action ensures φ when φ obtains whatever the other agents do; an action admits of φ when it leaves open the possibility that φ obtains (2024: 3558). Shi makes actions and their deontic status explicit, supplies the ensuring/admitting distinction, but does not represent the agent's relationship to the deontic constraint (2024: 3562).

This is the distinction the framework adopts in §2.1 and develops throughout §3: the *control requirement* is stated with C^⇢_A, *restricted entailment* with the ensuring relation, and the two incompatibility relations ⊥^↝_A and ⊥^⇢_A separate the conflicts that trigger *conflict resolution* from the resources available for resolving them. In institutional settings, where the outcomes of an agent's actions depend on the responses of colleagues, superiors, and oversight bodies, an agent will often be unable to ensure even the approximations φ' and ψ', and what responsible action requires is the choice of an action that admits of both and maximizes the prospect of each. This distinction is indispensable in collective settings, where unilateral guarantee is often unavailable; without it, the *control requirement* would render prospective responsibility in collective settings very nearly empty. What Shi's framework does not supply is any representation of the agent's relationship to the deontic constraints it encodes. The permitted set is given in the model. Nothing in the semantics represents whether the agent endorses the constraint, understands why the action is permitted, or grasps what turns on the choice. And conflict is handled by exclusion rather than by resolution, since the inference rule she describes as conflict-preventing is valid on the grounds that the

semantics precludes the situation in which the corresponding permissions could jointly obtain (2024: 3562).

4.1.2 Epistemic extensions of STIT

Subsequent epistemic extensions of STIT incorporate knowledge, practical knowledge, and the agent's awareness of prohibited action (Broersen 2008, 2011; Broersen and Ramírez Abarca 2018; Areces et al. 2023). The epistemic gap must therefore be stated more carefully than in §4.1. What these systems represent is the agent's knowledge of what their action guarantees and, in the *mens rea* work, the agent's awareness that what they are doing is prohibited. Epistemic Competence requires more than this. Section 3.2 argued, following Fischer and Ravizza and the Rosen–FitzPatrick–Zimmerman dispute, that responsibility requires a structured relationship to moral reasons: recognition of how reasons bear on one another, understanding of why one consideration outweighs another, and the capacity and opportunity to have recognized what one failed to recognize. Knowledge of what one's action produces is a component of this and does not exhaust it. An agent may know precisely what their action guarantees while lacking any grasp of why it matters. I do not claim that this further content could not be represented in an epistemic STIT framework. The claim is narrower: the epistemic condition that responsibility imposes differs from the epistemic condition these frameworks currently encode, and the difference is the difference between knowing what one is doing and understanding what is at stake in doing it.

4.2 Extending the Guidance Control Framework

Fischer and Ravizza (1998) provide the most developed philosophical account of what moral responsibility consists in, and it is an account centered, like the present one, on responsibility rather than obligation. The relationship between the two is one of extension and specification rather than replacement. Guidance control, introduced in §3.1, consists for Fischer and Ravizza in the action's

issuing from "the agent's own, moderately reasons-responsive mechanism" (1998: 39). Two features of their account bear directly on the present framework. It is an "actual-sequence" approach: the assessment turns on how the action actually came about rather than on the agent's alternatives. And it is mechanism-based rather than agent-based: responsibility attaches to the properties of the process that actually issues in action. The present framework shares both commitments. R_A(φ) evaluates whether the agent bears responsibility for φ by reference to the conditions under which the agent acts, rather than by reference to alternative possibilities, and the four conditions function as constraints on the mechanism through which responsibility is exercised. The substantive divergence lies in what each framework does with these shared commitments. Guidance control comprises moderate reasons-responsiveness and mechanism ownership. Moderate reasons-responsiveness requires that the mechanism issuing in action be "regularly receptive to reasons, some of which are moral reasons, and at least weakly reactive to reason" (1998: 82).

The present framework distributes this insight differently. What Fischer and Ravizza call regular receptivity corresponds to Epistemic Competence, which requires that the agent possess adequate understanding of φ's consequences, sufficient capacity to reason about φ's moral implications, and awareness of how φ relates to broader moral principles. What they call weak reactivity corresponds to *autonomous agency* and *practical feasibility* jointly: the agent must maintain practical influence and the outcome must be achievable. The framework then integrates these conditions into a structure that generates determinate consequences — conjunction, practical inheritance, conflict resolution, temporal bounding — where guidance control leaves them as constraints on a mechanism-type. Fischer and Ravizza identify reasons-responsiveness and mechanism ownership as conditions for moral responsibility without specifying what follows once these conditions are met. If an agent acts from their own, moderately reasons-responsive mechanism, their framework tells us that the agent is morally responsible, but it does not tell us how

their responsibilities interact, how they extend through practical inheritance, how they behave under conflict, or how they persist across time. The structural consequences traced in §3 supply these resources.

Fischer and Ravizza's mechanism-ownership condition requires agents to regard themselves as sources of their behavior and as appropriate targets of reactive attitudes on the basis of evidence (1998: 210–213). The present framework trades some of this richness for precision: *autonomous agency's* requirements are specified clearly enough to support the structural consequences that govern responsibility's behavior, while the ownership intuition is distributed across all four conditions, which jointly ensure that responsibility is attributed only when the conditions for genuine autonomous agency obtain. Where guidance control establishes whether the agent is morally responsible, the present framework asks how that responsibility behaves once established.

4.3 Counterfactual, Seeing-to-It, and Retrospective Responsibility

A substantial formal literature addresses related forms of responsibility from causal, counterfactual, and seeing-to-it perspectives. The present account concerns an agent's standing toward outcomes before they are settled, whereas this literature variously analyzes causal contribution, counterfactual dependence, and seeing-to-it achievement within particular decision structures. Chockler and Halpern (2004), Halpern (2016), and Braham and van Hees (2012) develop causal and counterfactual approaches to responsibility concerned with an agent's contribution to an outcome. Shi and Naumov (2025) provide a particularly important recent development in the analysis of seeing-to-it responsibility in multi-step decision schemes. Their work demonstrates that distinct forms of responsibility must be distinguished according to the relations among agents, actions, and outcomes. The significance of this work for the present proposal is twofold. First, it confirms that formal responsibility is not one undifferentiated problem. Second, it demonstrates why an account of responsibility in collective and sequential settings must distinguish different

relations among agents, actions, and outcomes. The present account is adjacent to this literature but addresses a different question. Counterfactual and seeing-to-it responsibility are evaluated in relation to outcomes and decision structures in which agents can prevent, guarantee, or contribute to what occurs. R_A(φ), by contrast, represents an agent's standing toward φ before the outcome is settled. The four conditions concern the agent's standing to hold and exercise that responsibility, not the degree to which the agent ultimately caused the outcome.

The epistemic difference is also important. In causal accounts, epistemic information can enter into the assessment of blame or the interpretation of an agent's causal contribution. In the present framework, epistemic competence is constitutive of the prospective responsibility relation itself. The question concerns whether the agent stood in the relevant responsible relation to the outcome before it happened, over and above how much blame they deserved for what happened. The distinction between ensuring and admitting reinforces the boundary. Shi and Naumov's (2025) treatment of seeing-to-it responsibility in multi-step decision schemes demonstrates the importance of strategic ability and achievement in collective settings. The present paper does not reproduce their semantics. It adopts the underlying lesson that prospective responsibility in collective settings cannot simply be identified with unilateral guarantee. A complete account of institutional responsibility would need to consider both prospective commitments and retrospective responsibility for what actually occurs. The present paper confines itself to the prospective side.

4.4 Could an Obligation-Centered Framework Absorb the Four Conditions?

Each of the four conditions can in principle be represented as a restriction on when a modality applies. The question is therefore not whether an enriched deontic formalism could reproduce the extension of R_A, but whether doing so would preserve the conceptual distinction the present account is intended to identify. The answer is yes only if the enriched formalism distinguishes two questions: whether an independently specified norm binds an agent, and whether an outcome falls

within an agent's practical field of responsible engagement. These questions may have identical answers in some cases, but they are not the same question, and the difference is reflected in the conditions under which each relation obtains.

This is why the present proposal should not be understood as a claim about the expressive inadequacy of deontic logic. An enriched deontic system could encode agency, epistemic state, feasibility, and temporal position. The issue is what those additions are representing. Once those conditions determine whether the agent stands in a responsibility relation to the outcome at all, the resulting operator is tracking a relation whose extension is fixed by the agent's practical standing toward the outcome, not merely by the binding force of a norm. The formal point is therefore compatible with expressive reducibility: two relations may be represented within the same formal language without thereby being identical in conceptual role.

There is a further reason the four conditions should not be treated as an arbitrary list. They become significant through their combination. The distinctive work of the present framework appears when an agent bears several prospective responsibilities that cannot all be maximally realized. A responsible resolution must remain within the agent's influence, must be informed by adequate understanding, must be feasible, and must remain responsive to changes across the period in which the conflict develops. A framework that encodes any three of these can represent important aspects of the case, but it cannot state the complete practical problem identified here. Nothing here entails formal inexpressibility. Each condition could be represented within an enriched deontic framework; the claim is instead that an operator tracking these conditions would represent a different relation from ordinary obligation. The distinction is conceptual before it is formal. It is that an enriched operator satisfying the four conditions would represent a different normative relation from ordinary obligation, because the conditions determine when an outcome falls within an agent's

prospective responsibility rather than determining when an independently specified norm binds the agent. The point is therefore conceptual before it is formal.

5. Responsible Reasoning Under Conflict: The Intelligence Officer

The structural differences identified above would be of limited interest if they made no practical difference to moral reasoning. This section shows that they do, through a case that exposes the distinctive contribution of responsibility-centered analysis.[4] An intelligence officer discovers that a domestic surveillance program, while successfully preventing multiple terrorist attacks, systematically exceeds its legal authorization. The program processes civilian data beyond its mandate and has demonstrably saved lives through early intervention. The officer must navigate competing demands: maintaining operational security, reporting legal violations, protecting civilian lives, and upholding legal compliance.

5.1 The Obligation-Centered Impasse

The relevant question is not merely whether the officer faces conflicting obligations. It is whether an obligation-centered representation can distinguish the different practical positions available to the agent within that conflict. Consider again the four demands: operational security, disclosure, protection of civilian lives, and legal compliance. Suppose that no available action guarantees both full operational secrecy and full disclosure, but that several available actions admit of realization of both. The conflict is therefore ensuring-incompatible but admitting-compatible. This distinction matters because the officer's practical problem is not exhausted by determining which norm has

[4] This scenario draws on the NSA's PRISM surveillance program revealed in 2013, which exceeded the scope of its authorization under Section 702 of the FISA Amendments Act. See Privacy and Civil Liberties Oversight Board (2014).

lexical or ordinal priority. The agent must instead identify what can still be achieved for both commitments given the structure of the available action-set.

A ranking of obligations can determine which demand is to prevail only if the relevant normative question is which commitment has greater priority. But that is not the only question raised by the case. Two actions can be equally permissible or equally preferred under a ranking while differing significantly in how well they preserve the competing outcomes. Conversely, an action that best serves one obligation may perform substantially worse with respect to another without thereby being deontically excluded. The practical problem is therefore not simply one of ranking the commitments. It is one of identifying an action whose effects remain within the jointly achievable practical space defined by both.

5.2 Responsibility-Centered Analysis

The responsibility-centered representation makes this practical structure explicit. The officer's position is represented as $R_A(S) \wedge R_A(P) \wedge R_A(L) \wedge R_A(W)$, but these responsibilities do not individually determine a unique action. Instead, their interaction with the officer's available actions determines which forms of discharge remain possible. Possibility establishes that each responsibility must be practically achievable. Effective whistleblowing, for example, requires channels, protections, and credibility that may not be available. If the officer has no effective means of reporting the violation, the responsibility for disclosure in the present sense may fail even though an independent obligation to report could remain.

The *control requirement* directs attention to the officer's actual practical situation. The officer can maintain security, report through oversight channels, protect lives, and modify the program's scope to improve legal compliance. For each responsibility the question is whether the action set contains elements that admit of the relevant outcome. The practical incompatibility relation identifies where

genuine conflicts arise. Security and disclosure are ensuring-incompatible, $\perp^{\rightsquigarrow}_{A}(S, W)$, since no available action ensures both full operational secrecy and full public disclosure. They are not admitting-incompatible: actions remain that leave open realization of each, and the discussion of *conflict resolution* below identifies two. Protection and legal compliance conflict similarly if the program's protective efficacy depends on data processing that exceeds its mandate. These are not logical contradictions. There are possible institutional arrangements in which both outcomes obtain; the incompatibility is local to the officer's present action-set.

Conflict Resolution transforms each conflict into a meta-responsibility. The officer's responsibility is to find an action that best approximates both security and disclosure within the constraints of the four conditions. Possible responses include reporting through classified oversight channels while preserving operational security, or modifying the program's scope to restore legal compliance while retaining as much threat-detection capacity as possible.

The distinction between ensuring and admitting matters here. Reporting to an oversight body does not ensure that the program will be corrected, nor does it ensure that operational security will survive the correction. Whether either outcome eventuates depends partly on the subsequent actions of other agents and institutions. The officer's action may therefore admit both approximations without ensuring either. Prospective responsibility does not disappear merely because the officer cannot unilaterally guarantee the result.

Temporal bounds then situate these responsibilities within specific timeframes. The officer's effective window may be shaped by institutional review cycles, political dynamics, or the development of a particular threat. Stability explains why institutional inconvenience does not by itself remove the responsibility. If the officer is reassigned to a position with no access to the program and no capacity to influence it, the practical condition may fail and the responsibility may

lapse. If the officer is merely threatened with career consequences for pursuing the matter, the responsibility does not automatically disappear.

5.3 What the Framework Shows

The case reveals the difference between identifying a normative conflict and representing an agent's practical position within that conflict. An obligation-centred analysis can represent the competing demands and can, given an appropriate ordering, classify some actions as preferable or required. But the ranking by itself does not represent the distinction between actions that guarantee a commitment, actions that merely admit of its realization, and actions that preserve several commitments simultaneously at less than their maximal level. That distinction is central to the officer's practical problem.

The responsibility-centred analysis begins from the agent's standing toward each outcome and then asks how those responsibilities interact with the available action-set. Security and disclosure are ensuring-incompatible because no available action guarantees both in full. They are nevertheless admitting-compatible because some actions leave open partial realization of each. The important practical fact is therefore not that one responsibility defeats the other, but that the agent retains a non-empty space of actions through which both commitments can still be pursued. As a result, the appropriate question becomes which of those actions best preserves the competing responsibilities under the conditions governing responsible agency.

This is the sense in which practical conflict is a problem of responsible approximation rather than merely a problem of normative ranking. The framework does not show that an enriched deontic theory could not encode such a relation. Instead its claim is explanatory: once the agent's responsibility is understood as a standing toward outcomes under conditions of agency, epistemic competence, feasibility, and temporal continuity, the structure of the action-space becomes part of the object that the theory must represent. A theory that represents only which obligation has priority

omits this dimension unless it supplements the normative relation with the very practical structure that the present account makes explicit. The significance of the case therefore lies not in showing that deontic reasoning is incapable of generating a compatible verdict. It lies in showing what such a verdict must be about if it is to capture responsible practical reasoning. The relevant assessment concerns not merely which norm wins, but how an agent situated within a constrained action-space can preserve competing commitments without abandoning either where some joint realization remains practically available.

6. Conclusion

This paper has argued that prospective responsibility is a distinct normative relation from obligation because it concerns an agent's standing toward outcomes under conditions of agency, epistemic competence, feasibility, and temporal-contextual adaptability. Four conditions give the distinction its structure: *autonomous agency*, *epistemic competence, practical feasibility,* and *temporal-contextual adaptability*. These conditions yield consequences concerning control, consistency, restricted entailment, possibility, conjunction, temporal bounds, stability, and conflict resolution. The paper does not offer a complete semantics for a new logic, rather it identifies the relations and structural consequences that a more complete formalization would have to capture.

The comparison with existing work also narrows the contribution. Horty's deontic STIT already integrates obligation with agency and temporal structure. Shi's action-based framework makes permitted actions formal objects and distinguishes actions that ensure outcomes from those that merely admit of them. Epistemic STIT incorporates knowledge into agency logic. Fischer and Ravizza provide a sophisticated philosophical account of the conditions of moral responsibility. Chockler and Halpern, Braham and van Hees, and Shi and Naumov develop formal accounts of counterfactual, causal, and seeing-to-it responsibility. The present proposal does not displace these frameworks. Its contribution lies in combining a prospective outcome-directed conception of

responsibility with a four-condition account of the agent's standing toward that outcome and, in particular, in applying that structure to practical conflict.

The proposal is not a logic that supersedes deontic ethics. It is an account of what a formal theory of prospective responsibility must represent if it is to distinguish the agent's practical responsibility for outcomes from the normative status of the actions through which those outcomes are pursued. The intelligence officer case illustrates why the distinction matters: when several responsibilities are practically incompatible, responsible reasoning may consist in finding an action that preserves the competing commitments as far as circumstances allow, whatever the verdict about which demand prevails.

Two limitations remain. First, the account of "nearest jointly achievable" approximation is not yet metrically specified and therefore cannot guarantee a unique resolution. Second, a complete formal theory would require an explicit semantics connecting the responsibility operator to the action, epistemic, and temporal structures discussed here. These are substantial limitations, but they do not undermine the narrower philosophical claim. The task here has been to identify the structure that such a formalization would need to preserve. Whether that structure can be given a complete semantics, and how it should interact with formal theories of practical reasoning, remains a question for further work.

References

Areces, Carlos, Valentin Cassano, Pablo F. Castro, Raul Fervari, and Andrés R. Saravia. "A Deontic Logic of Knowingly Complying." In Proceedings of the 22nd International Conference on Autonomous Agents and MultiAgent Systems, 2023.

Belnap, Nuel, Michael Perloff, and Ming Xu. Facing the Future: Agents and Choices in Our Indeterminist World. Oxford: Oxford University Press, 2001.

Braham, Matthew, and Martin van Hees. "An Anatomy of Moral Responsibility." Mind 121, no. 483 (2012): 601–634.

Bratman, Michael E. Intention, Plans, and Practical Reason. Cambridge, MA: Harvard University Press, 1987.

Bratman, Michael E. Structures of Agency: Essays. Oxford: Oxford University Press, 2007.

Broersen, Jan. "A Complete STIT Logic for Knowledge and Action, and Some of Its Applications." In Declarative Agent Languages and Technologies VI, 47–59. Berlin: Springer, 2008.

Broersen, Jan. "Deontic Epistemic STIT Logic Distinguishing Modes of Mens Rea." Journal of Applied Logic 9, no. 2 (2011): 137–152.

Broersen, Jan, and Aldo Iván Ramírez Abarca. "Formalising Oughts and Practical Knowledge Without Resorting to Action Types." In Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems, 1877–1879, 2018.

Cane, Peter. *Responsibility in Law and Morality*. Oxford: Hart Publishing, 2002.

Chisholm, Roderick M. "Contrary-to-Duty Imperatives and Deontic Logic." Analysis 24, no. 2 (1963): 33–36.

Chockler, Hana, and Joseph Y. Halpern. "Responsibility and Blame: A Structural-Model Approach." Journal of Artificial Intelligence Research 22 (2004): 93–115.

Fischer, John Martin, and Mark Ravizza. Responsibility and Control: A Theory of Moral Responsibility. Cambridge: Cambridge University Press, 1998.

FitzPatrick, William J. "Moral Responsibility and Normative Ignorance: Answering a New Skeptical Challenge." Ethics 118, no. 4 (2008): 589–613.

Frankfurt, Harry. "Alternate Possibilities and Moral Responsibility." The Journal of Philosophy 66, no. 23 (1969): 829–839.

Goodin, Robert E. *Utilitarianism as a Public Philosophy*. Cambridge: Cambridge University Press, 1995.

Halpern, Joseph Y. Actual Causality. Cambridge, MA: MIT Press, 2016.

Hansen, Jörg. "The Paradoxes of Deontic Logic: Alive and Kicking." Theoria 67, no. 3 (2001): 214–229.

Horty, John F. Agency and Deontic Logic. Oxford: Oxford University Press, 2001.

Horty, John F., and Nuel Belnap. "The Deliberative STIT: A Study of Action, Omission, Ability, and Obligation." Journal of Philosophical Logic 24, no. 6 (1995): 583–644.

Makinson, David, and Leendert van der Torre. "Input/Output Logics." Journal of Philosophical Logic 29, no. 4 (2000): 383–408.

Prior, Arthur N. "The Paradoxes of Derived Obligation." Mind 63, no. 249 (1954): 64–65.

Privacy and Civil Liberties Oversight Board. Report on the Surveillance Program Operated Pursuant to Section 702 of the Foreign Intelligence Surveillance Act. Washington, DC: Government Printing Office, 2014.

Rosen, Gideon. "Culpability and Ignorance." Proceedings of the Aristotelian Society 103, no. 1 (2003): 61–84.

Ross, W. D. The Right and the Good. Oxford: Oxford University Press, 1930.

Scanlon, T. M. What We Owe to Each Other. Cambridge, MA: Harvard University Press, 1998.

Scanlon, T. M. Moral Dimensions: Permissibility, Meaning, Blame. Cambridge, MA: Harvard University Press, 2008.

Shi, Qi. "Agentive Permissions in Multiagent Systems." In Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence (IJCAI-24), 3558–3566, 2024.

Shi, Qi, and Pavel Naumov. "Responsibility in Multi-Step Decision Schemes." *Journal of Philosophical Logic* 54, no. 2 (2025): 493–531.

Shoemaker, David. The Architecture of Blame and Praise. Oxford: Oxford University Press, 2024.

van de Poel, Ibo. "The Relation Between Forward-Looking and Backward-Looking Responsibility." In Moral Responsibility: Beyond Free Will and Determinism, 37–52. Dordrecht: Springer, 2011.

Vincent, Nicole A. "A Structured Taxonomy of Responsibility Concepts." In *Moral Responsibility: Beyond Free Will and Determinism*, edited by Nicole A. Vincent, Ibo van de Poel, and Jeroen van den Hoven, 15–35. Dordrecht: Springer, 2011.

von Wright, Georg Henrik. "Deontic Logic." Mind 60, no. 237 (1951): 1–15.

Wallace, R. Jay. The Moral Nexus. Princeton, NJ: Princeton University Press, 2019.

Watson, Gary. "Two Faces of Responsibility." Philosophical Topics 24, no. 2 (1996): 227–248.

Watson, Gary. "Free Agency." The Journal of Philosophy 72, no. 8 (1975): 205–220.

Zimmerman, Michael J. "Moral Responsibility and Ignorance." Ethics 107, no. 3 (1997): 410–426.