

Convergent Research Map

Things You Can't Undo — The Binary Outcome Framework (BOF)

*Prepared under NMA-14 Rev.1 Governance · Dr. Olutoyese Oyelese, MD · Navigational Mind Inc.
Research conducted August 2026*

Governance Note

Pinned Correction #1 active throughout this report: BOF is anti-confabulation protection, not decision-theoretic optimisation. Do not read the arg-max in the formula as expected-utility maximisation. Every convergence below is mapped with this line held. Where a tradition resembles BOF on the surface but differs in fundamental architecture, that difference is named rather than smoothed.

Preface

Things You Can't Undo is a plain-language exposition of the Binary Outcome Framework (BOF). Its task is translational — it takes a formula that looks like physics and shows the reader that it is actually a discipline against self-deception: not a calculator for optimal choice, but a brake against confabulation. The research below maps five established traditions that independently arrive at structurally similar territory, and names precisely where BOF departs from or sharpens each one. Claims are labelled EVIDENCED (directly grounded in the document or retrieved sources) or INFERRED (warranted logical extension).

Construct 1: The Three-Box Framework — Separating Irreversible from Reversible Harm

⚠️ Pinned Correction #1 applies to this construct — see Governance Note above.

BOF'S CLAIM

Outcomes must be sorted into exactly three mutually exclusive boxes: desired outcome $P(D)$, fixable damage $P(Ur)$, and permanent damage $P(Uir)$. No blending permitted. The separation is the entire mechanism — it keeps the permanent outcome visible rather than letting narrative cancel it.

CONVERGENT WORK

Prospect Theory — Kahneman & Tversky (1979)

Loss aversion is a defining characteristic of prospect theory, whereby responses are stronger to losses than to equivalently sized gains. The minimum gain needed to balance an equal chance of losing is approximately twice the size of the loss. Prospect theory arrives at the same structural observation as BOF: gains and losses are not psychologically symmetrical and should not be placed on the same scale.

Rawls — Maximin Principle

Rawls's maximin holds that rational parties behind the veil of ignorance would choose principles that make the worst possible position as good as it can be — because the original position represents a unique, irreversible choice about the basic rules governing one's entire life. The structural move is identical: when irreversibility is at stake, protect the worst-case floor, not the best-case ceiling.

DIFFERENTIAL — WHAT BOF ADDS

BOF introduces the w weighting mechanism — a named, variable dial calibrating how much extra weight permanent harm receives. Rawls's maximin is binary; Prospect Theory uses a fixed loss-aversion coefficient. CRITICAL: Prospect Theory is a descriptive model of how minds naturally behave. BOF is a prescriptive discipline against the opposite failure mode — where narrative UNDER-weights permanent loss by blending it with upside. They converge in structure but target opposite errors.

Construct 2: The Anti-Blending Rule — 'The Upside Is Huge' Cannot Cancel a Permanent Box

BOF'S CLAIM

Blending — placing good outcomes and catastrophic outcomes on the same scale — is the mechanism by which confabulation operates. The story can make any number balance any other. Separation is the wall that keeps the story out.

CONVERGENT WORK

Taleb — Antifragility and Narrative Fallacy

Fragility implies more to lose than to gain — more downside than upside, unfavorable asymmetry. The first step toward antifragility consists in first decreasing downside, rather than increasing upside. Taleb's narrative fallacy names the same failure mode: the human tendency to create coherent stories from random events, causing us to overlook catastrophic risk. BOF's 'a story can make any number balance any other number' and Taleb's narrative fallacy are descriptions of the same move.

Kahneman — Cognitive Bias and Motivated Reasoning

Tversky and Kahneman's cognitive bias literature documents the systematic ways the mind arrives at preferred conclusions and then constructs post-hoc justifications. Confirmation bias, the planning fallacy, and overconfidence are all forms of the gap-filling machinery BOF is designed to interrupt.

DIFFERENTIAL — WHAT BOF ADDS

Taleb's solution is structural-architectural — barbell strategies, skin in the game, optionality. BOF's solution is epistemological-procedural: a discipline applied at the moment of reasoning, inside any situation, without requiring prior architectural decisions. [INFERRED: BOF is more portable and more individually applicable than Taleb's framework.]

Construct 3: $w > 1$ — Permanent Harm Always Counts Extra

BOF'S CLAIM

The weighting factor w , which must always exceed 1, ensures permanent harm is never treated as equivalent to reversible harm. As w rises, even a small probability of permanent damage can dominate the decision. The document explicitly leaves w open: it does not claim to know the number.

CONVERGENT WORK

The Precautionary Principle — Bioethics and Environmental Policy

Where there are threats of serious or irreversible damage, lack of full scientific certainty shall not be used as a reason for postponing preventive measures. Irreversibility triggers a categorical shift — uncertainty becomes a reason for more caution, not an excuse for inaction. The 1992 Rio Declaration encodes this asymmetry at the policy level.

Hippocratic Ethics — Primum Non Nocere

The precautionary principle is based on the longstanding ethical premise of medicine: it is better not to do something, or to do nothing, than to produce damage. First, do no harm. The asymmetric weighting of permanent harm — its categorical priority over potential gain — is the operative logic of Hippocratic ethics across millennia.

Stoic Premeditatio Malorum

Premeditatio malorum — the premeditation of evils — is a Stoic exercise of imagining things that could go wrong or be taken away from us. The structural purpose is identical to BOF's Box 3: force the permanent-loss scenario into visibility before the decision, so it cannot be cancelled by optimistic narrative.

DIFFERENTIAL — WHAT BOF ADDS

BOF's w is a named, variable, and explicitly open parameter. The Precautionary Principle gives a binary rule. Stoic premeditation is an attitudinal practice. BOF is the only framework that names the asymmetric weight as a dial, acknowledges it might vary by type of permanence (money vs. identity vs. relationship), and explicitly refuses to fill in a number — treating that openness as intellectual honesty rather than a gap.

Construct 4: Ethics as Architecture — 'Stay on the Board' is a Precondition, not a Moral Rule

BOF'S CLAIM

The reason to avoid permanent harm is not that it is morally wrong but that it terminates navigability. Once structure collapses, navigation stops. The framework calls this 'ethics as architecture rather than morality.' Moral rules get argued with — structural survival requirements do not.

CONVERGENT WORK

Herbert Simon — Bounded Rationality and Satisficing

Decision makers aim at satisficing — searching for alternatives good enough according to pre-established criteria. The system must continue to function — decision-making is a process that requires a decision-maker who persists. The structural priority is survival of the capacity to decide, not optimization of any single choice.

Stoic Dichotomy of Control

The Stoic avoids permanent loss not because it is morally bad, but because permanent loss removes the capacity for the virtuous action that constitutes a life well-lived. Seneca's voluntary poverty exercises are structural maintenance of the navigator's capacity to act. The goal is dynamic stability: bend, reorganize, don't fragment.

DIFFERENTIAL — WHAT BOF ADDS

BOF's 'ethics as architecture' is a genuine philosophical move distinguishable from moral philosophy AND from Simon's pragmatics. Not 'be good.' Not 'be optimal.' Not 'be satisfactory.' The document names it directly: it does not aim to maximize preferred outcomes; it structures action to reduce the likelihood of catastrophic or unacceptable outcomes, while preserving the conditions under which desired outcomes may emerge. [EVIDENCED — direct quotation from the document.]

Construct 5: Move Anyway — Caution and Courage as Simultaneous Design Features

BOF'S CLAIM

The framework requires movement without certainty. You are not permitted to wait for the fog to lift. It will not. Freezing is named as a failure state. Resolution: keep each step small enough that being wrong lands in Box 2, not Box 3. Contained. Then check, adjust, go again. This tension with the caution-about-permanence is named as deliberate design.

CONVERGENT WORK

Simon — Satisficing Requires Movement

The player, instead of seeking for a 'best' move, needs only to look for a 'good' move — and then make it. Waiting for certainty before moving is not a conservative strategy in Simon's framework; it is a failure to satisfy the decision-making requirement at all.

Keats — Negative Capability

Keats describes Negative Capability as being in uncertainties, mysteries, doubts, without any irritable reaching after fact and reason. Without it, action becomes dogmatic; with only it, action risks becoming delayed or diffused. It is a discipline before action — a way of ensuring that when we finally choose, we do so with humility, awareness, and a fuller sense of complexity. This is the exact tension BOF holds deliberately.

Taleb — Barbell Strategy

Keep most exposure safe; take risks only where downside is bounded. The barbell strategy achieves antifragility by first lowering exposure to negative tail risk, then maximizing chances of beneficial outcomes. This is the structural expression of BOF's 'small enough step that being wrong lands in Box 2.'

DIFFERENTIAL — WHAT BOF ADDS

BOF is the only framework that explicitly holds both imperatives — caution and courage — in the same architecture and names their tension as a design feature rather than a contradiction. Keats describes the attitude. Simon describes the decision procedure. Taleb describes the structural position. BOF integrates all three simultaneously: the epistemic attitude (tolerate uncertainty), the procedural constraint (Box 3 visible and weighted), and the action requirement (move anyway, step bounded). [INFERRED: this integration is BOF's most distinctive contribution in this cluster.]

Construct 6: Confabulation — The Central Problem BOF Is Solving

BOF'S CLAIM

The framework is not really about decisions. It is about the fact that we fill gaps with stories and don't notice we're doing it. 'Confabulation: making things up without knowing that you're making things up.' Everything else — the boxes, the formula, the w — is scaffolding around this one very human problem.

CONVERGENT WORK

Kahneman / Tversky — Cognitive Bias Literature

Tversky and Kahneman's cognitive bias literature documents the mechanisms of motivated reasoning — the systematic ways the mind arrives at conclusions it preferred and then constructs post-hoc justifications. Confirmation bias, the planning fallacy, and overconfidence are all confabulation in Kahneman's framework. The narrative fallacy is its most direct analog to BOF's target.

Taleb — Narrative Fallacy

Where Kahneman focuses on systematic errors the biased mind makes, Taleb focuses on the stories we retrospectively construct to make random outcomes seem inevitable. Both describe the same gap-filling machinery BOF is designed to interrupt — Kahneman from inside the cognitive process, Taleb from the pattern of outcomes.

Simon — Bounded Rationality as the Structural Cause

Simon names the structural reason confabulation happens: because we cannot hold all possible outcomes, we reduce the problem to what we can manage — and the reduction itself is where stories fill gaps that evidence cannot. BOF's three-box separation is a structural response: it holds the minimum necessary categories visible without requiring full information.

DIFFERENTIAL — WHAT BOF ADDS

What BOF adds is specific and practical: a structural mechanism, not a description of the problem. Knowing about confabulation does not prevent it — and the cognitive bias literature, though widely known, does not reliably improve decisions because awareness is insufficient. BOF provides a structural interrupt at the moment of reasoning: the three-box separation makes Box 3 resistant to narrative cancellation.

Summary: Where BOF Stands in This Terrain

BOF Construct	Strongest Convergent Tradition	What BOF Adds
Three-Box Framework	Prospect Theory (Kahneman/Tversky), Rawls (Maximin)	Structural separation as anti-narrative device; box isolation is the mechanism
Anti-Blending Rule	Taleb (narrative fallacy, antifragility)	Explicitly targets narrative as the blending agent; names the story as the failure mode
$w > 1$ Weighting	Precautionary Principle, Hippocratic ethics, Stoic premeditation	Names w as a variable dial; leaves calibration open as intellectual honesty
Ethics as Architecture	Simon (bounded rationality), Stoicism	Neither moral nor optimal — precondition. Preserves navigability rather than maximising outcomes
Move Anyway	Keats (Negative Capability), Simon, Taleb (barbell)	Holds caution AND courage in same architecture; names their tension as deliberate design
Confabulation as Central Problem	Kahneman/Tversky, Taleb, Simon	Structural interrupt rather than description — BOF blocks confabulation procedurally

Closing Note

The document's own framing of BOF as 'not a decision calculator — closer to a brake' is well-supported by the research. All six traditions surveyed describe the problem BOF is solving: the gap between what we know and what we decide. Where each tradition differs is in its proposed response.

Kahneman maps the error. Taleb restructures the exposure. Rawls builds the worst-case floor into justice theory. Simon settles for the good-enough move. Keats inhabits the uncertainty. The Precautionary Principle encodes the asymmetry into law. BOF does something none of these alone does: it provides a portable, individual-scale, procedural discipline that forces the irreversibility question into view at the moment of decision — regardless of whether you have restructured your life, changed your policies, or achieved philosophical equanimity. It is designed to work even when the story is most compelling.

That is the claim. The research confirms both its convergences and its distinctiveness.