

LEARNING THE VALUE OF VALUE LEARNING

ALEX JOHN LONDON & AYDIN MOHSENI

ABSTRACT. Standard decision frameworks addresses uncertainty about facts but assumes fixed values. We extend the Jeffrey–Bolker framework to model refinements in values and prove a value-of-information theorem for axiological refinement. In multi-agent settings, we establish that mutual refinement will characteristically transform zero-sum games into positive-sum interactions and yields Pareto-improving Nash bargains. These results show that a framework of rational choice can be extended to model value refinement and its associated benefits. By unifying epistemic and axiological refinement under a single formalism, we broaden the conceptual foundations of rational choice and illuminate the normative status of ethical deliberation.

1. INTRODUCTION

Agents in the real world, whether human or artificial intelligence, often must make decisions in the face of uncertainty. One component of such decisions is *epistemic*, involving knowledge of, or beliefs about, states of the world. Another component is *axiological*, involving judgments about the value or worth of various states of affairs and the actions that can produce them. Classic value of information theorems demonstrate the importance of epistemic refinement understood as gathering information in order to change the agent’s beliefs about states of the world. In particular, under certain conditions,¹ rational agents should prefer to acquire information about external conditions. In contrast, there is little discussion of the value of axiological refinement, understood as the process of gaining greater clarity about the nature and scope of the values that determine the relative value or worth of different states of affairs. In this paper we demonstrate the importance of axiological refinement—that pausing to refine one’s values, under quite general conditions, has strictly positive expected utility.

To demonstrate that axiological refinement cannot be collapsed into purely epistemic refinement, we provide results that go beyond individual decision-making contexts and show that there is strategic, not merely epistemic, value from consulting one’s values. In particular, we demonstrate that in two-player zero-sum games, value refinement transforms the interactions into ones that are positive-sum with moral certainty. Similarly, in Nash bargaining scenarios, value refinement tends to result in Pareto improvements for all involved parties. In these contexts, realistic agents who refine their preferences systematically outperform naïve expected utility maximizers, both individually and collectively.

Our results are important on several levels. First, the most compelling arguments for the normative recommendation to maximize expected utility rest on the presumption that the agent in question has a stable set of values that can

¹Specifically, when there is act-state independence and information is costless (Kadane et al. 2008; Skyrms 1985; Gibbard and Harper 1978).

be represented by a unique (up to positive affine transformation) utility function. Yet actual agents—humans, machines, and institutions—typically fall short of idealized expected utility maximization. Agents’ preferences or choice behavior may be intransitive, incomplete, context dependent in ways that violate independence axioms, or otherwise inconsistent. Further, agents frequently face uncertainty not just about external states of the world, but also about their own preferences and values.

This latter challenge raises a significant practical and normative question: when agents should treat their current utility function as authoritative—take direct action to maximize expected utility in light of those utilities—and when they should instead pause to reflect—to refine, revise, or better represent their values.

Second, although there is an intuitive sense in which ethical reflection is relevant to decision and game theory, this connection is often implicit. Our results make this connection explicit in a way that can be represented within the mathematical formalism of decision theory. This reveals an expanded scope for decision theory and also opens the door for more rigorous exploration of the interaction between epistemic and axiological uncertainty.

In §2, we discuss relevant previous work and situate our work in the literature. §3 introduces the Jeffrey–Bolker decision framework and details how we model value refinement. In §4, we prove a value-of-value-refinement theorem for a single agent. In §5, we show how refinement can sometimes dissolve apparent dilemmas involving incommensurable values. We extend the analysis to strategic settings in §6, demonstrating that mutual refinement turns zero-sum conflict into positive-sum cooperation, and in §7, where we derive guarantees of Pareto improvement in the context of Nash bargaining. In §8, we conclude with normative and methodological implications and outline directions for future work.

2. BACKGROUND

Good (1967) demonstrated that free information cannot decrease expected utility, establishing that rational agents should prefer more information when deciding under uncertainty. Skyrms (1990) expanded this framework, proving the expected utility of acting with additional information equals or exceeds that without it. Further work by Godfrey-Smith (1991), Fallis (2007), and Horowitz (2014) clarified boundary conditions and extended these results across decision contexts. The core insight remains: expected utility maximizers benefit from reducing uncertainty about world states before selecting actions.

In contrast to the extensive scholarship on epistemic uncertainty, axiological uncertainty—uncertainty about the nature, content, or scope of one’s own values—remains underexplored, despite its prevalence in real-world decision-making for both humans and artificial agents. The existence of value uncertainty has been recognized by a few authors. For example, Dewey and Tufts (1936, pp. 172–176) distinguish cases of weakness of will, where agents struggle to act in conformity with their judgments of right and wrong from cases where the agent is uncertain about how to reconcile their values in order to arrive at a determination about right and wrong. In this second case, they argue that what is required is not strength of will, but reflection and clarification of the values in question. On their view, such reflection is a core component of theorizing in ethics. But their work

does not provide a formal model of what such reflection should look like or how agents should approach decisions in the face of such axiological uncertainty.

Levi (1990) provides a detailed, formal model of how rational agents should approach decisions when they remain uncertain about how to reconcile all of their values. He argues that this situation can arise from two sources. The first is pragmatic—an agent may be able to reconcile their various values in principle but not know how to do so within the time required for a decision. The second is the possibility of deep incommensurability. In this case, an agent may not know how to reconcile their various values because there is no unique tradeoff function that reconciles the distinct claims or reasons for action associated with each of the values in question. To this we add a third, normative reason. Some theories in normative ethics hold that it might be possible to reconcile or trade-off particular values but that doing so is morally wrong. For example, Kantian moral theories hold that “things” can be assigned a price and so can be given a value in terms of an equivalent amount of some other thing but that rational agents have dignity and that it is a fundamental moral wrong to treat a rational agent as having a price (Bjorndahl et al. 2017).

In response to such situations, Levi argues that an agent should first step back, identify the values that are directly relevant to the choice, and then entertain all possible convex combinations of those values. Acts that are on or above the surface created by these convex combinations are admissible for choice, whereas those that fall below this surface are inadmissible. If the resulting set of admissible options contains more than one act, then second-order values can be applied to that set. If there is no conflict among the agent’s values at this second level, then they will induce an ordering of the admissible options. If there is conflict among these values as well, then the agent can iteratively apply the procedure just described. Levi refers to this ordered system of values as a value structure.

A value structure is, in a sense, a weaker notion than a utility function in that an agent who acts on the basis of a utility function can be represented as employing a value structure of a particular form—one where the value functions can be commensurated, for example by the assignment of weights to each value function, and so represented by a single all-told utility function that aggregates all values into univocal evaluations of any given state or action. However, when an agent chooses from a value structure in accordance with Levi’s principles there will be cases in which those decisions cannot be represented as choosing in accordance with any utility function.

Our results differ from the approach taken by Levi. In response to uncertainty about a set of values, Levi understands inquiry into one’s values as a process of identifying their relative weights. Until precise weights can be specified, agents entertain all convex weightings of those values. In contrast, we show that inquiry can result in an alteration to a value such that some intra-personal and some interpersonal conflicts can be resolved. As such, our results are sufficiently general that they are not limited to the stronger assumptions required for expected utility maximization and apply within Levi’s framework as well. This is demonstrated in section 5. This result is sufficient to indicate how our approach might also be relevant to theorists who attempt to accommodate incommensurability in other ways, such as parity or ambiguity.

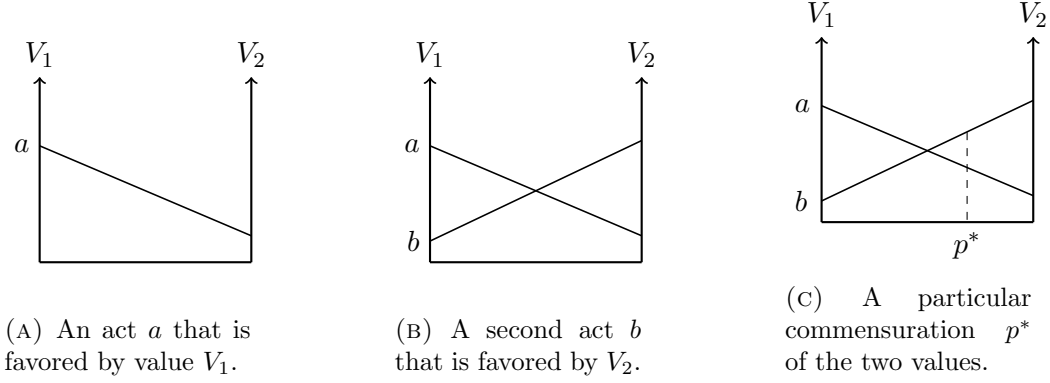


FIGURE 1. Visualization of the realization of distinct values V_1 and V_2 by two acts: a and b . The left vertical axis denotes the degree of realization of the first value V_1 . The right vertical axis denotes the degree of realization of the second value V_2 . The horizontal axis denotes the potential relative weight p of each value, and p^* denotes a particular commensuration of the two values into a single measure where V_2 composes a greater per-unit weight in the agent’s utility.

The work closest in spirit to ours is Bradley (2017) and Steele & Stefánsson (2021) on awareness growth. Bradley (2017) models awareness growth through expansion and refinement operations that partition existing propositions or introduce new sample space elements. These operations follow reverse Bayesianism, preserving the agent’s original likelihood ratios. Bradley’s refinement process parallels ours, but we extend it to decision-theoretic contexts involving utility uncertainty, not just probability uncertainty. This extension requires more general and relaxed Bayesian principles that we provide.² Steele & Stefánsson (2021) propose two principles for awareness growth: awareness reflection and preference awareness reflection. Awareness reflection applies the law of total probability to awareness changes: an agent’s current credence should equal their expected credence conditional on anticipated awareness changes. Preference awareness reflection requires that agents not expect their preference rankings to change. Our model extends awareness reflection to decision-theoretic contexts, satisfying preference awareness reflection while providing a more complete account of value learning in both individual and strategic settings.

Likewise, our results are distinct from problems of coherence across time including the way that certain “transformative” experiences can lead to various changes in agents (Ullmann-Margalit 2006; Paul 2014; Pettigrew 2015, 2019). In particular, our results apply to diachronic reasoning in both individual choice and strategic interactions and occur because commitments that an agent already endorses can be more precisely specified within particular decision contexts.

²We allow that refinement may change the probability of a proposition (though agents cannot ex ante expect change in any particular direction), shifting likelihood ratios between the refined proposition and others.

Value of information guarantees break down in game-theoretic contexts where act-state independence fails (Skyrms 1985; Gibbard and Harper 1978). When actions alter the relative probabilities of the realization states, additional information can decrease expected utility. Given this, one might conclude that value refinements offer no advantage in strategic settings. We demonstrate otherwise: learning about one’s values yields substantial benefits in game-theoretic contexts of extreme conflict.

A key question for decision theory asks when limited agents can implement expected utility maximization. Multiple arguments converge on this normative recommendation. Representation theorems (von Neumann and Morgenstern 1944; Savage 1954; Jeffrey 1965) show that if an agent’s choices satisfy fundamental axioms, her preferences can be represented by a unique utility function (up to positive affine transformation) coupled with a credence function satisfying probability axioms. Dutch book arguments (Ramsey 1931; de Finetti 1937; Kemeny 1955) prove agents violate probability axioms if and only if they accept guaranteed losses in fair bets. Accuracy arguments (Joyce 1998; Pettigrew 2016; Schoenfield 2017) establish that credences violating probability axioms are weakly dominated by conforming ones. Consequentialist arguments (Hammond 1988; McClennen 1990) demonstrate that maximizing expected utility in sequential decisions produces outcomes optimal by the agent’s standards.

When bounded agents can implement these recommendations remains unclear. Research on bounded rationality (Simon 1955; Gigerenzer and Selten 2002; Kahneman 2003) explores various departures from idealized rationality. We address a more focused question: what if an agent is uncertain about her values and can act to refine that understanding? We explore conditions where she should refine her situation understanding and values—in individual decisions, non-cooperative interactions, and Nash bargaining—before maximizing expected utility. Moreover, such refinements alter the way an agent represents the decision problem to be addressed. This connection between value specification and problem representation offers an important insight into the role that values play in problem formulation.

The Jeffrey-Bolker decision framework (Jeffrey 1965; Bolker 1966) suits this project well. Unlike Savage’s (1954) separation of states, acts, and outcomes, Jeffrey-Bolker employs propositions as the fundamental evaluation unit. Its atomless Boolean algebra structure allows preferences to be defined over propositions directly, with probability and desirability assigned to the same entities. This facilitates refinement operations where agents subdivide coarse propositions into finer ones as understanding develops. Because propositions in the algebra can be refined in well-defined ways, the framework’s flexibility makes it ideal for modeling epistemic and axiological refinements. That said, we conjecture that analogous results to those presented here can be proven in the decision frameworks of Savage, von Neumann-Morgenstern, and others.

3. SETUP & FORMAL PRELIMINARIES

3.1. A Fruitful Vignette: The Arborist and the Baker. We begin with a vignette to illustrate value refinement. Consider an Arborist A and Baker B who

jointly discover a rare orange in the wild. Both claim the fruit and find themselves in conflict over the resource.³

Initially, each sees two coarse-grained possibilities for joint action: c , to both cooperate and so divide the orange in some fashion, and $\neg c$, to both vie for all of the resource at the risk receiving none of it. Given their current understandings of those scenarios, they conceptualize the situation as apparently zero-sum, that is $EU_A(x) + EU_B(x) = 0$, and assign equal expected utility to their options:

$$EU_A(c) = EU_A(\neg c) \quad \text{and} \quad EU_B(c) = EU_B(\neg c)$$

This reflects their perception that any division merely redistributes fixed value, offering no advantage over conflict. Either action is rationalizable.

Before proceeding, each agent steps back and refines their representation of the situation. In practice, this can be through a process of reflection and introspection or conversation and third-party counsel. As a result, they discover that proposition c contains distinct possible realizations. They recognize that the arborist primarily values the fruit for the seeds because they provide a means of propagating a rare specimen, while the baker values the rind for its zest in specialty pastries. That is, c is decomposable into mutually exclusive propositions $(c \wedge f_A \wedge r_B) \vee (c \wedge f_B \wedge r_A)$ where f_i, r_i denotes apportioning the fruit or rind, respectively, to $i \in A, B$. This refinement transforms their joint action partition from $\mathbb{A}_0 = \{c, \neg c\}$ to a new, finer partition $\mathbb{A}_1 = \{c \wedge f_A \wedge r_B, c \wedge f_B \wedge r_A, \neg c\}$.

After refinement, their utilities over their actions are as follows:

$$\begin{aligned} EU_A(c \wedge f_A \wedge r_B) &> EU_A(c) = EU_A(\neg c) > EU_A(c \wedge r_A \wedge f_B), \\ EU_B(c \wedge f_A \wedge r_B) &> EU_B(c) = EU_B(\neg c) > EU_B(c \wedge r_A \wedge f_B). \end{aligned}$$

Relative to their initial valuation of the choice to cooperate and divide the resource, they discovered that there are both worse and better ways to do so. Indeed, since A and B are rational,⁴ they realize that, when they can refine their understanding of the decisions available to them, they should expect this sort of improvement to be the typical outcome! Indeed, in this case, the refinement of their shared conceptualization of the interaction transforms a zero-sum situation into a potentially positive-sum arrangement:

$$EU_A(c \wedge f_A \wedge r_B) + EU_B(c \wedge f_A \wedge r_B) > 0.$$

By consulting their values and distinguishing finer-grained realizations of their values, the parties discover a strict (ex ante) Pareto improvement—each can expect a better outcome without leaving the other expectably worse off.

This vignette illustrates one of our central contentions: when agents refine their values it can transform perceived zero-sum interactions into positive-sum arrangements and systematically improve individual and collective welfare. To make this case with greater precision, we begin with value refinement for a single agent, and progress carefully to strategic interactions.

³The vignette here is inspired by one used in Fisher, Ury, and Patton’s “Getting to Yes” (2011).

⁴In the sense that they are expected utility maximizers who also obey a reflection principle we will discuss shortly.

3.2. The Jeffrey–Bolker Decision Framework. The Jeffrey–Bolker decision framework boasts several features that make it well-suited to modeling value refinement. First, in Jeffrey–Bolker, acts, states, and consequences are all propositions in a single Boolean algebra. This allows the agent to assign probabilities to propositions about her own acts like anything else. This is essential to represent uncertainty about the structure of fine-grained actions that may result from refining her algebra. Collapsing the state/act/outcome distinction streamlines the analysis of value refinement, because every potential more detailed description of an option is already present in the algebra.

In Savage and vNM axioms, a decision problem needs to be setup so that it respects act-state independence. However, this is not required in Jeffrey–Bolker, as it is partition invariant—one can calculate expected values with respect to any partition and get the same answer. This is particularly important when one wants to simultaneously represent the prescription of competing decision theories that require different notions of act-state independence—e.g., causal, evidential, or functional. And it is important when one wishes to model decision situations where act-state independence is violated, as in cases of moral hazard or strategic interactions where agents’ actions are dependent on one another through sequential play or common knowledge.

Importantly, an agent’s representation of a decision situation is encoded by a $\langle \mathcal{A}, \mathbb{A}, U, P \rangle$ where an algebra $\mathcal{A}$ encodes the set of relevant possibilities she is entertaining, a partition $\mathbb{A} \subseteq \mathcal{A}$ encodes her action set, and U and P encodes her utilities and credences over the elements of $\mathcal{A}$. The agents can then have coarser $\mathcal{A}_0$ or finer-grained representations of the decision situation $\mathcal{A}_1$ such that $\mathcal{A}_0 \subset \mathcal{A}_1 \subset \mathcal{A}$.

Given our algebra $\mathcal{A}$ define the subset of non-empty sets as $\mathcal{A}' = \mathcal{A} \setminus \{\perp\}$, where $\perp$ is the bottom element of $\mathcal{A}$. $\mathcal{A}'$ is *atomless* if $\forall x \in \mathcal{A}', \exists y \in \mathcal{A}'$ such that $y \neq x$ and $y \subseteq x$. $\mathcal{A}$ is *complete* if every subset $X \subseteq \mathcal{A}$ has a supremum and infimum in $\mathcal{A}$.

Atomlessness guarantees our agent’s representation of the decision situation admits arbitrary precision. Every $x \in \mathcal{A}'$ can therefore be split into sharper propositions $x \wedge y$ and $x \wedge \neg y$. More generally, a proposition $x \in \mathcal{A}'$ can be decomposed into arbitrary finite partitions $x \equiv \bigvee_{i=1}^n y_i$. And the fact that actions are propositions in such an algebra guarantees that any refinement of an agent’s current action set can be the sensible objects of belief and value for an agent—that is, the domain of her credence and utility functions.

3.3. Formal Setup. Let $\mathcal{A}$ be a *complete, atomless* Boolean algebra providing an agent’s representation of the world and write $\mathcal{A}' = \mathcal{A} \setminus \{\perp\}$ for its non-empty elements. The agent’s preference relation $\succeq$ is defined on $\mathcal{A}'$ and satisfies four conditions:

Axiom 1 (Ordering). $\succeq$ is *reflexive, transitive, and complete*.

We interpret $A \succeq B$ as saying “A is at least as preferred as B,” with $\succ$ and $\sim$ defined in the standard way. Note that $\succeq$ is complete with respect to $\mathcal{A}'$, but we will later define truncated preference relations $\succeq_i$ that are defined on coarse-grained sub-algebras $\mathcal{A}_i$ of $\mathcal{A}$; in such cases, each $\succeq_i$ will constitute a complete pre-order with respect to $\mathcal{A}_i$.

Axiom 2 (Averaging). *Let $A, B \in \mathcal{A}'$ with $A \wedge B = \perp$.*

- (1) *If $A \succ B$, then $A \succ A \vee B \succ B$.*
- (2) *If $A \sim B$, then $A \sim A \vee B \sim B$.*

Averaging ensures that the value of a disjunction of two propositions lies between the two propositions. Intuitively, learning that A or B has occurred can be at most as good as the preferred proposition, and at worst as bad as the dispreferred proposition.

Axiom 3 (Impartiality). *Let $A, B \in \mathcal{A}'$ with $A \sim B$ and $A \wedge B = \perp$. If $A \vee C \sim B \vee C$ and $A \wedge C = B \wedge C = \perp$ while $A \approx C$ for some $C \in \mathcal{A}'$, then $A \vee D \sim B \vee D$ for every $D \in \mathcal{A}'$ such that $D \wedge (A \vee B) = \perp$.*

Impartiality provides the conditions for when any two disjoint propositions, A and B , are equiprobable for an agent—namely, when the agent valuing them equally implies that one would also value them equally when evaluated in disjunction with any third propositions with which they are both incompatible.

Axiom 4 (Continuity). *Let $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A}'$ satisfy $A_1 \subseteq A_2 \subseteq \dots$ and $\sup_n A_n = A^*$. If $B \succeq A^* \succeq C$, then there exists $N \in \mathbb{N}$ such that $B \succeq A_n \succeq C$ for all $n \geq N$.*

This ensures that the preference relation has no “holes” or “jumps” that would prevent continuous real-valued representation.

Bolker’s representation theorem then yields a probability measure $P : \mathcal{A} \rightarrow [0, 1]$ and a signed additive value measure $v : \mathcal{A} \rightarrow \mathbb{R}$ such that⁵

$$U(A) = \frac{v(A)}{P(A)} \quad \text{for every } A \in \mathcal{A}'.$$

For any partition $\{S_i\}$ the agent’s expected utility of A evaluated through that partition is

$$EU(A) = \sum_i U(A \wedge S_i) P(S_i | A).$$

That is, an agent whose preferences satisfy (A1-A4) can be represented as if she is maximizing expected utility in her choice of acts $A \in \mathbb{A}$ with a credence function $P : \mathcal{A} \rightarrow [0, 1]$ and utility function $U : \mathcal{A} \rightarrow \mathbb{R}$.

3.4. Modeling Value Refinement. We now extend the Jeffrey–Bolker framework by introducing a *refinement operation* on an agent’s representation of a decision problem. We begin with an idealized benchmark, then systematically develop the machinery for bounded agents who can refine the understanding and valuations of the actions available to them.

3.4.1. The Unbounded Benchmark. Consider first an unbounded agent who has evaluated all possible refinements of her decision problem. Her complete representation is $D = \langle \mathcal{A}, \mathbb{A}, P, U \rangle$ where:

- $\mathcal{A}$ is a complete, atomless Boolean algebra containing all conceivable propositions;

⁵The functions are unique when one further assumes unbounded utilities (Jeffrey 1965) or comparative probabilities (Joyce 1998) or equiconfidence judgments (Ahmed 2021).

- $\mathbb{A}$ is an act partition including all maximally fine-grained actions;⁶
- and $P : \mathcal{A} \rightarrow [0, 1]$ and $U : \mathcal{A}' \rightarrow \mathbb{R}$ are fully specified probability and utility functions.

This agent has already considered every relevant factor to arbitrary precision and settled her valuation of every possible act.⁷

3.4.2. The Bounded Agent. In contrast, the agents we study have not undergone arbitrary refinement. Their initial representation is $D_0 = \langle \mathcal{A}_0, \mathbb{A}_0, P_0, U_0 \rangle$ where:

- $\mathcal{A}_0 \subset \mathcal{A}$ is a coarse-grained subalgebra;
- $\mathbb{A}_0$ is the maximally fine-grained coarsening of $\mathbb{A}$ that is a subset of $\mathcal{A}_0$;⁸
- and P_0 and U_0 are restricted to $\mathcal{A}_0$ and $\mathcal{A}'_0$, respectively.⁹

This agent has not considered every factor, and their actions are restricted to those options they have considered. A key question is: How does such an agent refine their understanding of their actions $\mathbb{A}_0 \subset \mathbb{A}$ and relate to the value of refinements of their actions space that they has not yet realized—that is, how do they reason about propositions in $\mathcal{A} \setminus \mathcal{A}_0$?

3.4.3. Refinement Operations. We model bounded agents as exploring their decision problem by taking available acts and considering them in greater detail. That is, an agent takes a given action available to her A , and considers distinct ways it might be realized with respect to previously unconsidered propositions B_1 and B_2 . For example, our arborist might refine the act ‘split the resource’ into the more fine-grained partition consisting of ‘split the resource and give the rind to the baker’ and ‘split the resource and give the fruit to the baker’ This is modeled as follows.

Definition 5 (Binary Refinement). *Given an act partition $\mathbb{A}_0$ and a proposition $A \in \mathbb{A}_0$, a binary refinement of A produces the partition*

$$\mathcal{R}_A = \{A \wedge B_1, A \wedge B_2\}$$

where $B_1, B_2 \in \mathcal{A} \setminus \mathcal{A}_0$ are previously unconsidered propositions that partition A .¹⁰ This yields the refined algebra

$$\mathcal{A}_1 = \sigma(\mathcal{A}_0, \mathcal{R}_A),$$

and the refined act partition

$$\mathbb{A}_1 = (\mathbb{A}_0 \setminus \{A\}) \cup \mathcal{R}_A.$$

More generally, we can define a finitary refinement as follows.

⁶We will consider the finite case. That said, one could consider the infinite case, e.g., the set of all proper filters, or ultrafilters that are subsets of $\mathcal{A}$.

⁷Note that while an agent’s algebra may be atomless, their act partition may well be finite. This is the case we consider here. That is, one can imagine that for an agent there is a maximum level of precision in the control the can exercise in the realization of their acts, even if they can entertain even more fine-grained distinctions with respect to the state of the world.

⁸This will tend to be unique.

⁹And hence to $\mathbb{A}_0$ as $\mathbb{A}_0 \in \mathcal{A}$.

¹⁰That is, where $B_1 \vee B_2 = A$, $B_1 \wedge B_2 = \perp$, and neither B_1 nor B_2 is the bottom element $\perp$. This ensures refinements are non-trivial.

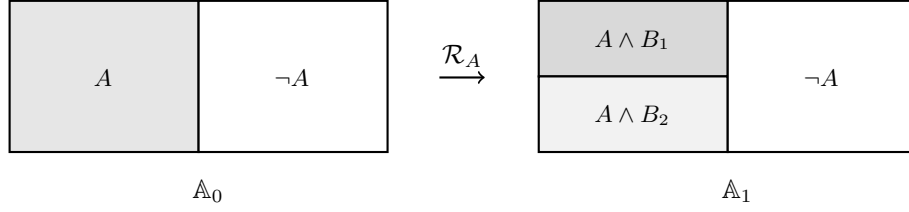


FIGURE 2. A binary refinement of act $A \in \mathbb{A}_0$. The initial act partition $\mathbb{A}_0$ consists of acts A and $\neg A$. Refinement produces $\mathbb{A}_1$ by splitting A into the more fine-grained acts $A \wedge B_1$ and $A \wedge B_2$.

Definition 6 (*k*-Ary Refinement). *A k -ary refinement of $A \in \mathbb{A}_0$ produces the partition*

$$\mathcal{R}_A = \{A \wedge B_1, \dots, A \wedge B_k\}$$

where $\{B_i\}_{i=1}^k \in \mathcal{A} \setminus \mathbb{A}_0$ are previously unconsidered propositions that partition $A \in \mathbb{A}_0$. This yields the refined algebra

$$\mathcal{A}_1 = \sigma(\mathcal{A}_0, \mathcal{R}_A),$$

and the refined act partition

$$\mathbb{A}_1 = (\mathbb{A}_0 \setminus \{A\}) \cup \mathcal{R}_A.$$

Since any k -ary refinement can be straightforwardly achieved through sequential binary refinements, we focus on the binary case throughout, understanding that our results hold for finite refinements in general.

Further, one can model unconsidered actions by introducing a catch-all proposition for actions not yet contemplated. This approach, developed for non-parametric Bayesian inference to model new hypotheses (Walker 2013; Balocchi et al. 2025), allows agents to discover new actions by refining the catch-all proposition.

Definition 7 (Expansion via Catch-All). *When an agent's act partition $\mathbb{A}_0$ contains a catch-all act $A_{CA} \in \mathbb{A}_0$, expansion of her act set is represented by the binary refinement $\mathcal{R}_{A_{CA}}$ of the catch-all act.*

When an agent employs a catch-all representing heretofore unconsidered acts, expansion of an agent's act partition¹¹ becomes a special case of refinement—refinement of the catch-all act proposition. This extends the scope of application of refinement to subsume the expansion of the agent's algebra.

3.4.4. Modeling Agent Attitudes Regarding Refinement. An agent contemplating refinement faces a decision under uncertainty—she must estimate refinement's value without knowing what distinctions she'll discover. We model this meta-uncertainty through a probability distribution over possible refinement outcomes.

When facing novel, uncertain problems, we expect greater returns from reflection than with familiar ones where prior consideration has identified optimal strategies. Such differential expectations reveal that we operate with tacit models of refinement's expected value across problem types. To capture these attitudes formally, we equip our agent with a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ where each $\omega \in \Omega$

¹¹This closely resembles Joyce (2017), where hypothesis sets grow, but differs in that the sample space remains fixed (new hypotheses already exist within the catch-all) and we consider axiological, not just epistemic, learning.

represents one possible outcome of the cognitive refinement process for a given act $A \in \mathbb{A}$ —one way the agent’s credences and utilities might settle after careful consideration.

We define the refinement outcome mapping $\xi^A : \Omega \rightarrow \mathbb{R}^2 \times \Delta^2$ by $\omega \mapsto (u_1(\omega), u_2(\omega), p_1(\omega), p_2(\omega))$, where $\Delta^2 = \{(p_1, p_2) \in [0, 1]^2 : p_1 + p_2 \in (0, 1)\}$.¹² Each component represents a post-deliberation assessment: $u_1(\omega) = U_1^\omega(A \wedge B_1)$ is the utility the agent will assign to the first refined act after deliberation in state ω , $u_2(\omega) = U_1^\omega(A \wedge B_2)$ is the utility she will assign to the second refined act, $p_1(\omega) = \bar{P}_1^\omega(A \wedge B_1)$ is her post-deliberation credence in the first refined proposition before rescaling, and $p_2(\omega) = \bar{P}_1^\omega(A \wedge B_2)$ is her credence in the second refined proposition before rescaling.

When contemplating refinement of act $A \in \mathbb{A}_0$, the agent maintains beliefs about what credences and utilities she’ll assign to the refined propositions. We capture this through a joint distribution μ_A over refinement outcomes defined as

$$\mu_A = \mathbb{P} \circ (\xi^A)^{-1}.$$

This is a probability measure on $\mathbb{R}^2 \times \Delta^2$ induced by $\mathbb{P}$ through the mapping ξ^A . For any measurable set $E \subseteq \mathbb{R}^2 \times \Delta^2$, we have $\mu_A(E) = \mathbb{P}(\{\omega \in \Omega : \xi^A(\omega) \in E\})$.

Since refinement may lead the agent to revise her assessment of $P_1(A)$ from $P_0(A)$ to $p_1(\omega) + p_2(\omega)$, maintaining probabilistic coherence requires rescaling. The post-refinement probability and utility functions become:

$$\begin{aligned} P_1^\omega(A \wedge B_1) &= \frac{p_1(\omega)}{Z(\omega)}, & P_1^\omega(A \wedge B_2) &= \frac{p_2(\omega)}{Z(\omega)}; \\ U_1^\omega(A \wedge B_1) &= u_1(\omega), & U_1^\omega(A \wedge B_2) &= u_2(\omega); \end{aligned}$$

where for old propositions $X \in \mathcal{A}_0$,

$$P_1^\omega(X) = P_0(X|\neg A)P_1(\neg A) + P_0(X|A)P_1(A)$$

and for new propositions $X \in \mathcal{A}_1$,

$$P_1^\omega(X) = P_1(X|\neg A)P_1(\neg A) + P_1(X|A)P_1(A);$$

where the normalization factor is

$$Z(\omega) = 1 - P_0(A) + p_1(\omega) + p_2(\omega)$$

ensuring that $\sum_{Y \in \mathbb{A}_1} P_1^\omega(Y) = 1$.

We assume that $\mu_A(\{p_1 + p_2 \in (0, 1)\}) = 1$, reflecting that the agent doesn’t expect refinement to reveal that A is either impossible or certain, ensuring that $Z(\omega)$ is non-zero (since $1 - P_0(A)$ is at least zero). We do not require that the probability of A after refinement $P_1(A)$ be same as its probability before refinement $P_0(A)$.¹³ This is because we want to allow that, through refinement, the agent may realize the proposition corresponding to the original act is actually more or less probable than initially believed. When this occurs, maintaining probabilistic coherence requires rescaling all credences using a normalizing factor $Z(\omega) \in \mathbb{R}^+$.

This framework enables us to model the agent’s *ex ante* attitudes regarding refinement’s potential value. For example, define the expected utility of act X

¹²Assume ξ^A is measurable. In general, unless otherwise stated, all functions are measurable.

¹³Though, as we will see, we do require that a rational agent’s credences prior to refinement satisfy that $\mathbb{E}[p_1 + p_2] = P_0(A)$ —the probability of an act A is preserved *in expectation*.

under representation D_i as $V_i(X) = P_i(X) \cdot U_i(X)$. The agent can now formulate and evaluate key propositions.

The probability of discovering a superior act through refining A is computed as follows. Let $A_0^* = \arg \max_{X \in \mathbb{A}_0} V_0(X)$ be the current optimal act. The probability that refining A yields something better is $\mathbb{P}[\max\{V_1(A \wedge B_1), V_1(A \wedge B_2)\} > V_0(A_0^*)]$. The expected value of refinement, denoted $\text{EVR}(A)$, equals

$$\mathbb{E}_{\mathbb{P}}[\max_{X \in \mathbb{A}_1(\omega)} V_1(X)] - V_0(A_0^*),$$

where $\mathbb{A}_1(\omega) = (\mathbb{A}_0 \setminus \{A\}) \cup \{A \wedge B_1, A \wedge B_2\}$ is the refined act partition.

The agent can also assess how likely refinement is to substantially change her beliefs through the probability $\mathbb{P}[|p_1(\omega) + p_2(\omega) - P_0(A)| > \epsilon]$ for some threshold $\epsilon > 0$. Despite expecting $\mathbb{E}[U_1(A)] = U_0(A)$ by the Refinement Reflection Principle, she can compute the risk of value deterioration as $\mathbb{P}[\max_{X \in \mathbb{A}_1(\omega)} V_1(X) < V_0(A_0^*)]$, which captures scenarios where refinement reveals the current best option is actually comprised of inferior sub-acts, and rescaling reduces the value of other options.

Using relative entropy, the agent can quantify expected epistemic improvement through $\mathbb{E}_{\mathbb{P}}[D_{KL}(P_1(\cdot) \| P_0(\cdot))]$, where the KL divergence is computed over the common support after extending P_0 to $\mathcal{A}_1$. If the agent has beliefs about which distinctions matter, she can condition on discovering particular types of refinements, computing $\mathbb{E}_{\mathbb{P}}[\max_{X \in \mathbb{A}_1(\omega)} V_1(X) \mid u_1(\omega) - u_2(\omega) > \delta]$ for some utility difference threshold δ .

3.4.5. The Refinement Reflection Principle. Importantly, an agent must decide *whether* to refine their understanding of an act in the first place. To do so, we require some structure to the agent's expectations regarding the results of reflection. To this end, we proffer two assumptions: a reflection-type principle relating a rational agent's current valuations of each act to its expected decomposition when refined; and a non-triviality assumption capturing when an agent is genuinely uncertain regarding the outcome of a possible refinement.

First, we establish the general principle for arbitrary finite refinements:

Definition 8 (Refinement Reflection Principle - General Form). *For any proposition $A \in \mathbb{A}_0$ and a k -ary refinement $\mathcal{R}_A = \{A \wedge B_1, \dots, A \wedge B_k\}$, a rational agent's credences obey:*

$$U_0(A) = \mathbb{E}_{\mu_A}[U_1(A)] = \mathbb{E}_{\mu_A} \left[\frac{\sum_{i=1}^k p_i \cdot u_i}{\sum_{i=1}^k p_i} \mid \sum_{i=1}^k p_i > 0 \right]$$

where $(u_1, \dots, u_k, p_1, \dots, p_k) \sim \mu_A$ with $u_i = U_1(A \wedge B_i)$ representing the utility of the i -th refined act, p_i denoting the raw (unnormalized) probability assigned to $A \wedge B_i$ after refinement.¹⁴

This principle states that the agent's current valuation of A equals what she expects her valuation will be after more careful consideration, accounting for the possibility that refinement may lead her to revise $P(A)$ itself from $P_0(A)$ to $\sum_{i=1}^k p_i$. The general RRP can be equivalently expressed using conditional probabilities.

¹⁴Note that we require $\sum_i p_i > 0$ almost surely under μ_A .

If we let $q_i = p_i / (\sum_j p_j)$ denote the conditional probability $P_1(B_i|A)$ after refinement, then $U_1(A) = \sum_{i=1}^k q_i \cdot u_i$, and the RRP states that $U_0(A) = \mathbb{E}_{\mu_A}[U_1(A)]$, where the expectation accounts for both the uncertainty in the conditional probabilities and the utilities.

Since any k -ary refinement can be achieved through sequential binary refinements (as noted in Section 3.3.3), we focus on the binary case throughout our analysis:

Definition 9 (Refinement Reflection Principle - Binary Case). *For any proposition $A \in \mathbb{A}_0$ and a binary refinement $\mathcal{R}_A = \{A \wedge B, A \wedge B_2\}$, a rational agent's credences obey:*

$$U_0(A) = \mathbb{E}_{\mu_A}[U_1(A)] = \mathbb{E}_{\mu_A} \left[\frac{p_1 \cdot u_1 + p_2 \cdot u_2}{p_1 + p_2} \middle| p_1 + p_2 > 0 \right]$$

where $(u_1, u_2, p_1, p_2) \sim \mu_A$ with $u_1 = U_1(A \wedge B)$ and $u_2 = U_1(A \wedge B_2)$ representing the utilities of the refined acts, p_1 and p_2 denoting the raw (unnormalized) probabilities assigned to $A \wedge B$ and $A \wedge B_2$ after refinement.¹⁵

This formulation parallels classical reflection principles where an agent's current credence equals her expected future credence, yet her actual future credence typically differs from the current one after consideration.¹⁶ Just as an agent's current credence should equal her expected future credence after conditioning, her current value assessment of a proposition in her algebra should equal her expected future assessment after refinement.¹⁷

Our second assumption ensures agents have genuine uncertainty about refinement outcomes.¹⁸

Definition 10 (Refinement Uncertainty). *An agent has refinement uncertainty if her distribution μ_A assigns positive probability to any open neighborhood in the outcome space. For the binary case, this means μ_A has continuous marginals on $\{(u_1, u_2, p_1, p_2) \in \mathbb{R}^2 \times [0, 1]^2 : p_1 + p_2 \in (0, 1)\}$.*

This ensures the agent has genuine uncertainty about how refinement will shape her credences and utilities by ruling out point-mass beliefs about refinement outcomes. The continuity assumption is crucial: it ensures that exact equalities like $u_1 = u_2$ occur with probability zero, enabling our strict inequality results. An agent satisfying refinement uncertainty believes that it is possible for refinement to reveal value.

Together with the RRP, these conditions ensure positive probability of discovering strict improvements through refinement. The key intuition for upcoming results is that if an agent has genuine uncertainty about refinement outcomes (refinement uncertainty), and their current valuation reflects the expected outcome

¹⁵Note that we require $p_1 + p_2 > 0$ almost surely under μ_A .

¹⁶Cf. van Fraassen (1984), Greaves and Wallace (2006), and Huttegger (2013).

¹⁷See (Dorst 2024) for a recent discussion of reflection principles and challenges to their normative status as standards for rationality.

¹⁸For now, we will also assume *refinement-state independence*: the decision to refine, by itself, carries no inherent information regarding the probabilities of existing propositions. That is, the act of deciding to think more carefully doesn't itself provide evidence, though the outcome of that consideration may lead to credence revision through rescaling. That said, one can imagine conditions under which this assumption will be relaxed.

of such consideration (RRP), then with positive probability they will identify superior options, and being rational, the agent selects these promising prospects where possible, thereby creating value through refinement of their act partition.

4. VALUE REFINEMENT FOR ONE AGENT

We can now apply our formal framework to analyze value refinement for a single agent. To summarize our setup: An agent with decision problem $D_0 = \langle \mathcal{A}_0, A_0, P_0, U_0 \rangle$ faces a choice among acts in the finite partition $\mathbb{A}_0$. While her current representation assigns determinate values $U_0(A)$ to each act $A \in \mathbb{A}_0$, she maintains uncertainty about how these values decompose when acts are refined. This uncertainty is captured by distributions $\{\mu_A\}_{A \in \mathbb{A}_0}$, where each μ_A encodes her beliefs about the utilities and probabilities of refined propositions, constrained by the Refinement Reflection Principle: $U_0(A) = \mathbb{E}_{\mu_A}[p_1 \cdot u_1 + p_2 \cdot u_2]$.

Before presenting the formal result, consider why value refinement generates expected value. When an agent chooses a coarse-grained act A , she commits to accepting whatever mixture of outcomes that act represents. This is like buying a bundle without knowing exactly what's inside—she gets the weighted average of the bundle's contents. Refinement unbundles the act, allowing her to see the components and choose among them. Since she can now select the best component rather than accepting the average, her expected utility improves.¹⁹

Consider an agent with initial action partition $A_0 = \{A_1, A_2, \dots, A_n\}$ and expected utilities $U_0(A_i)$ for each $A_i \in A_0$. An agent might consider refining one of her current best options, that is, a rationalizable act $A_j^\epsilon \in \arg \max_{A \in \mathbb{A}_0} U_0(A)$. Consider the refinement such an act $A_j \in A_0$ with respect to a previously unconsidered propositions $B_1, B_2 \in \mathcal{A} \setminus \mathcal{A}_0$.

The agent's uncertainty about the refined propositions is captured by the joint distribution μ_{A_j} over the outcomes of refinement $(u_1, u_2, p_1, p_2) \in \mathbb{R}^2 \times [0, 1]^2$.

By the Refinement Reflection Principle:

$$U_0(A_j) = \int_{\mathbb{R}^2 \times [0, 1]^2} \frac{p_1 \cdot u_1 + p_2 \cdot u_2}{p_1 + p_2} d\mu_{A_j}(u_1, u_2, p_1, p_2)$$

When the agent consults her values, she observes a specific realization $(u_1, u_2, p_1, p_2) \sim \mu_{A_j}$. Her post-refinement algebra becomes

$$\mathcal{A}_1 = \sigma(\mathcal{A}_0, \{B_1, B_2\})$$

and her post-refinement act partition becomes

$$\mathbb{A}_1 = (\mathbb{A}_0 \setminus \{A_j\}) \cup \{A_j \wedge B_1, A_j \wedge B_2\}.$$

Define $V_0 \equiv \max_{A_i \in \mathbb{A}_0} U_0(A_i)$ as the expected value of one of the agent rationalizable acts before refinement;²⁰ define $V'_0 \equiv \max_{A_i \in (\mathbb{A}_0 \setminus \{A_j\})} U_0(A_i)$ as the value of the agent's best act excluding the act she intends to refine from the set of options; and $V_1 \equiv \max\{V'_0, u_1, u_2\}$ as the value of the agent's best act after refinement.

¹⁹This relates to concept of option value. Just as a financial option has value because it provides the right but not the obligation to take an action, value refinement creates worth by transforming a commitment to an average into a choice among components.

²⁰A *rationalizable act* A^* in this context will be a member of the agent's best response set $A^* \in \arg \max_{A \in \mathbb{A}_0} U_0(A)$.

$$u(A) = \mathbb{E}_{\mu_A} [pu_1 + (1-p)u_2] < \mathbb{E}_{\mu_A} [\max\{u_1, u_2\}]$$



FIGURE 3. Refinement transforms commitment to an average of a coarse-grained bundle into the ability to select the best component among its fine-grained elements. Under RRP and refinement uncertainty, $\mathbb{E}_{\mu_A}[\max\{u_1, u_2\}] > \mathbb{E}_{\mu_A}[pu_1 + (1-p)u_2]$. The expected maximum of a non-uniform bundle exceeds its expected mean.

With preliminaries in place, we can state our main result.

Theorem 11 (Value of Value Refinement). *Consider an agent with decision problem $D_0 = \langle \mathcal{A}_0, \mathbb{A}_0, P_0, U_0 \rangle$ satisfying RRP. Let $A^* \in \arg \max_{A \in \mathbb{A}_0} U_0(A)$ be a rationalizable act, and suppose μ_{A^*} satisfies refinement uncertainty. Then the expected value of refinement is strictly positive:*

$$\mathbb{E}_{\mu_{A^*}}[V_1] > V_0,$$

where $V_0 = \max_{A \in \mathbb{A}_0} U_0(A)$ and $V_1 = \max\{U_1(A) : A \in \mathbb{A}_1\}$ denote, respectively, the agent's optimal expected utilities before and after refinement.²¹

The theorem establishes that value refinement creates strictly positive value in expectation whenever there is genuine uncertainty about how an act decomposes.

Several aspects merit emphasis. Refinement uncertainty ensures that $\mathbb{P}(u_1 = u_2) = 0$. This is crucial because if $u_1 = u_2$, then refinement would be pointless—the agent would be choosing between identical options. This guarantees that with probability one, the refined options differ, creating the possibility of selecting the better one.

The RRP ensures that the agent expects refinement to constitute a mean-preserving spread of utility. The agent's current valuation $U_0(A_j)$ already reflects her expectation about refined utilities. Together, these make it so the agent assigns positive probability to a refined act being an improvement on the original act. What refinement adds is not a predictable change in average value, but rather finer-grained choices and the consequent potential to avoid below-average realizations while realizing above-average ones.

The framework naturally extends to sequential or simultaneous refinements. If an agent can refine multiple acts or perform iterative refinements, each refinement operation strictly increases expected utility:

Corollary 12 (Monotonicity of Refinement Value). *Consider successive refinements $\mathcal{A}_0 \subset \mathcal{A}_1 \subset \dots \subset \mathcal{A}_k$ with corresponding optimal values $V_0, V_1, \dots, V_k$. If each refinement satisfies Theorem 11, then*

$$V_0 < V_1 < \dots < V_k.$$

²¹All proofs are provided in the appendix.

An agent need not fully resolve all uncertainty. Even partial refinement—distinguishing some but not all relevant propositions—creates value. Expected value increases whenever the refinement introduces meaningful choice.

The theorem and its corollary assume refinement incurs no cost. In practice, refinement requires time, cognitive effort, and may delay decision-making. We can therefore extend the result to incorporate such costs and characterize the resulting optimal refinement policies.

Suppose that each refinement incurs a fixed cost $c \in \mathbb{R}_+$. Let Δ_{R_i} denote the expected gain in value from the i -th refinement,

$$\Delta_{R_i} = \mathbb{E}[V_{i+1} - V_i \mid \mathcal{A}_i],$$

where V_i is the optimal value achievable with algebra $\mathcal{A}_i$.

Empirically, one might expect successive refinements to exhibit diminishing marginal returns: early distinctions address the largest uncertainties, and subsequent refinements reveal progressively smaller gains. This motivates the following assumption.

Definition 13 (Vanishing Returns). *A sequence $\{\Delta_{R_i}\}_{i \in \mathbb{N}}$ exhibits vanishing returns if $\Delta_{R_i} > \Delta_{R_{i+1}} \geq 0$ for all i , and $\lim_{i \rightarrow \infty} \Delta_{R_i} = 0$.*

This assumption captures the idea that the marginal benefits of refinement decrease as the agent exhausts the most valuable distinctions, and that the value of subsequent refinements tends toward zero.

Theorem 14 (Optimal Refinement with Fixed Costs). *Consider sequential refinements each costing $c > 0$. If $\{\Delta_{R_i}\}$ satisfies Definition 13, then:*

- (1) *If $c > \Delta_{R_0}$, the optimal policy is never to refine.*
- (2) *If $c \leq \Delta_{R_0}$, the optimal stopping time is $t^* = \max\{t \in \mathbb{N} : \Delta_{R_t} \geq c\}$, yielding net gain*

$$V_{t^*} - t^*c - V_0 = \sum_{i=0}^{t^*-1} (\Delta_{R_i} - c) > 0.$$

This theorem characterizes a simple and intuitive policy: continue refining as long as the expected marginal gain exceeds the cost, and stop when it no longer does. If the cost of the first refinement already outweighs its expected benefit, refinement is irrational. Otherwise, a unique stopping point t^* exists where marginal benefit falls below cost.

The structure mirrors classic results in optimal search theory and the economics of information acquisition (Stigler 1961; McCall 1970; Lippman and McCall 1976; Mortensen 1986). We can expect that agents will refine more extensively when decision stakes are high, costs are low, or initial uncertainty is large—precisely the conditions under which the value of distinguishing finer propositions is greatest.

Having established that refinement can increase an individual's expected utility, even when costly, we now turn to cases where the agent's values themselves conflict. Can refinement dissolve dilemmas between competing ends when no common metric exists?

5. RESOLVING DILEMMAS WITHOUT COMMENSURATION

Our model of value refinement is also relevant to theories of choice that deviate from expected utility maximization. Recall Levi's (1990) theory discussed earlier,

in which agents may be uncertain about how to commensurate or reconcile the various values to which they are committed. An agent with distinct values such as compassion and honesty might confront actions that realize those values to different degrees, requiring judgments regarding commensuration. For instance, how great a gain in harm aversion justifies how great a loss in fairness? While the agent may easily rank actions across each dimension of value, she may struggle to determine how to commensurate between the dimensions themselves.

Standard expected utility theory resolves such cases by commensuration: choose weights for each value and maximize the aggregate. Yet often the correct weights are unknown or normatively contested. We ask whether *value refinement* can sometimes eliminate the need for commensuration by revealing actions that are jointly better across multiple dimensions.

We extend our framework to accommodate agents with multiple distinct values. Let $V_1, \dots, V_k : \mathcal{A}' \rightarrow \mathbb{R}$ represent distinct value dimensions. Under standard aggregation, a weight vector $\mathbf{w} = (w_1, \dots, w_k)$ with $w_i \geq 0$ and $\sum_i w_i = 1$ yields

$$U(A) = \sum_{i=1}^k w_i V_i(A) \quad \text{for } A \in \mathcal{A}'.$$

Uncertainty about $\mathbf{w}$ or the normative inadmissibility of commensuration complicate choice when actions trade off values.

Consider the simplest case with two values, V_1 and V_2 , with initial partition $\mathbb{A}_0 = \{A, \neg A\}$ with opposing realizations:

$$V_1(A) > V_1(\neg A) \quad \text{and} \quad V_2(A) < V_2(\neg A).$$

The agent faces a dilemma: A is favored by V_1 , while $\neg A$ is favored by V_2 . Within expected utility theory, resolution requires determining the relative weight $w \in [0, 1]$ such that her utility becomes $U(A) = (1 - w)V_1(A) + wV_2(A)$. When $w = 0$, only V_1 matters; when $w = 1$, only V_2 matters; and when $0 < w < 1$, value V_1 matters $\frac{1-w}{w}$ times as much as V_2 . This confronts all the fundamental challenges of commensuration between potentially incommensurable values.

The possibility of value refinement opens a third path. Rather than resolving her uncertainty about the correct weight w , the agent can refine her understanding of the available actions themselves. As established in our earlier results, she can often expect to improve her expected utility through such refinement. However, in certain fortunate instances, refinement may eliminate the dilemma entirely by revealing an action that excels along both value dimensions, obviating the need for commensuration.

Following our framework, the agent fine-grains A by considering some previously unconsidered proposition $B \in \mathcal{A} \setminus \mathbb{A}_0$, producing the refined act partition

$$\mathbb{A}_1 = (\mathbb{A}_0 \setminus \{A\}) \cup \{A \wedge B, A \wedge \neg B\}.$$

Crucially, this refinement affects both value dimensions simultaneously. For each value function V_i , the agent maintains uncertainty about the refined values, captured by distributions $\mu_A^{(i)}$ over the triple $(v_{i,1}, v_{i,2}, q^{(i)})$ where $v_{i,1} = V_i(A \wedge B_1)$, $v_{i,2} = V_i(A \wedge B_2)$, and $q^{(i)} = P(B_1|A)$ represents the conditional probability.

The Refinement Reflection Principle extends naturally to each value dimension. For each $i \in \{1, 2\}$, we require $V_i(A) = \mathbb{E}_{\mu_A^{(i)}}[q^{(i)} \cdot v_{i,1} + (1 - q^{(i)}) \cdot v_{i,2}]$. This ensures

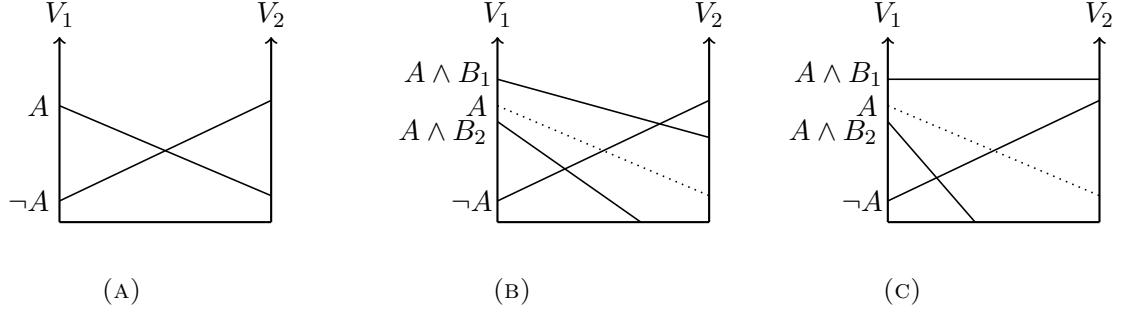


FIGURE 4. Visualization of how refinement can dissolve a value conflict with two value dimensions, V_1 , V_2 , and two acts, A , and $\neg A$. The left vertical axis denotes the degree of realization of the first value V_1 ; the right vertical axis denotes the degree of realization of the second V_2 . Figure 4a shows the dilemmas: A is favored by V_1 and $\neg A$ is favored by V_2 . Figures 4b and 4c show two possible results of refining act A : In 4b the dilemma remains, while in 4c a dominating action $A \wedge B$ is revealed making commensuration unnecessary.

that the agent's current valuation along each dimension equals her expected future valuation after refinement, maintaining coherence across the multiple values.

The key insight involves the concept of *multi-value dominance*.

Definition 15. An action A^* multi-value dominates a set $\mathcal{X}$ if

$$V_i(A^*) \geq \max_{A \in \mathcal{X}} V_i(A) \text{ for all } i,$$

with strict inequality for at least one i .

When such an action exists, the agent can choose it without determining the weight vector $\mathbf{w}$, since it remains optimal for any admissible weighting of the values. This represents a particularly elegant resolution to the commensuration problem—not by solving it, but by discovering it need not be solved.

Theorem 16 (Dilemma Resolution through Value Refinement). *Consider an agent with value functions $V_1, V_2 : \mathcal{A}' \rightarrow \mathbb{R}$ facing conflict where $V_1(A) > V_1(\neg A)$ and $V_2(A) < V_2(\neg A)$. Let the agent have independent refinement uncertainties $\mu_A^{(1)}, \mu_A^{(2)}$ over A . If the joint support includes realizations where one refined action multi-value dominates $\{\neg A, A \wedge B_1, A \wedge B_2\}$, then with positive probability, refinement resolves the dilemma by revealing a multi-value-dominating action.*

This theorem shows how refinement can fundamentally alter the structure of choice problems. The geometric intuition appears in Figure 4, which visualizes how the two values are realized by different actions. The horizontal axis represents the relative weight of each value, while the vertical axes show the degree to which each action realizes the respective values. Figure 4a depicts the initial dilemma where A_1 favors V_1 while $\neg A_1$ favors V_2 . Figures 4b and 4c illustrate two possible refinements of A_1 . In the first case, refinement provides more options but

preserves the fundamental tension between the values. In the second case, however, refinement reveals action $A_1 \wedge B_1$ that excels along both value dimensions, eliminating the need to determine the relative weight of the values.

The resolution occurs because, in expectation, refinement constitutes a multi-dimensional mean-preserving spread. Just as single-valued refinement creates option value by allowing selection of above-average components, multi-valued refinement can reveal components that simultaneously exceed the average across multiple dimensions. When this occurs, the agent discovers an action that remains optimal regardless of how she might weight her values, transforming an apparent dilemma requiring difficult commensuration judgments into a choice with a clear solution.

This possibility is not guaranteed—refinement may simply provide additional options without revealing a dominating action. Yet when the joint distribution of refined values includes the possibility of multi-value dominance, there exists positive probability that refinement will thread the needle between competing values. The result demonstrates that value refinement affects not merely the expected utility of available actions but can alter the fundamental structure of choice problems themselves. When refinement reveals multi-value dominating actions, it allows agents to bypass rather than resolve their uncertainty about value weights, providing a third path between the horns of a dilemma.

Having established the value of refinement for individual agents, including its potential to resolve apparent dilemmas between incommensurable values, we now turn to strategic settings. Classical value of information results break down in game-theoretic contexts where actions influence the probability of states, potentially making additional information harmful. We demonstrate that value refinement maintains its benefits even in these challenging environments, beginning with the case of two-agent zero-sum games.

6. VALUE REFINEMENT IN TWO-PLAYER ZERO-SUM GAMES

Having established the benefits of value refinement for individual agents, we now examine strategic environments where classical information acquisition can be harmful. Value-of-information guarantees break down in game-theoretic contexts where actions influence state probabilities (Skyrms 1985; Gibbard and Harper 1978). When players' actions are interdependent, additional information can decrease expected utility by producing unfavorable responses from opponents. Given this breakdown, one might expect value refinement to offer no advantage in strategic settings. We demonstrate otherwise: a *unilateral* value refinement generically creates first-order gains at equilibrium, and—because realizations need not preserve zero-sum structure—these gains can manifest as positive-sum opportunities.

Consider a baseline two-player 2×2 game G_0 with actions $\{A^1, \neg A^1\}$ and $\{A^2, \neg A^2\}$. At $\varepsilon = 0$ the game is zero-sum, $U_1^0 + U_2^0 \equiv 0$, and admits a unique interior mixed Nash equilibrium $(\alpha^*, \beta^*) \in (0, 1)^2$.²² Player 1 then *refines* action A^1 into $\{A^1 \wedge B_1, A^1 \wedge B_2\}$ and the realized payoffs of the refined rows are perturbed

²²Given the initial, zero-sum structure, this will be the min-max solution to the game (von Neumann and Morgenstern 1944).

			A^2		$\neg A^2$		
			A^1	$\neg A^1$			
Original Game G_0	A^1	$(v, -v)$	$(\alpha, -\alpha)$				
	$\neg A^1$	$(\beta, -\beta)$	$(\gamma, -\gamma)$				
			A^2		$\neg A^2$		
$A^1 \wedge B_1$			$(-v + \epsilon_{11}^2, v + \epsilon_{11}^1)$	$(-\alpha + \epsilon_{21}^2, \alpha + \epsilon_{21}^1)$			
Refined Game G_1	$A^1 \wedge B_2$	$(-v + \epsilon_{12}^2, v + \epsilon_{12}^1)$	$(-\alpha + \epsilon_{22}^2, \alpha + \epsilon_{22}^1)$				
	$\neg A^1$	$(-\beta, \beta)$	$(-\gamma, \gamma)$				

FIGURE 5. Transformation of a 2×2 zero-sum game into a 2×3 game through refinement. The original action profile (A^1, A^2) with payoffs $(v, -v)$ expands into a 2×2 subgame (upper-left quadrant of refined game) where payoffs are perturbed by independent noise terms ϵ_{ij}^i .

at order $\varepsilon > 0$. Formally, for $C \in \{B_1, B_2\}$ and $a^2 \in \{A^2, \neg A^2\}$,

$$\begin{aligned} U_1^\varepsilon(A^1 \wedge C, a^2) &= U_1^0(A^1, a^2) + \varepsilon \eta_C^1(a^2), & U_i^\varepsilon(\neg A^1, a^2) &= U_i^0(\neg A^1, a^2). \\ U_2^\varepsilon(A^1 \wedge C, a^2) &= U_2^0(A^1, a^2) + \varepsilon \eta_C^2(a^2), \end{aligned}$$

Given RRP, the random vectors $\{\eta_C^i(A^2), \eta_C^i(\neg A^2)\}$ encoding the agent's uncertainty about the changes to individual payoffs have mean zero, are absolutely continuous (so ties occur with probability 0). We also assume, for simplicity, that payoff perturbations are independent of one another.²³ Thus, while G_0 is zero-sum, a realized G_ε is *generically* non-zero-sum; zero-sum holds only in expectation over refinement uncertainty.²⁴

Figure 5 illustrate the refinement. Here, row player uliterally refines act A^1 into $\{A^1 \wedge B_1, A^1 \wedge B_2\}$ with the expected payoffs represented in terms of continuous, symmetric, and independent perturbations of those of the original act. With this, new Nash equilibria have the potential to emerge—potentially better or worse than before. However, rational agents will best respond, take advantage of opportunities for gain. This will tend to transform the best equilibrium payoffs from zero-sum to positive sum.

Theorem 17 (Zero-Sum Escape from Unilateral Value Refinement). *Consider a baseline 2×2 zero-sum game G_0 with two agents facing a zero-sum base game G_0 . Let both players have independent refinement uncertainty over their rationalizable acts and satisfy RRP such that their uncertainty regarding refinement outcomes is representable as defined over perturbations $\{\eta_C^i\}$ to their payoffs that are independent, continuous, and mean-zero. If a player unilaterally refines such an act A to $\{A \wedge B_1, A \wedge B_2\}$, then both player's expected gain at the welfare-optimal Nash equilibrium W_i is strictly positive:*

$$\mathbb{E}_{\mu_A}[W_i^*(\varepsilon) - W_i^*(0)] > 0 \text{ for } i \in \{1, 2\}.$$

²³Note that, if the players have strong beliefs that are correlated such that any refinement will preserve the zero-sum structure, then the game is guaranteed to remain the same and our results do not hold.

²⁴Of course, the assumption that the resulting interaction is likely to be non-zero-sum will not always be appropriate. For example, in games like chess, poker, and the like, we artificially induce a zero-sum structure—any additional actions will continue to be zero-sum within the frame of those games. The rigid maintenance of zero-sum structures is less likely, however, in the context of interactions in markets, diplomacy, or evolution.

Theorem 18 (Zero-Sum Escape from Unilateral Value Refinement). *Consider a baseline 2×2 zero-sum game G_0 with unique interior mixed Nash equilibrium $(\alpha^*, \beta^*) \in (0, 1)^2$. Let both players have refinement uncertainty over their rationalizable acts and satisfy RRP such that their uncertainty regarding refinement outcomes is representable as defined over perturbations $\{\eta_C^i\}$ to their payoffs that are independent, continuous, and mean-zero. Let A^1 be such act such it may be refined to $\{A^1 \wedge B_1, A^1 \wedge B_2\}$. Define the welfare-optimal Nash equilibrium as the equilibrium maximizing total surplus:*

$$(\alpha^w(\varepsilon), \beta^w(\varepsilon)) = \arg \max_{(\alpha, \beta) \in \mathcal{E}(\varepsilon)} [U_1^\varepsilon(\alpha, \beta) + U_2^\varepsilon(\alpha, \beta)]$$

where $\mathcal{E}(\varepsilon)$ is the set of Nash equilibria of G_ε . Then both players' expected gains at the welfare-optimal equilibrium are strictly positive:

$$\mathbb{E}[W_i(\varepsilon) - W_i(0)] > 0 \quad \text{for } i \in \{1, 2\}$$

where $W_i(\varepsilon)$ denotes player i 's payoff at the welfare-optimal equilibrium.

Even when only one player refines, the best equilibrium payoff for that player increases at a strictly positive first-order rate. Intuitively, the refinement is expected to split A^1 into two variants; against β^* on average one is better, and equilibrium stability keeps the opponent's mix near β^* . Positivity follows from $\mathbb{E}[\max\{X_{A^1 \wedge B_1}, X_{A^1 \wedge B_2}\}] = \frac{1}{2}\mathbb{E}[|X_{A^1 \wedge B_1} - X_{A^1 \wedge B_2}|] > 0$ for i.i.d. continuous mean-zero changes to payoffs $X_{A^1 \wedge B_1}, X_{A^1 \wedge B_2}$.

The broader lesson is that apparent zero-sum conflicts are fragile, or non-robust to fine-graining of one's options. Unilateral refinement typically reveals locally superior implementations of an action, and equilibrium responds smoothly: the opponent's best reply barely shifts at first order, so the refiner captures the revealed advantage. Because realized games need not be zero-sum, these gains often coincide with *positive-sum* equilibria, creating coordination possibilities absent at the coarse level.

Nidtya

7. VALUE REFINEMENT IN NASH BARGAINING

We now turn to analyze value refinement in Nash bargaining with a constrained resource. When a one-dimensional split of a resource is refined, it reveals separable dimensions. Refinement makes these dimensions explicit and allocable independently, expanding the feasible utility set. We demonstrate that bargaining solutions in the expanded set strictly dominate those in the original set in expectation for both agents when they value the dimensions differently.

Consider a Nash bargaining game $\mathcal{G}_0 = (\mathcal{N}, \mathcal{X}_0, (u_0^1, u_0^2), \mathcal{F}_0, d)$ where $\mathcal{N} = \{1, 2\}$ denotes the two agents, $\mathcal{X}_0 = \{x \in [0, 1]\}$ represents allocations with agent 1 receiving fraction x and agent 2 receiving fraction $1 - x$, and $u_0^i : [0, 1] \rightarrow \mathbb{R}$ are their utility functions. The feasible utility set is:

$$\mathcal{F}_0 = \{(u_0^1(x), u_0^2(1 - x)) : x \in [0, 1]\} \subset \mathbb{R}^2$$

with disagreement point $d = (d_1, d_2)$. We assume each u_0^i is continuous, increasing, and concave in the share, ensuring $\mathcal{F}_0$ is compact, convex, and comprehensive.²⁵

²⁵A set $\mathcal{F} \subset \mathbb{R}^2$ is *comprehensive* above d if whenever $(u_1, u_2) \in \mathcal{F}$ and $d_1 \leq u'_1 \leq u_1$, $d_2 \leq u'_2 \leq u_2$, then $(u'_1, u'_2) \in \mathcal{F}$.

The Nash bargaining solution is the canonical solution concept for such games and the one that uniquely satisfies the requirements of Pareto optimality, symmetry, independence of irrelevant alternatives, and invariance to affine transformations of utility.²⁶²⁷ The Nash bargaining solution on $(\mathcal{F}_0, d)$ is:

$$f^{\text{NB}}(\mathcal{F}_0, d) = \arg \max_{(u_1, u_2) \in \mathcal{F}_0} (u_1 - d_1)(u_2 - d_2)$$

yielding allocation x^* with utilities $u_0^1(x^*)$ and $u_0^2(1 - x^*)$ for players 1 and 2, respectively.

As before, refinement partitions the actions available to the agents. Here, it transforms the Nash bargaining game $\mathcal{G}_0$ into a refined game

$$\mathcal{G}_1(\omega) = (\mathcal{N}, \mathcal{X}_1, (u^1(\cdot, \omega), u^2(\cdot, \omega)), \mathcal{F}_1(\omega), d).$$

This is formally equivalent to simultaneous binary refinements of the continuous set of possible allocations in such a way that preserves the basic features of the feasible utility set. Specifically, the allocation space expands from one-dimensional to two-dimensional evaluations of the resource:

$$\mathcal{X}_1 = \{(x_1, x_2) \in [0, 1]^2\}$$

where agent 1 receives (x_1, x_2) and agent 2 receives $(1 - x_1, 1 - x_2)$. Each dimension can now be allocated independently.

The refined feasible utility set becomes:

$$\mathcal{F}_1(\omega) = \{(u^1(x_1, x_2, \omega), u^2(1 - x_1, 1 - x_2, \omega)) : (x_1, x_2) \in [0, 1]^2\}$$

One can represent pre-refinement allocations in the refined space as ‘bundles’ as follows:

$$\mathcal{B}(\omega) = \{(u^1(x, x, \omega), u^2(1 - x, 1 - x, \omega)) : x \in [0, 1]\}$$

where, by construction, $\mathcal{B}(\omega) \subseteq \mathcal{F}_1(\omega)$ for all ω . This set inclusion is fundamental: the refined game allows all bundled allocations plus additional possibilities where dimensions are allocated independently.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space representing all possible refinements. Each $\omega \in \Omega$ represents one specific realization of how utilities decompose across dimensions. For each agent $i \in \{1, 2\}$, we define measurable functions:

- $w_j^i : \Omega \rightarrow [0, 1]$ for $j \in \{1, 2\}$ represent agent i ’s weights on each dimension, with $w_1^i(\omega) + w_2^i(\omega) = 1$
- $v_j^i : \Omega \rightarrow C^2([0, 1], \mathbb{R}_+)$ represent agent i ’s utility functions for each dimension j , with each $v_j^i(\cdot, \omega)$ strictly increasing and concave

Each agent has a credal distribution μ^i over Ω capturing their uncertainty about these components. We assume μ^1 and μ^2 are independent and satisfy refinement uncertainty: they are continuous and mutually absolutely continuous

²⁶Formally: (i) *Pareto efficiency*: no $u' \in \mathcal{F}$ with $u' \geq u$ (componentwise) and $u' \neq u$; (ii) *Symmetry*: if $\mathcal{F}$ and d are symmetric under player permutation, then $u_1 = u_2$; (iii) *Invariance*: unchanged under strictly positive affine transformations $u_i \mapsto a_i u_i + b_i$; (iv) *Independence of irrelevant alternatives*: if u is the solution for $(\mathcal{F}, d)$ and $u \in \mathcal{F}' \subseteq \mathcal{F}$, then u is also the solution for $(\mathcal{F}', d)$. These jointly imply $u = \arg \max_{v \in \mathcal{F}, v \geq d} (v_1 - d_1)(v_2 - d_2)$.

²⁷Other solution concepts include the Kalai-Smorodinsky solution (which maintains proportional gains), the egalitarian solution (which equalizes gains), and the utilitarian solution (which maximizes the sum of player utilities).

with Lebesgue measure on an open subset of the parameter space. Let $\nu = \mu^1 \times \mu^2$ denote the product measure.

Upon refinement at realization ω , utilities become additively separable:

$$u^i(x_1, x_2, \omega) = w_1^i(\omega) \cdot v_1^i(x_1, \omega) + w_2^i(\omega) \cdot v_2^i(x_2, \omega)$$

The weight divergence angle $\theta(\omega) = \arccos(\mathbf{w}^1(\omega) \cdot \mathbf{w}^2(\omega))$ where $\mathbf{w}^i = (w_1^i(\omega), w_2^i(\omega))$ measures preference misalignment. Since weights are non-negative and sum to one, $\theta \in [0, \pi/2]$, with $\theta = 0$ indicating identical weights and $\theta = \pi/2$ indicating orthogonal (complementary) preferences.

Application of the Refinement Reflection Principle here requires that for each allocation $x \in [0, 1]$:

$$u_0^i(x) = \mathbb{E}_{\mu^i}[u^i(x, x, \omega)] = \mathbb{E}_{\mu^i}[w_1^i(\omega) \cdot v_1^i(x, \omega) + w_2^i(\omega) \cdot v_2^i(x, \omega)]$$

This states that pre-refinement utilities equal expected utilities under forced bundling—when both dimensions must be allocated in the same proportion. The RRP ensures coherence between the original and refined games, ensuring that the agents' current valuations of a joint allocation are consistent with their subjective expected values for its decomposition.

We can now state our main result for Nash bargaining games.

Theorem 19 (Value of Refinement in Nash Bargaining). *Consider a Nash bargaining game $\mathcal{G}_0 = (\mathcal{N}, \mathcal{X}_0, (u_0^1, u_0^2), \mathcal{F}_0, d)$ where agents allocate a one-dimensional resource. Let agents have independent refinement uncertainties μ^1, μ^2 satisfying RRP with continuous distributions. If with positive $\nu = \mu^1 \times \mu^2$ probability the agents' weight vectors differ ($\mathbf{w}^1(\omega) \neq \mathbf{w}^2(\omega)$), then refinement produces strict Pareto improvement in expected payoffs:*

$$\mathbb{E}_\nu[u_1^{i*}(\omega)] > u_0^{i*} \quad \text{for each } i \in \{1, 2\}$$

where u_0^{i*} and $u_1^{i*}(\omega)$ denote agent i 's payoffs in the Nash solutions of $\mathcal{G}_0$ and $\mathcal{G}_1(\omega)$ respectively.

The theorem establishes that both agents strictly benefit in expectation from dimensional discovery. The mechanism combines two key insights. First, refinement expands the allocation space from the line segment $\{(x, 1 - x) : x \in [0, 1]\}$ to the square $\{(x_1, x_2, 1 - x_1, 1 - x_2) : (x_1, x_2) \in [0, 1]^2\}$, enabling specialization where agents trade dimensions rather than accepting proportional division. Second, the Nash bargaining solution's multiplicative structure ensures that when the feasible set expands and the Nash product increases, both agents benefit from the improvement.

The proof (see Appendix) establishes that the bundled allocation (x^*, x^*) satisfies the first-order conditions for optimality in the refined game only on a measure-zero set of parameters. Maintaining proportional division requires a knife-edge condition where the marginal rates of substitution across dimensions perfectly align—an event occurring with probability zero under continuous distributions. When this condition fails, the Nash solution of the refined game strictly improves the Nash product, and the structure of Nash bargaining ensures both agents share in these gains.

We consider the magnitude of gains from refinement. In particular, we find that the benefits of refinement increase with the divergence in agents' preferences:

Proposition 20 (Gains Increase with Preference Divergence). *Consider the setting of Theorem 19 with linear per-dimension utilities $v_j^i(x_j, \omega) = a_j^i(\omega) \cdot x_j$. The refined Nash product $N_1^*(\omega)$ is increasing in the divergence angle $\theta(\omega) = \arccos(\mathbf{w}^1(\omega) \cdot \mathbf{w}^2(\omega))$ and is maximized when $\theta = \pi/2$ (orthogonal preferences).*

With additive dimensions of value, optimal allocation involves complete specialization: each dimension goes entirely to the agent who values it more highly relative to their valuation of other dimensions. As preference divergence increases, agents increasingly specialize in different dimensions, raising both agents' utilities and the Nash product. Maximum gains occur with orthogonal (complementary) preferences, where each agent receives only the dimension they value.

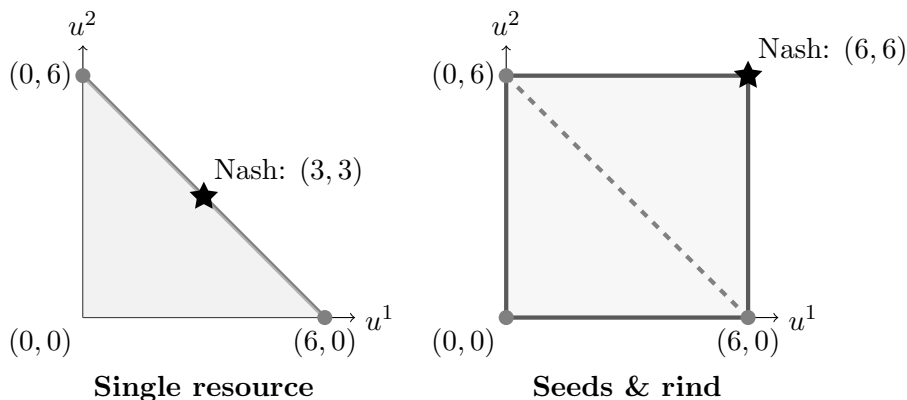


FIGURE 6. Refinement expands the feasible set from a line (left) to a square (right). The dashed line shows bundled allocations corresponding to $\mathcal{B}(\omega)$, where both dimensions must be split proportionally. The expanded shaded region represents the full refined feasible set $\mathcal{F}_1(\omega)$ where dimensions can be allocated independently, enabling the Pareto-superior outcome at $(6, 6)$.

Our result complements classical theorems on the value of information in bargaining and the gains from trade literature. Myerson (1991) showed how private information about valuations affects bargaining outcomes, while Myerson and Satterthwaite (1983) characterized when efficient trade is possible under asymmetric information. Our contribution differs: even with common knowledge of the game structure, discovering latent dimensional structure creates value through expanded allocation possibilities. The gains arise not from resolving informational asymmetries but from recognizing that what appeared as a single-dimensional allocation problem actually permits multi-dimensional trades. This connects to the broader literature on mechanism design with multi-dimensional types (Rochet and Choné 1998) and multi-issue bargaining Fershtman (1990), though our focus on refinement as discovery of dimensional structure is novel.

This application of our refinement framework to Nash bargaining reveals several key insights. First, in the context of Bargaining, refinement can operate as dimensional discovery: the cardinality of the feasible set enlarges because allocations may vary independently across latent dimensions.

With this, both agents can benefit in expectation from pausing to refine their understanding of what they’re negotiating over. Just as individual agents benefit from refining their conception of available acts, bargaining parties benefit from discovering the dimensional structure of their allocations. Once refinement reveals the richer structure of the allocation space, possible opportunities for Pareto improvements emerge for what appeared to be a problem of one-dimensional division.

Further, the magnitude of gains is controlled by preference divergence. As $\theta(\omega)$ increases from 0 (identical preferences) to $\pi/2$ (orthogonal preferences), the potential for mutually beneficial specialization grows. This provides a quantitative measure of how much value refinement can create.

8. IMPLICATIONS AND FUTURE DIRECTIONS

Our results give a mathematical twist to the injunction at Delphi to “know thyself.” If you pause to reflect on your values before you act then, under broad conditions, value refinement at a reasonable cost is *ex ante* beneficial for a rational agent (Theorem 11), can address classical dilemmas about incommensurable values (Theorem 16), turns zero-sum conflicts into positive-sum opportunities (Theorem 17), and expands Nash-bargaining frontiers so that each party can secure a strictly higher expected payoff (Theorem 19). These theorems carry normative, methodological, and practical consequences that reach from individual reflection to multi-agent coordination and institutional design.

Decision makers, whether they are individual human agents, AI systems, organizations or institutions must rely on their values and knowledge of the world to make decisions. Unlike the idealized agents of textbooks, these decision makers can be uncertain about whether their current representation and understanding of their values is adequate. There is thus a tension between taking the time to learn and acting on the values and information at hand. One of the lessons of the present work is that when decision makers find themselves in situations of conflict, whether from individual moral dilemmas or coordination problems in group decisions, it can be beneficial to avoid the temptation to move immediately to tradeoffs. Rather, taking conflict as an occasion to reflect on one’s values can open up the possibility that conflicts can be averted or mitigated and that dilemmas can be avoided or resolved. Failure to reflect on one’s values on such occasions can entrench suboptimal values and lock decision makers into perspectives that do not adequately reflect their own ultimate interests.

The work also extends the scope of decision theory. Savage and von Neumann–Morgenstern treat the act–state–outcome partition as fixed; by contrast, the atomless Boolean algebra of the Jeffrey–Bolker framework lets that partition be indefinitely refined in a single model. We provide a natural formalism for how this refinement might proceed. This flexibility allows us to formalize a deliberative process that is ubiquitous in practice yet underappreciated in traditional theory: interrogating and sharpening one’s evaluative map prior to choice. In doing so, we supply a rational-choice foundation for heuristics that urge parties to separate positions from interests and to explore unconsidered options for mutual gain.

Classical coherence arguments—Dutch books, accuracy dominance, sequential optimality—assume from the outset that an agent possesses a settled utility function. We show that taking such settlement for granted is itself a substantive decision: foregoing refinement can be strictly dominated by clarifying one’s values

first. Rational choice therefore has a two-stage character: refine, then maximize. By revealing that finer descriptions of available acts can expose alternatives that dominate on every value dimension, our analysis shows that value pluralism need not entail decision paralysis, while monism may sacrifice long-run efficacy for short-term consistency.

The results we present are sufficiently general that they apply to approaches to decision making that are weaker than expected utility in that they do not require commensurability of all objects of choice. Our results alone do not say anything in favor or against any of these approaches.

However, our results give rise to what we call the Paradox of Bounded Optimization, which we regard as the computational analogue of the paradox of hedonism. The optimal response to this paradox likely has important implications for the way that realistic agents should think about their utility function.

For our purposes, the paradox of hedonism can be understood as a *reductio* on the combination of two seemingly intuitive claims. The first is the axiological claim that pleasure is the only intrinsic good. The second is the prudential claim that an agent who cares about their own wellbeing should directly seek to maximize the amount of pleasure in their life. The paradox derives from the claim that if an agent pursues an activity purely for the goal of deriving pleasure from it, they are unlikely to achieve their own prudential goal. The reason is that in order to derive pleasure from that activity, the agent must value something other than pleasure, contradicting the axiological claim. For example, deriving pleasure from a game of chess requires that the agent value winning, or elegant play, or bold openings or cunning defenses. Valuing something other than pleasure allows the agent to derive pleasure from an activity but contradicts the axiological claim that only pleasure is intrinsically valuable.

In the broadest sense, our results demonstrate that there are cases where a bounded agent who seeks to maximize expected utility cannot do so by the direct application of the utility function they accept at that time. Rather, the agent will do better by pausing to reflect on the underlying values that are relevant to the problem at hand and of which their overall utility is a function.

On a theoretical level, this aligns with pluralistic approaches to value theory, which hold that acknowledging and consulting a diversity of ultimate values can lead to better outcomes in the long run. This contrasts with monistic approaches, where decision-making is governed by a single utility function formulated to encode choice behavior and avoid inconsistencies. While monistic systems may be less vulnerable to the sorts of incoherence indicated by Dutch book and money-pump arguments, for bounded agents they can fail to capture the full richness of agent's uncertainty regarding their underlying values, or of boundedly rational solutions to decision problems. Refining values not only enriches the state space of possible outcomes, but also transforms the character of strategic interactions.

Moreover, this approach challenges certain reductive interpretations of utility. If utility functions are merely mathematical constructs reflecting underlying values, then what matters are those values themselves. For agents who do not already know all of their values, consulting their values is critical for better realizing those values, as opposed to merely optimizing within a coarse-grained or incomplete framework.

In this sense, the problem of value refinement echoes the paradox of hedonism: just as directly pursuing pleasure may lead to its absence, so too acting on the agent's current utility function may lead to systematically suboptimal outcomes, even by one's own light. Such an agent would be better off reflecting on and refining the values from which that utility function is constructed.

Good-style value-of-information theorems guarantee that learning about the world is worth its price. We have provided the parallel guarantee for learning about, and refining, one's own values. Agents who treat their utility functions as provisional hypotheses, open to disciplined revision, can expect to fare better both individually and collectively. Rationality, then, is not merely the calculation performed with a fixed utility function; it is the iterative work of making that function worthy of calculation.

Our findings thus highlight the importance of investigating ultimate values and refining their representation in decision-making frameworks. By doing so, agents can systematically improve their outcomes and transform adversarial interactions into cooperative ones, advancing both individual and collective well-being.

Important limitations remain. Our results are formulated within the Jeffrey–Bolker decision framework, owing to its congenial treatment of acts as propositions—however, analogous results could be formulated in the decision frameworks of Savage, von Neumann and Morgenstern, Anscombe and Aumann, and others. We assumed negligible cognitive and temporal costs; introducing stochastic refinement costs would permit an optimal-stopping analysis of how much reflection suffices. The strategic results relied on common probabilistic beliefs; relaxing that assumption may attenuate the positive-sum guarantee, and bounding the resulting effect sizes remains an open problem. Furthermore, in iterated interactions, early-round refinement shapes later payoffs and information flows; richer dynamics—including signaling about one's value-learning intentions—await investigation.

APPENDIX A. MATHEMATICAL APPENDIX

Proof of Theorem 11 (Value of Value Refinement). Let $V_0 = \max_{A \in \mathbb{A}_0} U_0(A)$ denote the agent's optimal value before refinement, let $V'_0 = \max_{A \in \mathbb{A}_0 \setminus \{A^*\}} U_0(A)$ denote the best value excluding the refined act, and let $V_1 = \max\{V'_0, u_1, u_2\}$ denote the optimal value after refinement, where $(u_1, u_2, p_1, p_2) \sim \mu_{A^*}$ with $u_1 = U(A^* \wedge B_1)$, $u_2 = U(A^* \wedge B_2)$, and we define $p = \frac{p_1}{p_1 + p_2}$ as the conditional probability $P(B|A^*)$.

We must show that $\mathbb{E}_{\mu_{A^*}}[V_1] > V_0$. Since A^* is rationalizable, we have $U_0(A^*) = V_0$. By the Refinement Reflection Principle,

$$U_0(A^*) = \mathbb{E}_{\mu_{A^*}}[p \cdot u_1 + (1 - p) \cdot u_2].$$

The expected gain from refinement equals

$$\begin{aligned} \mathbb{E}_{\mu_{A^*}}[V_1] - V_0 &= \mathbb{E}_{\mu_{A^*}}[\max\{V'_0, u_1, u_2\}] - \mathbb{E}_{\mu_{A^*}}[p \cdot u_1 + (1 - p) \cdot u_2] \\ &\geq \mathbb{E}_{\mu_{A^*}}[\max\{u_1, u_2\}] - \mathbb{E}_{\mu_{A^*}}[p \cdot u_1 + (1 - p) \cdot u_2]. \end{aligned}$$

For any realization (u_1, u_2, p) , we have

$$\max\{u_1, u_2\} \geq p \cdot u_1 + (1 - p) \cdot u_2,$$

with equality only when $u_1 = u_2$. Since μ_{A^*} has continuous marginals, we know that $\mathbb{P}(u_1 = u_2) = 0$. Hence the inequality is strict almost surely. By dominated

convergence,

$$\mathbb{E}_{\mu_{A^*}}[\max\{u_1, u_2\}] > \mathbb{E}_{\mu_{A^*}}[p \cdot u_1 + (1 - p) \cdot u_2].$$

Therefore $\mathbb{E}_{\mu_{A^*}}[V_1] > V_0$, as desired. $\square$

Proof of Theorem 14 (Optimal Refinement with Fixed Costs). The agent faces sequential decisions where at stage i , having incurred total cost $i \cdot k$, she chooses whether to refine by paying k for expected gain Δ_{R_i} or to stop and act.

If $k > \Delta_{R_0}$, then by vanishing returns we have $k > \Delta_{R_i}$ for all $i \geq 0$. Each refinement would yield negative net value, so the optimal policy is immediate action.

If $k \leq \Delta_{R_0}$, we define the stopping time

$$t^* = \max\{t \in \mathbb{N} : \Delta_{R_t} \geq k\}.$$

This stopping time is well-defined since $\Delta_{R_0} \geq k$ and $\lim_{i \rightarrow \infty} \Delta_{R_i} = 0 < k$.

By construction, for $i < t^*$ we have $\Delta_{R_i} > k$, yielding positive net gain from refinement. At $i = t^*$ we have $\Delta_{R_{t^*}} \geq k$, yielding non-negative net gain. For $i > t^*$ we have $\Delta_{R_i} < k$, yielding negative net gain. By backward induction, the optimal policy refines exactly t^* times. The total net gain equals

$$\sum_{i=0}^{t^*-1} \Delta_{R_i} - t^* \cdot k = \sum_{i=0}^{t^*-1} (\Delta_{R_i} - k) > 0,$$

since each term satisfies $\Delta_{R_i} - k \geq 0$ with strict inequality for $i < t^*$. $\square$

Proof of Theorem 16 (Dilemma Resolution). Let $(v_{1,1}, v_{1,2}, p_1) \sim \mu_A^{(1)}$ and $(v_{2,1}, v_{2,2}, p_2) \sim \mu_A^{(2)}$ represent refinement uncertainties for V_1 and V_2 respectively. An action multi-value dominates when it achieves the maximum along both value dimensions. This occurs when either

$$(D1): \quad v_{1,1} \geq \max\{V_1(\neg A), v_{1,2}\} \text{ and } v_{2,1} \geq \max\{V_2(\neg A), v_{2,2}\}$$

$$(D2): \quad v_{1,2} \geq \max\{V_1(\neg A), v_{1,1}\} \text{ and } v_{2,2} \geq \max\{V_2(\neg A), v_{2,1}\}$$

with at least one strict inequality in each case.

By hypothesis, the support of the joint distribution includes realizations satisfying (D1) or (D2). Since the marginal distributions are continuous, there exists an open set $\Omega^* \subset \mathbb{R}^4$ with $\mathbb{P}((v_{1,1}, v_{1,2}, v_{2,1}, v_{2,2}) \in \Omega^*) > 0$ where multi-value dominance occurs.

When (D1) holds, action $A \wedge B_1$ satisfies

$$U_w(A \wedge B_1) = (1 - w)V_1(A \wedge B_1) + wV_2(A \wedge B_1) \geq U_w(A')$$

for all $A' \in \{\neg A, A \wedge B_2\}$ and all weights $w \in [0, 1]$. A similar statement holds for (D2) with $A \wedge B_2$. Therefore we have $\mathbb{P}(\text{dilemma resolved}) = \mathbb{P}((D1) \cup (D2)) > 0$, as desired. $\square$

Proof of Theorem 17 (Zero-Sum Escape). We consider a baseline 2×2 zero-sum game G_0 with actions $\{A^1, \neg A^1\} \times \{A^2, \neg A^2\}$ and payoffs satisfying $U_1^0 + U_2^0 \equiv 0$. The unique interior mixed Nash equilibrium is (α^*, β^*) , where α^* denotes player 1's probability on A^1 and β^* denotes player 2's probability on A^2 . Player 1 refines action A^1 into two variants $\{A^1 \wedge B_1, A^1 \wedge B_2\}$, transforming the game into a 3×2 game G_ε . The refinement introduces perturbations $\eta_C^i(a^2)$ for $i \in \{1, 2\}$,

$C \in \{B_1, B_2\}$, and $a^2 \in \{A^2, \neg A^2\}$, which are independent across players and variants, have continuous distributions, and by RRP satisfy $\mathbb{E}[\eta_C^i(a^2)] = 0$. Our goal is to show that both players benefit in expectation at the welfare-optimal Nash equilibrium of the perturbed game.

We begin by characterizing the structure of the perturbed game. After player 1's refinement, G_ε becomes a 3×2 game with payoffs:

For $(A^1 \wedge C, a^2)$ where $C \in \{B_1, B_2\}$ and $a^2 \in \{A^2, \neg A^2\}$:

$$U_1^\varepsilon(A^1 \wedge C, a^2) = U_1^0(A^1, a^2) + \varepsilon \eta_C^1(a^2),$$

$$U_2^\varepsilon(A^1 \wedge C, a^2) = U_2^0(A^1, a^2) + \varepsilon \eta_C^2(a^2) = -U_1^0(A^1, a^2) + \varepsilon \eta_C^2(a^2).$$

And for $(\neg A^1, a^2)$, we have $U_i^\varepsilon(\neg A^1, a^2) = U_i^0(\neg A^1, a^2)$.

Since η_C^1 and η_C^2 are independent, the perturbed game is generically non-zero-sum.

To understand the equilibrium structure near $\varepsilon = 0$, we examine which refined variant player 1 prefers. Define the payoff difference between refined variants for player 1:

$$\begin{aligned} \Delta_1(\beta, \varepsilon) &= U_1^\varepsilon(A^1 \wedge B_1, \beta) - U_1^\varepsilon(A^1 \wedge B_2, \beta) \\ &= \varepsilon[\beta(\eta_{B_1}^1(A^2) - \eta_{B_2}^1(A^2)) + (1 - \beta)(\eta_{B_1}^1(\neg A^2) - \eta_{B_2}^1(\neg A^2))]. \end{aligned}$$

At the baseline equilibrium mix β^* , player 1 was indifferent between A^1 and $\neg A^1$. After refinement, define:

$$X = \eta_{B_1}^1(A^2)\beta^* + \eta_{B_1}^1(\neg A^2)(1 - \beta^*) - [\eta_{B_2}^1(A^2)\beta^* + \eta_{B_2}^1(\neg A^2)(1 - \beta^*)].$$

Since the η terms are continuous and independent, $\mathbb{P}(X = 0) = 0$. Thus, in expectation, one variant is strictly preferred at β^* . Let $C^* = \arg \max_{C \in \{B_1, B_2\}} U_1^\varepsilon(A^1 \wedge C, \beta^*)$.

We now characterize equilibria for small ε . For small $\varepsilon > 0$, any Nash equilibrium near (α^*, β^*) has player 1 mixing between at most two actions: $\{A^1 \wedge C^*, \neg A^1\}$. This reduces to analyzing a 2×2 subgame.

The indifference conditions for an interior equilibrium are:

$$\begin{aligned} \phi_1(\alpha, \beta, \varepsilon) &= U_1^\varepsilon(A^1 \wedge C^*, \beta) - U_1^\varepsilon(\neg A^1, \beta) = 0, \\ \phi_2(\alpha, \beta, \varepsilon) &= U_2^\varepsilon(\alpha, A^2) - U_2^\varepsilon(\alpha, \neg A^2) = 0, \end{aligned}$$

where for player 2's condition, α now represents the probability on $A^1 \wedge C^*$ (with remaining mass on $\neg A^1$).

By the implicit function theorem, since $\nabla_{(\alpha, \beta)} \phi|_{(\alpha^*, \beta^*, 0)}$ is non-singular (from uniqueness of the baseline equilibrium), there exists a unique continuation $(\alpha(\varepsilon), \beta(\varepsilon))$ with:

$$\begin{aligned} \alpha(\varepsilon) &= \alpha^* + \varepsilon \alpha_1 + o(\varepsilon), \\ \beta(\varepsilon) &= \beta^* + \varepsilon \beta_1 + o(\varepsilon). \end{aligned}$$

To analyze player 1's payoff at first order, note that player 1's expected payoff at this equilibrium is:

$$V_1(\varepsilon) = \alpha(\varepsilon)U_1^\varepsilon(A^1 \wedge C^*, \beta(\varepsilon)) + (1 - \alpha(\varepsilon))U_1^\varepsilon(\neg A^1, \beta(\varepsilon)).$$

Since player 1 is indifferent at equilibrium: $U_1^\varepsilon(A^1 \wedge C^*, \beta(\varepsilon)) = U_1^\varepsilon(\neg A^1, \beta(\varepsilon))$, we have:

$$V_1(\varepsilon) = U_1^\varepsilon(\neg A^1, \beta(\varepsilon)) = U_1^0(\neg A^1, \beta(\varepsilon)).$$

We can also evaluate using the fact that C^* was chosen optimally:

$$V_1(\varepsilon) - V_1(0) = \alpha^* \varepsilon \mathbb{E}[\max\{\eta_{B_1}^1 \cdot \beta^* + \eta_{B_1}^1 \cdot (1 - \beta^*), \eta_{B_2}^1 \cdot \beta^* + \eta_{B_2}^1 \cdot (1 - \beta^*)\}] + o(\varepsilon).$$

Since $\mathbb{E}[\max\{Y_1, Y_2\}] > 0$ for independent, continuous, mean-zero Y_1, Y_2 , player 1 gains at first order.

Turning to player 2's payoff, the expected payoff is:

$$V_2(\varepsilon) = \alpha(\varepsilon)U_2^\varepsilon(A^1 \wedge C^*, \beta(\varepsilon)) + (1 - \alpha(\varepsilon))U_2^\varepsilon(\neg A^1, \beta(\varepsilon)).$$

The perturbations affecting player 2's payoffs are:

$$U_2^\varepsilon(A^1 \wedge C^*, a^2) = -U_1^0(A^1, a^2) + \varepsilon \eta_{C^*}^2(a^2).$$

At first order:

$$V_2(\varepsilon) - V_2(0) = \alpha^* \varepsilon [\beta^* \eta_{C^*}^2(A^2) + (1 - \beta^*) \eta_{C^*}^2(\neg A^2)] + o(\varepsilon).$$

Since $\eta_{C^*}^2$ is independent of $\eta_{C^*}^1$ and has mean zero with continuous distribution:

$$\mathbb{E}[V_2(\varepsilon) - V_2(0)] = \alpha^* \varepsilon \cdot 0 + o(\varepsilon) = o(\varepsilon).$$

The above analysis considered a particular equilibrium continuation. We now demonstrate that welfare optimization creates strict gains for both players in expection. The total welfare at any equilibrium is:

$$W(\varepsilon) = V_1(\varepsilon) + V_2(\varepsilon).$$

At the baseline, $W(0) = V_1(0) + V_2(0) = 0$ (zero-sum game).

For the perturbed game, the sum of payoffs for action profile $(A^1 \wedge C, a^2)$ is:

$$U_1^\varepsilon(A^1 \wedge C, a^2) + U_2^\varepsilon(A^1 \wedge C, a^2) = \varepsilon [\eta_C^1(a^2) + \eta_C^2(a^2)].$$

Since η_C^1 and η_C^2 are independent and continuous, with positive probability we have $\eta_C^1(a^2) + \eta_C^2(a^2) > 0$ for some (C, a^2) . This creates positive-sum outcomes.

The welfare-optimal equilibrium selects among all Nash equilibria to maximize total surplus. With positive probability, there exists an equilibrium placing positive weight on positive-sum outcomes, yielding:

$$\mathbb{E}[W(\varepsilon)] > 0.$$

It remains to show that both players expect to benefit at the welfare-optimal equilibrium. At the welfare-optimal equilibrium with $W(\varepsilon) > 0$, suppose player 2's payoff decreased: $W_2(\varepsilon) < W_2(0) = -V_1(0)$. Then:

$$W_1(\varepsilon) = W(\varepsilon) - W_2(\varepsilon) > 0 - (-V_1(0)) = V_1(0).$$

But player 1 could guarantee at least $V_1(0) + \varepsilon \cdot \mathbb{E}[\max\{X_{B_1}, X_{B_2}\}] + o(\varepsilon)$ by optimal play. If the welfare-optimal equilibrium gave player 2 less than her baseline payoff, player 1 would have incentive to deviate, contradicting equilibrium.

By continuity of the Nash correspondence and the fact that both players can guarantee approximately their baseline payoffs (through strategies close to (α^*, β^*)), the welfare gains must be shared. Specifically, for small ε :

$$W_i(\varepsilon) \geq W_i(0) + c_i \varepsilon + o(\varepsilon)$$

where $c_1, c_2 > 0$ and $c_1 + c_2 = \mathbb{E}[\max_{C, a^2} [\eta_C^1(a^2) + \eta_C^2(a^2)]] > 0$.

Therefore, $\mathbb{E}[W_i(\varepsilon) - W_i(0)] > 0$ for both $i \in \{1, 2\}$. $\square$

Proof of Theorem 19 (Nash Bargaining). We consider the baseline Nash bargaining game $\mathcal{G}_0$ where agents divide a one-dimensional resource with allocation $x \in [0, 1]$ giving agent 1 share x and agent 2 share $1 - x$, with utilities $u_0^i(x)$ that are continuous, increasing, and concave. The Nash solution selects a unique x^* , with payoffs $u_0^{i*} = u_0^i(x^*)$.

Refinement reveals two latent dimensions of the resource. Each agent i has weights $\mathbf{w}^i(\omega) = (w_1^i(\omega), w_2^i(\omega))$ and strictly increasing, concave per-dimension utilities $v_j^i(\cdot, \omega)$, all drawn independently from continuous distributions. A refined allocation $(x_1, x_2) \in [0, 1]^2$ yields

$$u_1^i(x_1, x_2, \omega) = w_1^i(\omega) v_1^i(x_1, \omega) + w_2^i(\omega) v_2^i(x_2, \omega).$$

The Refinement Reflection Principle requires

$$u_0^i(x) = \mathbb{E}_{\mu^i}[u_1^i(x, x, \omega)] \quad \text{for all } x \in [0, 1].$$

We assume that the agents' weight vectors differ with positive probability.

We now show that refinement strictly increases each agent's expected Nash payoff. Let x^* be the Nash solution in the unrefined game $\mathcal{G}_0$. For each refinement state ω , define

$$b_i(\omega) := u_1^i(x^*, x^*, \omega),$$

the payoff agent i would receive if the allocation in the refined game were kept bundled. By the Refinement Reflection Principle,

$$\mathbb{E}_\nu[b_i(\omega)] = u_0^{i*}.$$

To compare refinement to the baseline, write each agent's gain in state ω as

$$\Delta_i(\omega) := u_1^{i*}(\omega) - b_i(\omega).$$

Because (x^*, x^*) is feasible for every refined game $\mathcal{G}_1(\omega)$, the refined Nash solution can only do weakly better, so $\Delta_i(\omega) \geq 0$ for all ω .

The next step is to identify when this inequality is strict. Bundling is Nash-optimal in the refined game only if the two agents' marginal rates of substitution across the revealed dimensions match exactly at (x^*, x^*) . This condition depends smoothly on a finite list of state variables (the agents' weights and marginal utilities).

Since these variables are independently and continuously distributed, the probability that the required equality holds is zero. Thus there is a set of states, call it Ω^+ , with strictly positive probability, on which bundling is not optimal for the refined game.

On any state in Ω^+ where the agents' weight vectors differ, their marginal rates of substitution differ at (x^*, x^*) . Together with strict monotonicity and concavity of each v_j^i , this ensures the existence of a small, feasible reallocation across dimensions that raises both agents' utilities relative to bundling.

Since the Nash product is strictly increasing in each argument above the disagreement point, and the feasible set contains a point that strictly Pareto-dominates $(b_1(\omega), b_2(\omega))$, the Nash maximizer must lie strictly above $(b_1(\omega), b_2(\omega))$ in both coordinates. Hence, for all $\omega \in \Omega^+$,

$$\Delta_i(\omega) > 0 \quad \text{for each } i.$$

Finally, because $\Delta_i(\omega) \geq 0$ everywhere and strictly positive on a set of positive probability, its expectation is strictly positive:

$$\mathbb{E}_\nu[\Delta_i(\omega)] > 0.$$

Using the earlier identity for the expected bundled payoff gives

$$\mathbb{E}_\nu[u_1^{i*}(\omega)] = \mathbb{E}_\nu[b_i(\omega)] + \mathbb{E}_\nu[\Delta_i(\omega)] = u_0^{i*} + \mathbb{E}_\nu[\Delta_i(\omega)] > u_0^{i*}.$$

Thus each agent strictly gains in expectation from refinement. $\square$

Proof of Proposition 20 (Preference Divergence). With linear utilities $v_j^i(x_j, \omega) = a_j^i(\omega)x_j$, the Nash product under allocation (x_1, x_2) becomes

$$N(x_1, x_2) = [w_1^1 a_1^1 x_1 + w_2^1 a_2^1 x_2 - d_1][w_1^2 a_1^2 (1 - x_1) + w_2^2 a_2^2 (1 - x_2) - d_2].$$

Since N is linear in each x_j when the other is fixed, the first-order condition $\partial N / \partial x_j = 0$ cannot hold at interior points. The maximum must occur at a corner of $[0, 1]^2$.

The allocation rule assigns dimension j entirely to agent 1 if $\frac{w_j^1 a_j^1}{w_j^2 a_j^2} > \frac{U^1}{U^2}$ and to agent 2 otherwise, where U^i denotes equilibrium utility minus disagreement.

When $\theta = 0$ and weight vectors are identical, both ratios $\frac{w_j^1}{w_j^2}$ are equal, eliminating gains from specialization. As θ increases toward $\pi/2$, the weight ratios diverge, enabling increasingly efficient specialization. At $\theta = \pi/2$ with orthogonal preferences, complete specialization occurs where each agent receives only their valued dimension, maximizing the Nash product.

Therefore $N^*(\theta)$ is weakly increasing in $\theta \in [0, \pi/2]$ and achieves its maximum at $\theta = \pi/2$, as desired. $\square$

REFERENCES

- Ahmed, A. (2021). *Evidential Decision Theory*. Elements in Decision Theory and Philosophy. Cambridge: Cambridge University Press.
- Balocchi, C., S. Favaro, and Z. Naulet (2025). Bayesian nonparametric inference for “species-sampling” problems. *Statistical Science*. To appear; see IMS Statistical Science Future Papers. Preprint available as arXiv:2203.06076.
- Bjorndahl, A., A. J. London, and K. J. S. Zollman (2017, April). Kantian decision making under uncertainty: Dignity, price, and consistency. *Philosophers’ Imprint* 17(7), 1–22.
- Bolker, E. D. (1966). Functions resembling quotients of measures. *Transactions of the American Mathematical Society* 124(2), 292–312.
- Bradley, R. (2017). *Decision Theory with a Human Face*. Cambridge: Cambridge University Press. Part IV develops the treatment of unawareness and awareness growth as algebra changes.
- de Finetti, B. (1937). La prévision: Ses lois logiques, ses sources subjectives. *Annales de l’Institut Henri Poincaré* 7, 1–68.
- Dewey, J. and J. H. Tufts (1936). *Ethics* (Revised edition ed.). New York: H. Holt.
- Dorst, K. (2024). Reflection without idealisation. manuscript, 2024.
- Fallis, D. (2007). Attitudes toward epistemic risk and the value of experiments. *Studia Logica* 86(2), 215–246.

- Fershtman, C. (1990). The importance of the agenda in bargaining. *Games and Economic Behavior* 2(3), 224–238.
- Gibbard, A. and W. L. Harper (1978). Counterfactuals and two kinds of expected utility. In *Foundations and Applications of Decision Theory*. Reidel.
- Gigerenzer, G. and R. Selten (2002). *Bounded Rationality: The Adaptive Toolbox*. MIT Press.
- Godfrey-Smith, P. (1991). Signal, decision, action. *Journal of Philosophy* 88(12), 709–722.
- Good, I. J. (1967). On the principle of total evidence. *The British Journal for the Philosophy of Science* 17(4), 319–321.
- Greaves, H. and D. Wallace (2006). Justifying conditionalisation. *Mind* 115(459), 607–651.
- Hammond, P. J. (1988). Consequentialist foundations for expected utility. *Theory and Decision* 25(1), 25–78.
- Horowitz, S. (2014). Immoderately rational. *Philosophical Studies* 167(1), 41–56.
- Huttegger, S. M. (2013). In defense of reflection. *Philosophy of Science* 80(3), 413–433.
- Jeffrey, R. C. (1965). *The Logic of Decision*. McGraw-Hill.
- Joyce, J. M. (1998). A nonpragmatic vindication of probabilism. *Philosophy of Science* 65(4), 575–603.
- Kadane, J. B., M. J. Schervish, and T. Seidenfeld (2008). Is ignorance bliss? *Journal of Philosophy* 105(1), 5–36.
- Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review* 93(5), 1449–1475.
- Kemeny, J. G. (1955). Fair bets and inductive probabilities. *Journal of Symbolic Logic* 20(3), 263–273.
- Levi, I. (1990). *Hard choices: Decision making under unresolved conflict*. Cambridge University Press.
- Lippman, S. A. and J. J. McCall (1976). The economics of job search: A survey. *Economic Inquiry* 14(2), 155–189.
- McCall, J. J. (1970). Economics of information and job search. *Quarterly Journal of Economics* 84(1), 113–126.
- McClellenn, E. F. (1990). *Rationality and Dynamic Choice: Foundational Explorations*. Cambridge University Press.
- Mortensen, D. T. (1986). Job search and labor market analysis. In O. Ashenfelter and R. Layard (Eds.), *Handbook of Labor Economics*, Vol. 2, pp. 849–919. Elsevier.
- Myerson, R. B. (1991). *Game Theory: Analysis of Conflict*. Cambridge, Massachusetts: Harvard University Press.
- Myerson, R. B. and M. A. Satterthwaite (1983). Efficient mechanisms for bilateral trading. *Journal of Economic Theory* 29(2), 265–281.
- Paul, L. A. (2014). *Transformative experience*. OUP Oxford.
- Pettigrew, R. (2015). Transformative experience and decision theory.
- Pettigrew, R. (2016). *Accuracy and the Laws of Credence*. Oxford University Press.
- Pettigrew, R. (2019). *Choosing for changing selves*. Oxford University Press.
- Ramsey, F. P. (1931). Truth and probability. In *The Foundations of Mathematics and Other Logical Essays*. Routledge.

- Rochet, J.-C. and P. Choné (1998). Ironing, sweeping, and multidimensional screening. *Econometrica* 66(4), 783–826.
- Savage, L. J. (1954). *The Foundations of Statistics*. Wiley.
- Schoenfield, M. (2017). The accuracy and rationality of imprecise credences. *Noûs* 51(4), 667–685.
- Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics* 69(1), 99–118.
- Skyrms, B. (1985). *Choice and Chance: An Introduction to Inductive Logic*. Wadsworth.
- Skyrms, B. (1990). The value of knowledge. In *Minnesota Studies in the Philosophy of Science*, Volume 14, pp. 245–266. Minneapolis: University of Minnesota Press.
- Steele, K. and H. O. Stefánsson (2021). *Beyond Uncertainty: Reasoning with Unknown Possibilities*. Cambridge Elements in Decision Theory and Philosophy. Cambridge, UK: Cambridge University Press.
- Stigler, G. J. (1961). The economics of information. *Journal of Political Economy* 69(3), 213–225.
- Ullmann-Margalit, E. (2006). Big decisions: Opting, converting, drifting¹. *Royal Institute of Philosophy Supplements* 58, 157–172.
- van Fraassen, B. C. (1984). Belief and the will. In *Midwest Studies in Philosophy IX*, pp. 428–463.
- von Neumann, J. and O. Morgenstern (1944). *Theory of Games and Economic Behavior*. Princeton University Press.
- Walker, S. G. (2013). Bayesian inference with misspecified models. *Journal of Statistical Planning and Inference* 143(10), 1621–1633.