Morality: A Battle Royale

Paul Studtmann[*] Shyam Gouri Suresh[†] Bryce Wiedenbeck[‡]

August 12, 2024

Preliminary and Incomplete.
Please do not cite or circulate.

Abstract

This paper presents a large-scale simulation study evaluating the relative success of competing moral theories in strategic interactions. Rather than relying on abstract philosophical argumentation, we adopt a game-theoretic framework in which agents governed by distinct moral principles interact across the entire range of symmetric 2-player, 2-strategy, normal form games, including the Prisoner's Dilemma, Stag Hunt, and Hawk-Dove. We analyze 32 moral theories, including five foundational frameworks—average consequence utilitarianism, Kantian deontology, selfishness, maxi-min, and empathy-based morality—as well as 26 synthesized moralities formed by combining these base theories. Through Monte Carlo simulations, we determine which moral rules yield the highest expected material payoffs, both in homogeneous interactions (where all agents follow the same rule) and in mixed populations that include selfish agents.

Our findings reveal that synthesized moralities systematically outperform their foundational counterparts, highlighting the practical advantages of moral autonomy. In particular, the synthesis of utilitarianism and deontology (UD) emerges as the *optimific champiom*, achieving the highest payoffs under the assumption of universal adherence. However, in mixed societies where agents must coexist with selfish individuals, a different moral rule, what we call the *realist's champion*, prevails: the synthesis of deontology and selfishness (SD).

These results have significant implications for moral philosophy, challenging the traditional focus on isolated moral theories and suggesting that moral rules are best understood in relation to their interactive success. The study also

[*]Davidson College. pastudtmann@davidson.edu
[†]Davidson College. shgourisuresh@davidson.edu
[‡]Davidson College. brwiedenbeck@davidson.edu

underscores the essential role of deontology in viable moral syntheses, raising fundamental questions about the evolutionary robustness and practical viability of moral systems.

1 Introduction

The contentious nature of moral disputations often renders moral philosophy reminiscent of a boxing match, with arguments for one moral view battling arguments against it, and vice versa. But what if, rather than pitting reasons for and against individual theories, we could stage a contest between the moral theories themselves? Which theory would emerge victorious? This intriguing question naturally prompts another: What would such a contest look like?

Let us suppose that it is not an actual boxing match, but rather a structured interaction between agents, each adhering to one of the competing moral rules. Victory in this context would be determined by which agent achieves the highest expected payoff in the interaction. The winning theory, then, would be the one whose principles lead to better outcomes. With this framework in place, we can return to the central question: Which moral theory would triumph?

In this paper, we present the results of extensive simulations designed to answer precisely this question across a wide array of interactive contexts. These simulations examine 32 distinct moral theories. One might reasonably ask where such a wealth of theories originates, given that far fewer are prominently discussed in the philosophical canon. In answer to this question, we begin with five foundational moral theories that are deeply embedded in the tradition: average consequence utilitarianism (U), Kantian deontology (D), selfishness (S), maxi-min (M), and empathy-based morality (E). These theories represent well-established frameworks within moral philosophy.

Traditionally, philosophers have considered these theories in isolation. However, the idea of a synthesis—familiar from Hegel—invites us to explore combinations of these moralities. Can disparate theories be synthesized into new hybrid frameworks? The answer, as Gouri Suresh and Studtmann (2023) demonstrate, lies in a general class of models of moral autonomy. In these models, agents choose not only strategic actions but also the moral principles governing those actions. Within this framework, synthesizing moral theories becomes straightforward. Combinations can be constructed by incorporating multiple theories into the strategic reasoning of the agent. For example, one might synthesize selfishness and empathy (SE), or create a hybrid of utilitarianism, deontology, and maxi-min (UDM), and so on.

The result is a taxonomy of 26 possible synthesized moralities, which, when added to the original five foundational theories, yields a total of 31. Because we examine every possible pairing of these 31 moralities, we examine the results of 961 simulations. What we present in this paper, then, is not merely a metaphorical boxing match between individual theories, but a battle royale encompassing the entire spectrum of moral frameworks. Through this analysis, we uncover which principles are most effective in a broad range of interactive scenarios, shedding new light on the strengths and weaknesses of competing moralities.

We present the results of simulations in which agents interact across the full spectrum of symmetric dyadic simultaneous perfect information games, which have the form displayed in Table 1.

	$\mathbf{S}_1$	$\mathbf{S}_2$
$\mathbf{S}_1$	A, A	B, C
$\mathbf{S}_2$	C, B	D, D

Table 1. Symmetric Interaction

By addressing the entire class of symmetric dyadic games, our analysis offers a level of generality that surpasses most game-theoretic studies of morality, which typically focus on specific games, such as the Prisoner's Dilemma, the Stag Hunt, or the Nash Bargaining Game. The class of games we consider encompasses the first two of these games as well as others, including Hawk-Dove. The significance of these games in shaping moral systems is widely recognized. According to Curry et al. (2019), these games, together with the Nash Bargaining game, serve as foundational elements of moral systems across diverse cultures. Likewise, Harms and Skyrms (2008) emphasize that understanding the evolution of morality hinges on explaining behaviors such as cooperation in the Prisoner's Dilemma, choosing stag in the Stag Hunt, and achieving equal splits in the symmetric bargaining game. This study aims to identify a universal principle that yields optimal outcomes across all dyadic symmetric games, rather than focusing on rules specific to individual types of symmetric interactions.

The structure of this paper is as follows. In Section 2, we provide a detailed exposition of the models that form the basis of the analysis. In section 3, we discuss the simulation design and methodology, outlining the framework and assumptions that underpin our analysis. The appendix presents a comprehensive table that captures the expected values of interactions for every possible pairing of agent types.

Sections 4 and 5 discuss two particularly significant rankings derived from the data. The first ranking assesses the expected values of interactions in which both agents adhere to the same moral rule. This ranking is especially relevant for identifying what Parfit (2011) calls the *optimific* rule, the rule that yields the best results under the assumption of universal adherence. From this ranking, we will identify the *optimific champion*, the rule that triumphs when interacting solely with others of its kind.

The second ranking examines the expected values of the interactions in which one agent follows a given rule, X, while the other operates according to the rule of selfishness. This ranking is essential for identifying the optimal rule in a more realistic context, where interactions with self-interested agents are inevitable. The winner of this ranking can be considered a more robust victor, overcoming the challenges of competing against a fundamentally different adversarial type. We shall refer to this victor as the *realist's champion*, in contrast to the optimific champion, who benefits from the more favorable conditions of interaction with its own kind.

We conclude with philosophical reflections on the broader implications of these findings.

2 Models

Our modeling framework comprises two primary components. First, we define five base moralities by extremizing parameters that represent distinct moral preferences. Second, we extend standard game–theoretic models to incorporate autonomous moral choice.

2.1 Base Moralities

2.1.1 Kantian Morality

To capture Kantian morality, we adopt the Kantian preference parameter, κ_i, as introduced by Alger and Weibull (2013). In their formulation, an agent's preferences are modified by incorporating a Kantian component. Formally, the utility function for agent i is given by

$$V_i(a_i, a_{-i}) = (1 - \kappa_i)\, u_i(a_i, a_{-i}) + \kappa_i\, u_i(a_i, a_i, \ldots, a_i),$$

where:

- $u_i(a_i, a_{-i})$ is the conventional material payoff received when agent i chooses action a_i while the remaining agents choose a_{-i}, and
- $u_i(a_i, a_i, \ldots, a_i)$ is the payoff that would result if agent i's action were universalized (i.e., adopted by all agents).

The parameter $\kappa_i \in [0, 1]$ quantifies the degree to which agent i internalizes the normative implications of his action. When $\kappa_i = 0$, the agent's utility is purely self-regarding (the typical "homo economicus"), whereas when $\kappa_i = 1$, the utility depends solely on the universalized outcome. Alger and Weibull refer to such an agent as *homo kantiensis*, whose decision criterion aligns with Kant's categorical imperative—that is, one should act only on maxims that one wishes to become universal laws.

To illustrate the effect of this transformation, when the *homo kantiensis* utility is applied to the game in table 1, each agent's evaluation is determined solely by the outcome associated with the action they choose. In particular, the payoff for action $S1$ is given by the diagonal entry A, and for action $S2$ by D. The transformed payoff matrix is given in Table 2.

	$\mathbf{S}_1$	$\mathbf{S}_2$
$\mathbf{S}_1$	A, A	A, D
$\mathbf{S}_2$	D, A	D, D

Table 2. Deontological Interaction

2.1.2 Average Consequence Utilitarianism

To incorporate utilitarianism into a game, one may augment a player's own payoff by a fraction of the other player's payoff. Let $U_1(x, y)$ and $U_2(x, y)$ denote the material payoffs for Players 1 and 2, respectively. One formulation of altruistic preferences is

$$V(x, y) = \frac{U_1(x, y) + \alpha\, U_2(x, y)}{1 + \alpha}, \tag{1}$$

where $\alpha \geq 0$ represents the degree of altruism. (This utility function is equivalent to that in Levine (1998) upon setting the spite parameter λ to zero and normalizing by $(1 + \alpha)$.) When $\alpha = 1$, this utility function becomes the following:

$$V(x, y) = \frac{U_1(x, y) + U_2(x, y)}{2}, \tag{2}$$

The matrix in Table 3 results from applying this utility function to the matrix in Table 1.

	Player 2	
Player 1	Col 1	Col 2
Row 1	A, A	$\left(\frac{B+C}{2}, \frac{B+C}{2}\right)$
Row 2	$\left(\frac{B+C}{2}, \frac{B+C}{2}\right)$	D, D

Table 3. Utilitarian Interaction

2.1.3 Empathy

An alternative perspective views altruism as a trade-off between self-interest and concern for others. Let $\varepsilon \in [0, 1]$ denote an agent's level of empathy, with higher values reflecting greater concern for the other agent's well-being. Under this specification, the utility function becomes

$$V(x, y) = (1 - \varepsilon)\, U_1(x, y) + \varepsilon\, U_2(x, y). \tag{3}$$

When $\varepsilon = 1$, the agent is entirely other-regarding and their utility function becomes the following.

$$V(x, y) = U_2(x, y). \tag{4}$$

The matrix in Table 4 results from applying this utility function to the matrix in Table 1.

	$\mathbf{S}_1$	$\mathbf{S}_2$
$\mathbf{S}_1$	A, A	B, C
$\mathbf{S}_2$	C, B	D, D

Table 4. Empathetic Interaction

2.1.4 Maxi-Min Preferences via a CES Utility Function

We further generalize our framework by incorporating a maxi-min criterion through a constant elasticity of substitution (CES) utility function. Specifically, consider the

CES aggregator for two outcomes x and y:

$$U(x,y) = \left[\delta\, x^{-\rho} + (1-\delta)\, y^{-\rho}\right]^{-1/\rho},$$

where $\delta \in [0,1]$ is a weight parameter and $\rho \geq 0$ governs the degree of substitution between x and y. In the limit as $\rho \to \infty$, the aggregator converges to the Leontief (min) function:

$$\lim_{\rho \to \infty} U(x,y) = \min\{x,y\}.$$

This extreme case corresponds to a maxi-min utility specification in which the agent evaluates outcomes solely based on the worst-off payoff, reflecting a strict form of egalitarianism. The matrix in Table 5 results from applying this utility function to the matrix in 1.

	$\mathbf{S}_1$	$\mathbf{S}_2$
$\mathbf{S}_1$	A, A	Min[B, C]
$\mathbf{S}_2$	Min[C, B]	D, D

Table 5. Maxi-Min Interaction

2.1.5 Selfishness

Finally, pure selfishness is represented by the classical game–theoretic model in which each agent maximizes their own material payoff without regard for the payoff of others. In this case, the payoff matrix in Table 1 (without transformation) reflects the interaction of two selfish agents.

2.2 Autonomous Moral Choice

Traditional game theory posits that players choose among available strategies and that the combination of these strategies determines their payoffs. In our extended framework, players simultaneously select an action strategy and a moral principle to evaluate outcomes.

Consider a standard normal-form game with N players, where each player i has an action strategy space

$$S_i = \{s_1^i, s_2^i, s_3^i, \ldots\},$$

	Ds_1	Ds_2	As_1	As_2
Ds_1	A,A	A,D	A,A	A,$(B+C)/2$
Ds_2	D,A	D,D	D,$(B+C)/2$	D,D
As_1	A,A	$(B+C)/2$,D	A,A	$(B+C)/2$,$(B+C)/2$
As_2	$(B+C)/2$,A	D,D	$(B+C)/2$,$(B+C)/2$	D,D

Table 6. Autonomy to Choose Between Deontology and Altruism

and a payoff function $u_i(s_1, \dots, s_N)$. We denote the game by

$$g = [N; \{S_1, \dots, S_N\}; \{u_1(\cdot), \dots, u_N(\cdot)\}] .$$

To incorporate moral agency, we introduce a morality space for each player

$$M_i = \{m_1^i, m_2^i, m_3^i, \dots\},$$

which represents the alternative moral principles available. The transformed strategy space for player i is then given by the Cartesian product

$$T_i = M_i \times S_i,$$

so that a player's overall decision involves choosing both a moral principle and an action strategy. Under this transformation, a player's payoffs depend not only on the action strategies selected by all players but also on the moral principles they adopt.

A comprehensive discussion of models incorporating autonomous moral choice can be found in Gouri-Suresh and Studtmann (2025). Here, we illustrate the idea of a synthesis of moralities with a matrix in Table 6, which represents the synthesis of utilitarianism and deontology, a moral rule that plays a crucial role in the ensuing discussion.

3 Simulations

To determine the expected material payoffs for each possible pairing of moral rules, we conducted Monte Carlo simulations. For each pairing, we generated the corresponding payoff matrix (e.g., Table 6 represents two agents following the synthesis of utilitarianism and deontology) and ran 100,000 simulation rounds. In each round, we assigned random values between -1 and 1 to the variables, assuming a uniform distribution to reflect the absence of prior assumptions about their distribution. To prevent degenerate games, we introduced slight perturbations when variables were assigned identical values. For each round, we identified all equilibria and computed the expected material payoff, as represented in Table 1. When multiple equilibria were present, we assumed each was equally likely to occur.

With this description of the models and simulations in place, we can now articulate the precise assumptions underlying our approach. This discussion clarifies both the breadth of our results and their limitations. First, we examine moral rules within the full range of two-player, two-choice *symmetric* interactions. This level of generality is mathematically natural and includes several well-studied games that are highly relevant to morality, such as the Prisoner's Dilemma, Stag Hunt, and Hawk-Dove. The restriction to symmetric games is also motivated by the fact that Kantian deontology is not well defined in non-symmetric contexts. While one might suggest that asymmetric games could be symmetrized in advance, an approach employed, for instance, by Alger and Weibull (citation needed), we find that such symmetrization introduces a consequentialist aspect into the framework, which is at odds with deontological reasoning.

Second, we assume that the variables are uniformly distributed, which we take to be the most natural assumption in the absence of priors about the interactive context. We recognize, however, that this restriction limits the conclusions that can be drawn, and we plan to explore alternative distributions in future work. Lastly, we assume that multiple equilibria are equally likely to occur, a mathematically and analytically well-justified assumption, though one that could be modified in various ways.

Despite these limitations, we believe our results are sufficiently broad and compelling to warrant serious consideration.

4 The Optimific Champion

Parfit (2011) has argued that the correct moral rule is optimific, that is, it produces the best consequences under the assumption that everyone follows it. Our simulations provide a definitive answer regarding which rule, among those we examine, is optimific. In fact, we argue that it will be challenging to find a rule that outperforms the one we identify. However, the simulations reveal much more than just the optimific champion. They offer a number of insights, which we shall discuss systematically. The following is a complete ranking of the expected material outcomes for interactions in which both agents adhere to the same moral rule.

1. **UD = UDM**
2. **DM**
3. **SUD = EUD = SUDM = EUDM**
4. **SD = ED = SDM = EDM**
5. **SDE = SDEM = SUDE = SUDEM**
6. **SUM = EUM**
7. **SU = EU = UM = SEUM**
8. **D = U = SE = SEM = SEU**
9. **E = S = EM = SM**
10. **M**

4.1 Syntheses Outperform Foundational Moralities

One of the most striking insights derived from the simulations is that a synthesis of moral rules tends to outperform the foundational moralities from which it is constructed. This result is of profound significance. While Parfit (2011) has argued for a synthesis of utilitarianism, Kantian deontology, and contractualism, the broader idea that moral rules should be synthesized has been largely neglected by moral philosophers. The traditional focus has been on defending one foundational rule at the expense of others. Our findings demonstrate that this approach is misguided.

At the bottom of the ranking, at level (10), lies the foundational rule of maxi-min (M), which achieves the lowest expected material payoffs. This result is unsurprising; even Rawls (1971), who championed maxi-min in political philosophy, acknowledged that it might not be suitable for many situations. Our simulations confirm this: In symmetric dyadic perfect information interactions, maxi-min performs the worst among the rules we examined.

Above maxi-min, at level (9), there are two other foundational rules: selfishness (S) and empathy (E). Interestingly, these two achieve the same expected material outcomes in interactions. This may seem counterintuitive, but upon closer examination, the reason for the equivalence becomes clear: In cases where both agents act according to the same rule, two purely selfish agents achieve outcomes identical to those of two empathetic agents because in the latter case, the interest of one agent is as effective as their own self-interest in the former case. Moreover, two syntheses, selfishness combined with maxi-min (SM) and empathy combined with maxi-min (EM), are also on level (9). This points to a general characteristic of maxi-min: its transformative effect on outcomes is often minimal when synthesized with other moral rules. As maxi-min focuses exclusively on minimizing the worst outcome, it tends to be dominated by other moral considerations, rendering its impact negligible in most cases. However, there are exceptions to this generalization.

At the next level of the ranking, level (8) we encounter two foundational rules, deontology (D) and utilitarianism (U). Surprisingly, these achieve identical expected material outcomes despite the distinct nature of their respective utility functions. This is one of the more intriguing results from the simulations, one that has been proven analytically and discussed in detail by Gouri-Suresh and Studtmann (2024). Alongside these foundational rules, three syntheses also appear at this level of the ranking: empathy combined with selfishness (SE), empathy combined with selfishness and maxi-min (SEM), and empathy combined with selfishness and utilitarianism (SUE). Of particular interest is the fact that empathy combined with selfishness performs as well as deontology and utilitarianism. This highlights an overlooked synthesis of moral rules that achieves outcomes equivalent to those of the two most celebrated foundational moral theories. The fact that the synthesis of empathy, selfishness, and maxi-min achieves the same expected outcomes is an instance of the general tendency we mentioned above for maxi-min not to make a substantive difference to the outcomes. However, it is interesting that in this case the addition of utilitarianism does not improve outcomes, as can be seen by the fact that the synthesis of selfishness, empathy, and utilitarianism also achieves outcomes equal to the synthesis without utilitarianism.

Now, the very next level of the ranking, level (7), and all the levels above it contain only syntheses of moralities. This is a demonstration of the power of synthesis or, in this context, equivalently, the power of moral choice. Game theory has long demonstrated that strategic choices can improve agents' outcomes. As it turns out, what can be called moral choices – that is, the choice between competing moral frameworks – significantly improves outcomes as well.

Historically, Kant emphasized autonomy as a choice between the law of selfishness and the moral law derived from the categorical imperative. Kant's emphasis on autonomy emerged from an *a priori* analysis, largely divorced from considerations of consequences. In contrast, our simulations focus on the consequences of autonomous decision-making. Nonetheless, we reach a parallel conclusion: autonomy is essential for achieving superior material outcomes. However, our findings extend Kant's conclusion, showing that various forms of autonomy - beyond the autonomy to choose between the law of selfishness and the universal law - can improve outcomes in a wide range of interactions. This conclusion should not come as a surprise. If strategic decision-making improves expected outcomes, then allowing agents an additional layer of choice, one between moralities, naturally further optimizes these outcomes.

4.2 Deontology is Synergistic

The history of moral philosophy is marked by well-trodden debates in which moral theories are treated as independent entities, their consequences analyzed and compared both to one another and to our moral intuitions. Although this traditional approach has its merits, it precludes an exploration of the properties moral theories might exhibit when synthesized with others. This is unfortunate since it may turn out that certain foundational moral theories prove their worth in syntheses, whereas other moral theories do not. The above rankings reveal that this is indeed the case. For they show that Kantian deontology possesses a unique synergistic quality: When combined with other moral frameworks, the expected payoffs of the resulting synthesis exceed those of the constituent moralities considered in isolation. The same is not true of the other moral theories.

The synergistic potential of Kantian deontology is abundantly illustrated in its synthesis with maxi-min. As noted earlier, maxi-min, when considered alone, produces the lowest expected payoffs among the moralities under consideration. However, the synthesis of maxi-min and deontology (MD) achieves the second-highest expected payoffs. This represents a dramatic leap from maxi-min's rank at level (10) and deontology's at level (8) to a shared position at level (2). Although this is the most striking example of synergy, every synthesis involving deontology exhibits this phenomenon. This synergistic effect is not confined to syntheses with only two moralities but extends to more complex syntheses as well. For example, the synthesis of deontology, utilitarianism, and selfishness (SUD), yields higher expected outcomes than not only its individual components - deontology (D), selfishness (S) and utilitarianism (U) - but also the simpler synthesis (SU), which excludes deontology.

One might be tempted to attribute these effects to the introduction of autonomy rather than to deontology itself. However, the rankings clearly demonstrate that this is not the case. Consider, for instance, the synthesis of utilitarianism, selfishness, and empathy (SUE), which achieves the same expected outcome as utilitarianism alone. This indicates that the synthesis does not exceed the performance of its components. One might wonder what accounts for this synergistic effect. Although we do not have a formal answer to this question, it is plausible that the synergistic properties of deontology stem from its emphasis on universal rule-following. Deontological frameworks grounded in universalizing determine moral actions through the outcomes of rules that everyone adheres to. This intrinsic universality likely underpins the synergistic effects observed in the simulations.

That rule following correlates with higher expected outcomes is, on reflection, entirely intuitive. Human activity, unlike that of other animals, is deeply shaped by the adherence to rules. Language, economic exchange, courtship, and virtually every aspect of human life is governed by rules. It is thus reasonable to speculate that the exceptional cooperative nature of human societies can be traced to our capacity for rule-following. The simulation results lend strong support to this hypothesis, suggesting that deontology - particularly in the form of universal rule following - occupies a central position in the structure of morality, which, as many authors have argued, facilitates human cooperation.

The centrality of deontology in moral frameworks is further underscored by its presence in all the highest-ranked syntheses (levels 1–4). In contrast, no other moral theory enjoys such a consistent representation. Although utilitarianism features in some high-performing syntheses, such as the pairing of utilitarianism and deontology (UD), it also appears in several non-synergistic frameworks at lower levels (6–8). This is not to downplay the importance of utilitarianism; indeed, it forms part of the highest-performing synthesis. Rather, it highlights that utilitarianism, without the structural benefits conferred by deontology, lacks the capacity to elevate foundational moral frameworks to the highest levels of performance.

4.3 UD is the Optimific Champion

We now turn to the first major goal of our inquiry: identifying the optimific champion. As the rankings indicate, there are two contenders for this title: (1) the synthesis of utilitarianism and deontology (UD) and (2) the synthesis of utilitarianism, deontology and maxi-min (UDM). However, as previously noted, the inclusion of maxi-min often contributes little to the effectiveness of a synthesis, frequently leaving the outcomes unchanged. This is the case here: adding maxi-min to the synthesis does not improve the expected outcomes. Thus, we can declare the synthesis of utilitarianism and deontology (UD) as the true optimific champion.

The status of UD as the optimific champion underscores the fundamental importance of utilitarianism and deontology as moral frameworks. However, one might wonder whether it is possible to construct a synthesis that outperforms UD. Such a moral rule would be of interest if for no other reason than that it would have to be composed at least in part by a foundational morality distinct from either deontology or utilitarianism. And it would be of interest to discover such a foundational morality. However, although such a synthesis may exist, it is not immediately apparent what form it would take. The payoff matrix for UD shown in Table 6, considered along with the objective material payoffs of the original game, demonstrates that UD comes very close to achieving the theoretical maximum expected payoff and that it will be difficult to do better.

The near-optimality of UD becomes evident when we compare the original symmetric form in Table 1 with the matrix in Table 6. In the original game, the total expected outcome can take one of three values: $2A$, $2D$, or $B+C$, yielding six possible orderings. Since the orderings for cases where $A > D$ are symmetrical to those where $D > A$, it suffices to consider three scenarios:

1. $A > D > (B + C)$,
2. $A > \frac{B+C}{2} > D$,
3. $\frac{B+C}{2} > A > D$.

In scenarios (1) and (2), the optimal outcome occurs when both agents coordinate on the action yielding R. In scenario (3), the best outcome arises when the agents anti-coordinate.

Examining the matrix in Table 6 reveals that the individual payoffs correspond to A, D, or $\frac{B+C}{2}$. Thus, the solution space for the game mirrors the ordering of total expected outcomes in the original symmetric game. In cases (1) and (2), agents playing UD coordinate on the strategy that yields R, which is the maximum payoff under those conditions. In case (3), the game transforms into an anti-coordination game featuring three equilibria: two pure-strategy asymmetric equilibria and one symmetric mixed-strategy equilibrium. In scenario (3), the asymmetric equilibria yield the highest total payoffs; hence, if the agents anti-coordinate, they achieve the maximum possible payoff. However, the existence of the mixed-strategy equilibrium prevents UD from reaching that maximum, since its expected payoff is lower than that of the pure strategies.

This limitation, the mixed strategy equilibrium, poses a challenge in devising a synthesis that outperforms UD. Any such synthesis would need to offer a mixed equilibrium superior to that of UD while maintaining the same partitioning of the solution space. Given that the mixed strategy for UD arises directly from its payoff structure, modifying the structure to enhance the mixed equilibrium would misalign the solution space from the ordering of the maximum possible material payoffs derived from the original symmetric form. Although such a synthesis might be attainable, its precise formulation remains an open mathematical question.

A potential resolution to this limitation is to impose a restriction whereby agents choose only pure-strategy equilibria. This hypothetical rule, which we call *Pure Strategy Deontarianism*, is admittedly artificial because it mathematically enforces a constraint on agent behavior. Nonetheless, it guarantees that agents achieve the maximum possible expected payoff. This feature renders Pure Strategy Deontarianism a valuable theoretical benchmark, serving as an absolute measure of optimality against which all other moral rules can be evaluated. The table below summarizes the percentages of the maximum possible payoff achieved at each level of the ranking.

Rank Level	**Percentage of Maximum Payoff**
1	0.945
2	0.902
3	0.889
4	0.860
5	0.843
6	0.745
7	0.736
8	0.727
9	0.605
10	0.496

On a standard grading scale, moralities at levels (6) to (10), which include all the foundational moralities, as well as all the syntheses that contain utilitarianism but not deontology, perform poorly, earning grades ranging from C to F. In contrast, syntheses that include deontology, all of which occur on levels (1) to (5), achieve grades between

A and B, further highlighting the strength of frameworks that incorporate deontology.

5 The Realist's Champion

If determining the correct moral rule required only an examination of cases in which all agents act according to the same rule, the results of the previous section would provide substantial insight, identifying a moral rule – UD – which, if not the correct one, at least approximates it. However, restricting the analysis to such cases would overlook a critical dimension of morality: its capacity to foster cooperation by combating selfishness. A widely accepted view among empirically oriented researchers of morality is that one of the core functions of morality - if not its defining function - is to counteract selfishness. (Curry et al. (2019), Krebs (2011), Smyth (2017), de Waal (2014), Schroder (2000)) Therefore, it is imperative to examine to what extent a given morality, whether foundational or synthetic, serves this purpose.

In general, there are two ways that a moral rule can be said to combat selfishness. First, it may incline its adherents toward prosocial rather than self-serving behavior. Second, a moral rule can be assessed by its capacity to coexist with selfishness in a mixed society. Suppose that we envision a population composed of agents adhering to the rule of selfishness and others following a competing moral rule X. If the population reaches an equilibrium where no agent has an incentive to switch rules, the proportion of agents adhering to X serves as a measure of the rule's efficacy against selfishness. A rule that can only sustain 10% adherence in such an equilibrium clearly combats selfishness less effectively than a rule that commands 60%.

In this analysis, we focus on the latter measure, the percentage of a mixed population in equilibrium that adheres to a moral rule. As we shall see, one result that we obtain is that these two different ways of measuring the effectiveness of a moral rule against selfishness come apart. The rules that combat it best in the second sense are not the best at combating it in the first sense.

To begin, note an immediate result concerning how each moral rule performs against selfishness in pairwise interactions:: selfishness always outperforms any moral rule in direct competition. If there were to be a 1 on 1 champion in a 'battle royale' of moral rules, selfishness would emerge victorious. The strength of selfishness in 1 on 1 interactions, however, ends up being the source of a weakness, one that, as we shall see, opens up the possibility that a moral rule, even though it loses 1 on 1 to selfishness, can nonetheless emerge victorious in a contest with it: When a selfish agent interacts with another selfish agent, the expected outcomes are low. As the table in the previous section shows, selfish agents interacting with other selfish agents achieve only 60% of the maximum possible payoffs. When turned against itself, the ferocity of selfishness leads to expected outcomes that earn a grade of D-.

Selfishness's poor outcomes when interacting with itself opens the door for alternative moralities to exist in a society with it. Although a selfish agent consistently outperforms any agent adhering to a pro-social rule in direct, 1 on 1 interactions, it remains possible for that rule to outperform selfishness when measured against the outcomes selfishness achieves when interacting with itself. This turns out to be critical. In a mixed society, selfish agents do not exclusively engage with agents of other moral persuasions; they must also contend with interactions among themselves. The same holds true for agents following a moral rule: they interact both with selfish agents and with others adhering to their own rule.

As a result, if a moral rule performs better against selfishness than selfishness does against itself - and also achieves higher outcomes in interactions among its own adherents than selfishness achieves against selfishness - it can establish a stable presence within a mixed society even though selfishness defeats it 1 on 1. To put the matter colloquially, although every moral rule is bound to lose to selfishness 1 on 1, some of them hold their own against it by achieving higher payoffs than selfish agents achieve when interacting with selfish agents. The ability of moral agents to hold their own against selfish agents, combined with their superior outcomes in intragroup interactions, enables them to persist in equilibrium with selfishness.

Now, to properly evaluate the extent to which a moral rule can hold its own against selfishness, it is necessary to consider the nature of the matching process within a mixed society. One might assume assortative matching, where agents following a moral rule are more likely to interact with others who follow the same rule. Although this assumption is common in many studies, it presents a problem in assessing the true resilience of a moral rule against selfishness. Assortative matching artificially increases the performance of moral rules by creating favorable interaction conditions, as moral agents fare better when paired with others of their kind. To avoid this bias, we assume a random matching process in which agents interact without preference for adherence to rules. Only under such conditions can the true strength of a moral rule against selfishness be accurately assessed.

Under random matching, only eight of the 31 moral rules analyzed outperform selfishness against itself. Just as in the case of agents interacting with other agents of the same type, when agents of a type, X, successfully interact against selfishness, so too will agents that are a synthesis of type X and maxi-min. There are eight moral rules that do better against selfishness than selfishness does against itself: D, DM, SD, SDM, UD, UDM, SUD, SUDM. These eight rules are of significant interest, since they are the only rules of those we have examined that can successfully perform the function of combating selfishness. Were one to suppose that any viable moral rule must be able to perform this function, then we have discovered a list of viable moral rules.

Although each of these eight rules is viable, they are not equally effective in combating selfishness. From the simulations, we can determine the proportion of agents in a mixed population who adhere to each of these rules in equilibrium, as shown in the following table.

Strategy	**Percentage**
D	25%
DM	37%
UD	37%
UDM	37%
SUD	44%
SUDM	44%
SD	60%
SDM	60%

As in the case with the optimific champion, this ranking allows us to pronounce the realist's champion. However, it also reveals many interesting insights, which we discuss below.

5.1 Deontology is Essential to Morality

In the history of moral philosophy, theories have often been treated as adversaries, each vying for dominance in a zero-sum contest. Within such a framework, the notion of a moral rule being essential to morality makes little sense; a rule is either correct or it is not. However, this dynamic shifts when one considers moral syntheses, particularly within mixed societies. Here, only a subset of moralities proves viable, and it becomes an open question whether a foundational moral rule might underlie every viable synthesis. If such a rule exists, it would be essential to morality.

Our analysis reveals that deontology is the only foundational moral rule present in every viable synthesis. Moreover, deontology is viable on its own: in a mixed society of selfish agents and purely deontological agents, the latter comprise 25% of the population in equilibrium. Yet deontology's true strength lies in its synergistic properties. Not only does it feature in every viable synthesis, but these syntheses outperform the constituent moralities from which they are constructed.

This centrality of deontology is surprising. Moral philosophy has long depicted deontology and utilitarianism as equals in a philosophical deadlock, with neither side able to decisively triumph. However, we contend that such a depiction is misleading. By focusing on moral syntheses in mixed societies, it becomes evident that deontology occupies a fundamental role that utilitarianism does not. This is not to dismiss the relevance of utilitarianism to morality; there are viable syntheses that include it. However, utilitarianism should not be regarded as an equal to deontology.

5.2 Maxi-Min: An Unexpected Equal to Utilitarianism

Maxi-min, as a decision-making rule, has relatively few advocates. Rawls famously incorporated it into political philosophy, yet even he acknowledged its limitations outside specific contexts. In contrast, utilitarianism has enjoyed widespread and enduring support. To some extent, our simulations support this difference in status. When examined independently in the optimific case, utilitarianism does outperform maxi-min in our simulations: utilitarianism ranks two levels higher than maxi-min, which sits at the very bottom.

However, a different picture emerges when one considers syntheses. Two striking facts stand out. First, as we have already seen, the synthesis of deontology and utilitarianism (SD) and the synthesis of deontology and maxi-min (DM) occupy the top two levels in the optimific ranking. This shows that in the optimific case, the two are almost as good as one another, though there is admittedly a slight advantage to utilitarianism. But in mixed societies, these two syntheses achieve near-identical equilibrium adherence rates, each commanding approximately 37% of the population. Although this result has not been proven analytically, an exact identity is within the margin of error. Even if there is no exact identity, the proximity of these outcomes underscores their comparable performance.

Importantly, however, maxi-min surpasses utilitarianism in mixed societies in one key respect. Among viable moral rules, we distinguish between weak viability, the mere ability to persist in equilibrium alongside selfishness, and strong viability, which requires moral agents to outnumber selfish agents in equilibrium. Only NK and NKR achieve strong viability, satisfying what we term the "non-Quixote condition." This ensures that morality is not relegated to a noble but embattled minority, perpetually overshadowed by selfishness. Maxi-min's inclusion in NKR grants it a role in a strongly viable synthesis, a distinction utilitarianism lacks.

Thus, while utilitarianism may appeal intuitively, it falls short in its ability to be part of a non-quixotic morality. Utilitarianism is relegated to tilting at windmills. Maxi-min, by contrast, despite its modest reputation, is a component of a strongly viable, and hence non-quixotic, moral synthesis. Now, one might object that in the synthesis SDM maxi-min is not really doing any work. It is really SD that is the strongly viable morality and maxi-min is riding its coattails. In response, we note that adding utilitarianism to SD lowers the expected proportion of adherents in equilibrium. In equilibirum, SDU is only weakly viable. So, maxi-min, though it does not add anything to SD, at least has the good sense not to interfere with a good thing. Utilitarianism does not.

5.3 The Best Way to Combat Selfishness: Allowing the Choice to Be Selfish

Two criteria measure the ability of a moral rule to combat selfishness: (1) its capacity to induce prosocial behavior; and (2) its capacity to maintain adherence in mixed societies at equilibrium. These criteria can conflict. Our simulations reveal that moral rules that incorporate selfishness as a choice perform better under the second criterion, even if it comes at the expense of prosocial behavior.

Consider deontology. Purely deontological agents always act to maximize group outcomes, consistently cooperating, for example, in the Prisoner's Dilemma. In contrast, agents who adhere to a synthesis of deontology and selfishness choose strategically, sometimes cooperating, and sometimes defecting. Although the synthesis of deontology and selfishness is thus less pro-social than deontology, the ability to choose to be selfish allows the synthesis to outperform pure deontology in mixed societies.

This insight challenges Kant's categorical imperative, which demands one always adhere to universalizable rules. If success in mixed societies, measured by equilibrium adherence, is the criterion for moral correctness, then the synthesis of deontology and selfishness emerges as superior to pure deontology. Hence, it is false that one should always act according to a universalizable rule. Kant might object that this conclusion rests on consequentialist considerations alien to pure morality. However, this synthesis requires agents to navigate the tension between selfishness and the moral law, a decision Kant himself emphasized as central to moral autonomy. In this sense, the synthesis aligns with Kantian principles while accommodating consequentialist realities.

5.4 SD is the Realist's Champion

The idea that humans face a fundamental conflict, one between their own selfish desires and the demands of the moral law, has been a central theme in philosophy since Plato. It is therefore striking that the synthesis of selfishness and deontology (SD) emerges as the realist's champion. (We assume that SD rather than SDR is the realist's champio despite their tying in mixed societies for reasons that we have already discussed.) Agents adhering to SD must navigate a choice between actions grounded in universalizable principles and actions motivated purely by self-interest. This synthesis is particularly notable because deontology, as defined within game theory, imposes a symmetry constraint on the reasoning of an agent. A deontological agent selects optimal symmetrical outcomes, while a selfish agent considers all outcomes but prioritizes those maximizing their own material welfare. SD thus compels agents to choose between viewing the world through the lens of symmetry or through the lens of asymmetry. This tension embodies an axiological struggle: If symmetry represents the good and asymmetry the bad, then NK obliges agents to choose between the good and the bad; or if one does not wish to view selfishness as the bad, it obliges a choice between the good and the selfish. In this framework, a strategic decision transcends mere material outcomes, becoming in addition an axiological choice.

Several philosophers have criticized evolutionary and game-theoretic accounts of morality for neglecting crucial dimensions, particularly autonomy. (Buchanan and Powell (2018), Korsgaard (2006) and Nagel (1979)). What is perhaps most surprising about the results of this study is that the agents who outperform selfishness, and thus the agents most evolutionarily successful against selfish counterparts, are autonomous. This finding simultaneously acknowledges the limitations of standard game-theoretic and evolutionary accounts and yet integrates a deeper understanding of moral agency into game theory.

Beyond integrating moral autonomy into game theory, these results highlight an aspect of morality often overlooked by philosophers who approach human existence through a moral lens. Under random matching, SD achieves equilibrium adherence rates of 60%, increasing to 70% with assortative matching. These figures exceed those of all other moralities we examined. Despite SD's impressive performance, 30–40% of the population in a mixed society will still adhere to the rule of selfishness. Such agents lack intrinsic motivation to act morally; their seemingly moral actions are driven solely by external incentives, such as fear of punishment or social ostracism. Absent these motivations, they consistently defect. Hence, the attempt by some philosophers to view humans under a single lens, as all having the capacity to be motivated by moral considerations, may be misguided. It may be that human agency is inherently dual, with some agents who have an internal motivation to act morally and others who do not.

6 Conclusions

Our investigation has explored a battle royale of moral theories through extensive simulations, identifying two champions: the optimific champion — the synthesis of utilitarianism and deontology (UD) — and the realist's champion — the synthesis of deontology and selfishness (SD). The optimific champion, UD, is a rule that produces the highest material outcomes under the assumption of universal adherence. In contrast, the realist's champion, SD, demonstrates a much more robust form of moral success: It can not only coexist with selfishness in equilibrium but can attract the majority of agents in mixed populations.

But which champion should we consider the correct moral rule? This question demands metaethical considerations that extend beyond the scope of this paper. Nonetheless, we offer some observations that merit consideration.

6.1 Optimific Rules

Optimific rules, we contend, are ideally suited for two kinds of situations. First, if agents know that they will all adhere to the same rule, they have reason to choose an optimific rule. This scenario might describe a utopian society—such as a notion of heaven—in which, for example, angels interact according to the pure-strategy version of UD or some equivalent rule. Although the possibility of such a society may be questioned, there is another context where possibility is not at issue: that of two hypothetical, idealized agents engaged in the derivation of the social contract. Incorporating an optimific rule into this process guarantees that the resulting theory of justice aligns with a rule that attains the maximum possible payoffs across a morally significant set of interactions, arguably representing an ideal form of justice.

Of course, neither heaven nor the idealized conditions for deriving a theory of justice exist in the actual world. Thus, the question remains whether an optimific rule should guide real-world situations, where agents may not act morally. Several philosophers have explored this issue at length (citations). Here, we simply note that if one regards morality as inherently utopian—meaning that one ought always to look toward the ideal even in non-ideal circumstances—then the optimific rule stands as an appropriate guide in an imperfect world. On the other hand, if one believes that moral rules should be sensitive to real-world imperfections, it might be more advisable to adopt the rule that performs best in the presence of selfish agents.

6.2 The Importance of Viability

The notion of viability, the capacity of a moral rule to sustain itself against selfishness, is an inherently evolutionary concept. If one makes the minimal assumption that a mixed society will evolve until it reaches its equilibrium, a viable rule is one that can evolve from a monomorphic society of selfish agents. If a mutation within such a society occurs to an agent acting according to a viable morality, then the society will evolve to an equilibrium that contains some non-zero percentage of non-selfish agents. The eight rules we identified as viable, with SD emerging as the most robust, can evolve in this way. And they can do so, remarkably, without any of the standard explanatory mechanisms for the evolution of cooperative behavior, such as kin selection, indirect reciprocity, punishment, group selection, etc. They can evolve without any such mechanisms because they all share the property that they achieve better outcomes against selfishness than selfishness achieves against itself.

One may object that viability, understood in this evolutionary sense, is not a necessary criterion for a moral rule. It is conceivable after all that morality is fundamentally non-biological since it is applicable in contexts untouched by evolutionary pressures, such as a hypothetical heavenly society. However, even if one rejects viability as a necessary condition for morality, the concept remains of interest. Rules that can persist and thrive against selfishness in evolutionary contests are, by their robustness alone, deserving of careful philosophical consideration.

If, however, one accepts that evolutionary robustness is indispensable to morality, i.e., that a moral rule *must* survive in competition with selfishness, then viability becomes a central moral concept. And it provides a straightforward test for proposed moral rules: can they co-exist with selfish agents in equilibrium under random matching? If not, then at best they articulate ideals suited to contexts insulated from evolutionary pressures, contexts far removed from the biological and social realities of human life. For such settings, the optimific rule, UD, may suffice. But in this world, a viable moral rule is required and SD emerges as the strongest contender.

6.3 Final Reflection

The contest between moral rules staged in this paper mirrors, in some ways, the history of moral philosophy itself, a history defined by the tension between different moral principles and between moral principles and self-interest. Our findings offer resolutions to these tensions, resolutions that are made possible by allowing for the synthesis of moralities. UD as the optimific rule demonstrates that utilitarianism and deontology, two different moral principles, can, when synthesized, yield near optimal outcomes under conditions of universal adherence. SD, as the realist's champion, reveals that deontology, when synthesized with selfishness, thrives even in mixed societies.

Appendix

In this appendix, we first provide the equations for the analytic computation of the expected value for two agents acting according to Pure Strategy Deontarianism. We then provide a table displaying the expected material payoffs for pairwise interactions between all agent types. In the table, each entry consists of two numbers: the first corresponds to the row agent and the second to the column agent.

A Maximum Total Expected Payoffs

When two players have the autonomy to choose between altruism and deontology but are restricted to pure strategies, their total expected payoff is at a maximum. Equation 5 shows the expression for Z_{MTEP}, the expected value of the average material welfare for the two players following this moral rule. Note that $y(A, B, C, D)$ is the average material welfare for given values of A, B, C, and D. We computed the value of Z_{MTEP} analytically and confirmed it numerically.

$$\begin{aligned}
Z_{MTEPAvg} &= \int_{A=-1}^{1}\int_{B=-1}^{1}\int_{C=-1}^{1}\int_{D=-1}^{1} y(A,B,C,D)\,dD\,dC\,dB\,dA \\
X &= \frac{B+C}{2}, \\
y &= \begin{cases} A, & \text{if } A \geq X \geq D, \\ A, & \text{if } A \geq D > X, \\ \frac{1}{2}(B+C), & \text{if } X > A \geq D, \\ \frac{1}{2}(B+C), & \text{if } X > D > A, \\ D, & \text{if } D > A \geq X, \\ D, & \text{if } D > X > A, \end{cases} \\
Z_{MTEP} &= \frac{11}{24} = 0.458
\end{aligned} \tag{5}$$

B Table for Pairwise Interactions

Table 7. Table of Expected Material Outcomes in Pairwise Interactions

	N	**U**	**K**	**E**	**R**	**NU**
N	(27.79, 27.79)	(41.98, 25.23)	(49.99, 33.29)	(48.94, 15.51)	(31.70, 24.27)	(40.22, 27.23)
U	(25.23, 41.98)	(33.33, 33.33)	(45.79, 45.86)	(41.96, 25.28)	(33.38, 33.40)	(29.23, 37.23)
K	(33.29, 49.99)	(45.86, 45.79)	(33.31, 33.31)	(50.04, 33.31)	(43.34, 43.32)	(37.48, 48.30)
E	(15.51, 48.94)	(25.28, 41.96)	(33.31, 50.04)	(27.78, 27.78)	(24.27, 31.67)	(17.19, 46.79)
R	(24.27, 31.70)	(33.40, 33.38)	(43.32, 43.34)	(31.67, 24.27)	(22.68, 22.68)	(31.14, 34.74)
NU	(27.23, 40.22)	(37.23, 29.23)	(48.30, 37.48)	(46.79, 17.19)	(34.74, 31.14)	(33.76, 33.76)
NK	(34.31, 43.87)	(47.62, 34.13)	(49.99, 33.29)	(54.47, 21.99)	(43.47, 35.29)	(42.25, 39.08)
NE	(19.04, 45.03)	(32.20, 32.13)	(39.98, 39.98)	(45.05, 19.02)	(32.24, 32.20)	(26.14, 38.26)
NR	(27.79, 27.79)	(41.98, 25.23)	(49.99, 33.29)	(48.94, 15.51)	(31.70, 24.27)	(40.45, 27.50)
UK	(31.58, 49.83)	(43.27, 43.27)	(45.79, 45.86)	(49.82, 31.64)	(43.32, 43.34)	(35.38, 45.28)
UE	(17.19, 46.79)	(29.28, 37.22)	(37.48, 48.36)	(40.22, 27.26)	(31.15, 34.75)	(21.72, 42.06)
UR	(25.23, 41.98)	(33.33, 33.33)	(45.79, 45.86)	(41.96, 25.28)	(33.38, 33.40)	(29.54, 37.49)
KE	(21.99, 54.47)	(34.22, 47.58)	(33.31, 50.04)	(43.91, 34.33)	(35.34, 43.46)	(24.93, 50.80)
KR	(31.52, 47.78)	(43.86, 43.82)	(43.32, 43.34)	(47.80, 31.57)	(41.28, 41.28)	(35.58, 46.24)
ER	(15.51, 48.94)	(25.28, 41.96)	(33.31, 50.04)	(27.78, 27.78)	(24.27, 31.67)	(17.48, 47.00)
NUK	(33.01, 48.07)	(44.53, 36.64)	(48.30, 37.48)	(52.32, 23.53)	(42.53, 38.47)	(39.35, 41.56)
NUE	(19.04, 45.03)	(32.20, 32.13)	(39.98, 39.98)	(45.05, 19.02)	(32.24, 32.20)	(26.25, 38.36)
NUR	(27.23, 40.22)	(37.23, 29.23)	(48.30, 37.48)	(46.79, 17.19)	(34.74, 31.14)	(33.89, 33.89)
NKE	(25.05, 50.57)	(37.97, 37.90)	(39.98, 39.98)	(50.58, 25.03)	(38.03, 37.99)	(31.80, 43.63)
NKR	(34.31, 43.87)	(47.62, 34.13)	(49.99, 33.29)	(54.47, 21.99)	(43.47, 35.29)	(42.25, 39.08)
NER	(19.04, 45.03)	(32.20, 32.13)	(39.98, 39.98)	(45.05, 19.02)	(32.24, 32.20)	(26.25, 38.37)
UKE	(23.52, 52.33)	(36.70, 44.52)	(37.48, 48.36)	(48.07, 33.05)	(38.49, 42.54)	(27.66, 47.40)
UKR	(31.58, 49.83)	(43.27, 43.27)	(45.79, 45.86)	(49.82, 31.64)	(43.32, 43.34)	(35.38, 45.28)
UER	(17.19, 46.79)	(29.28, 37.22)	(37.48, 48.36)	(40.22, 27.26)	(31.15, 34.75)	(21.85, 42.17)
KER	(21.99, 54.47)	(34.22, 47.58)	(33.31, 50.04)	(43.91, 34.33)	(35.34, 43.46)	(24.93, 50.80)
NUKE	(25.05, 50.57)	(37.97, 37.90)	(39.98, 39.98)	(50.58, 25.03)	(38.03, 37.99)	(31.80, 43.63)
NUKR	(33.01, 48.07)	(44.53, 36.64)	(48.30, 37.48)	(52.32, 23.53)	(42.53, 38.47)	(39.35, 41.56)
NKER	(25.05, 50.57)	(37.97, 37.90)	(39.98, 39.98)	(50.58, 25.03)	(38.03, 37.99)	(31.80, 43.63)
NUER	(19.04, 45.03)	(32.20, 32.13)	(39.98, 39.98)	(45.05, 19.02)	(32.24, 32.20)	(26.33, 38.44)
UKER	(23.52, 52.33)	(36.70, 44.52)	(37.48, 48.36)	(48.07, 33.05)	(38.49, 42.54)	(27.66, 47.40)
NUKER	(25.05, 50.57)	(37.97, 37.90)	(39.98, 39.98)	(50.58, 25.03)	(38.03, 37.99)	(31.80, 43.63)

	NK	**NE**	**NR**	**UK**	**UE**	**UR**
N	(43.87, 34.31)	(45.03, 19.04)	(27.79, 27.79)	(49.83, 31.58)	(46.79, 17.19)	(41.98, 25.23)
U	(34.13, 47.62)	(32.13, 32.20)	(25.23, 41.98)	(43.27, 43.27)	(37.22, 29.28)	(33.33, 33.33)
K	(33.29, 49.99)	(39.98, 39.98)	(33.29, 49.99)	(45.86, 45.79)	(48.36, 37.48)	(45.86, 45.79)
E	(21.99, 54.47)	(19.02, 45.05)	(15.51, 48.94)	(31.64, 49.82)	(27.26, 40.22)	(25.28, 41.96)
R	(35.29, 43.47)	(32.20, 32.24)	(24.27, 31.70)	(43.34, 43.32)	(34.75, 31.15)	(33.40, 33.38)
NU	(39.08, 42.25)	(38.26, 26.14)	(27.50, 40.45)	(45.28, 35.38)	(42.06, 21.72)	(37.49, 29.54)
NK	(39.44, 39.44)	(45.87, 30.56)	(34.31, 43.87)	(48.37, 32.87)	(50.82, 24.90)	(47.62, 34.13)
NE	(30.56, 45.87)	(33.14, 33.14)	(19.30, 45.25)	(37.46, 37.39)	(38.29, 26.10)	(32.43, 32.36)
NR	(43.87, 34.31)	(45.25, 19.30)	(27.88, 27.88)	(49.83, 31.58)	(47.00, 17.48)	(42.21, 25.56)
UK	(32.87, 48.37)	(37.39, 37.46)	(31.58, 49.83)	(43.27, 43.27)	(45.28, 35.44)	(43.27, 43.27)
UE	(24.90, 50.82)	(26.10, 38.29)	(17.48, 47.00)	(35.44, 45.28)	(33.78, 33.78)	(29.59, 37.47)
UR	(34.13, 47.62)	(32.36, 32.43)	(25.56, 42.21)	(43.27, 43.27)	(37.47, 29.59)	(33.69, 33.69)
KE	(21.99, 54.47)	(30.55, 45.87)	(21.99, 54.47)	(32.95, 48.33)	(39.12, 42.25)	(34.22, 47.58)
KR	(31.52, 47.78)	(38.00, 38.02)	(31.52, 47.78)	(43.86, 43.82)	(46.28, 35.60)	(43.86, 43.82)
ER	(21.99, 54.47)	(19.28, 45.26)	(15.91, 48.95)	(31.64, 49.82)	(27.53, 40.46)	(25.61, 42.20)
NUK	(38.21, 43.60)	(42.68, 33.23)	(33.01, 48.07)	(45.28, 35.38)	(47.40, 27.66)	(44.53, 36.64)
NUE	(30.56, 45.87)	(33.25, 33.25)	(19.42, 45.34)	(37.46, 37.39)	(38.39, 26.22)	(32.55, 32.48)
NUR	(39.08, 42.25)	(38.37, 26.25)	(27.63, 40.56)	(45.28, 35.38)	(42.17, 21.86)	(37.61, 29.69)
NKE	(30.56, 45.87)	(38.62, 38.62)	(25.05, 50.57)	(37.46, 37.39)	(43.66, 31.77)	(37.97, 37.90)
NKR	(39.44, 39.44)	(45.87, 30.56)	(34.25, 43.81)	(48.37, 32.87)	(50.82, 24.90)	(47.62, 34.13)
NER	(30.56, 45.87)	(33.25, 33.25)	(19.42, 45.35)	(37.46, 37.39)	(38.39, 26.22)	(32.54, 32.47)
UKE	(24.90, 50.82)	(33.19, 42.71)	(23.52, 52.33)	(35.44, 45.28)	(41.58, 39.37)	(36.70, 44.52)
UKR	(32.87, 48.37)	(37.39, 37.46)	(31.58, 49.83)	(43.27, 43.27)	(45.28, 35.44)	(43.27, 43.27)
UER	(24.90, 50.82)	(26.22, 38.39)	(17.61, 47.11)	(35.44, 45.28)	(33.90, 33.90)	(29.74, 37.60)
KER	(21.99, 54.47)	(30.55, 45.87)	(22.14, 54.31)	(32.95, 48.33)	(39.12, 42.25)	(34.22, 47.58)
NUKE	(30.56, 45.87)	(38.62, 38.62)	(25.05, 50.57)	(37.46, 37.39)	(43.66, 31.77)	(37.97, 37.90)
NUKR	(38.21, 43.60)	(42.68, 33.23)	(33.01, 48.07)	(45.28, 35.38)	(47.40, 27.66)	(44.53, 36.64)
NKER	(30.56, 45.87)	(38.62, 38.62)	(25.05, 50.57)	(37.46, 37.39)	(43.66, 31.77)	(37.97, 37.90)
NUER	(30.56, 45.87)	(33.33, 33.33)	(19.50, 45.42)	(37.46, 37.39)	(38.46, 26.30)	(32.62, 32.55)
UKER	(24.90, 50.82)	(33.19, 42.71)	(23.52, 52.33)	(35.44, 45.28)	(41.58, 39.37)	(36.70, 44.52)
NUKER	(30.56, 45.87)	(38.62, 38.62)	(25.05, 50.57)	(37.46, 37.39)	(43.66, 31.77)	(37.97, 37.90)

	KE	**KR**	**ER**	**NUK**	**NUE**	**NUR**
N	(54.47, 21.99)	(47.78, 31.52)	(48.94, 15.51)	(48.07, 33.01)	(45.03, 19.04)	(40.22, 27.23)
U	(47.58, 34.22)	(43.82, 43.86)	(41.96, 25.28)	(36.64, 44.53)	(32.13, 32.20)	(29.23, 37.23)
K	(50.04, 33.31)	(43.34, 43.32)	(50.04, 33.31)	(37.48, 48.30)	(39.98, 39.98)	(37.48, 48.30)
E	(34.33, 43.91)	(31.57, 47.80)	(27.78, 27.78)	(23.53, 52.32)	(19.02, 45.05)	(17.19, 46.79)
R	(43.46, 35.34)	(41.28, 41.28)	(31.67, 24.27)	(38.47, 42.53)	(32.20, 32.24)	(31.14, 34.74)
NU	(50.80, 24.93)	(46.24, 35.58)	(47.00, 17.48)	(41.56, 39.35)	(38.36, 26.25)	(33.89, 33.89)
NK	(54.47, 21.99)	(47.78, 31.52)	(54.47, 21.99)	(43.60, 38.21)	(45.87, 30.56)	(42.25, 39.08)
NE	(45.87, 30.55)	(38.02, 38.00)	(45.26, 19.28)	(33.23, 42.68)	(33.25, 33.25)	(26.25, 38.37)
NR	(54.47, 21.99)	(47.78, 31.52)	(48.95, 15.91)	(48.07, 33.01)	(45.34, 19.42)	(40.56, 27.63)
UK	(48.33, 32.95)	(43.82, 43.86)	(49.82, 31.64)	(35.38, 45.28)	(37.39, 37.46)	(35.38, 45.28)
UE	(42.25, 39.12)	(35.60, 46.28)	(40.46, 27.53)	(27.66, 47.40)	(26.22, 38.39)	(21.86, 42.17)
UR	(47.58, 34.22)	(43.82, 43.86)	(42.20, 25.61)	(36.64, 44.53)	(32.48, 32.55)	(29.69, 37.61)
KE	(39.46, 39.46)	(31.57, 47.80)	(43.91, 34.33)	(24.93, 50.80)	(30.55, 45.87)	(24.93, 50.80)
KR	(47.80, 31.57)	(41.28, 41.28)	(47.80, 31.57)	(35.58, 46.24)	(38.00, 38.02)	(35.58, 46.24)
ER	(34.33, 43.91)	(31.57, 47.80)	(27.86, 27.86)	(23.53, 52.32)	(19.40, 45.36)	(17.62, 47.10)
NUK	(50.80, 24.93)	(46.24, 35.58)	(52.32, 23.53)	(40.70, 40.70)	(42.68, 33.23)	(39.35, 41.56)
NUE	(45.87, 30.55)	(38.02, 38.00)	(45.36, 19.40)	(33.23, 42.68)	(33.44, 33.44)	(26.46, 38.55)
NUR	(50.80, 24.93)	(46.24, 35.58)	(47.10, 17.62)	(41.56, 39.35)	(38.55, 26.46)	(34.10, 34.10)
NKE	(45.87, 30.55)	(38.02, 38.00)	(50.58, 25.03)	(33.23, 42.68)	(38.62, 38.62)	(31.80, 43.63)
NKR	(54.47, 21.99)	(47.78, 31.52)	(54.31, 22.14)	(43.60, 38.21)	(45.87, 30.56)	(42.25, 39.08)
NER	(45.87, 30.55)	(38.02, 38.00)	(45.36, 19.40)	(33.23, 42.68)	(33.43, 33.43)	(26.46, 38.54)
UKE	(43.59, 38.26)	(35.60, 46.28)	(48.07, 33.05)	(27.66, 47.40)	(33.19, 42.71)	(27.66, 47.40)
UKR	(48.33, 32.95)	(43.82, 43.86)	(49.82, 31.64)	(35.38, 45.28)	(37.39, 37.46)	(35.38, 45.28)
UER	(42.25, 39.12)	(35.60, 46.28)	(40.57, 27.66)	(27.66, 47.40)	(26.42, 38.57)	(22.08, 42.34)
KER	(39.46, 39.46)	(31.57, 47.80)	(43.85, 34.28)	(24.93, 50.80)	(30.55, 45.87)	(24.93, 50.80)
NUKE	(45.87, 30.55)	(38.02, 38.00)	(50.58, 25.03)	(33.23, 42.68)	(38.62, 38.62)	(31.80, 43.63)
NUKR	(50.80, 24.93)	(46.24, 35.58)	(52.32, 23.53)	(40.70, 40.70)	(42.68, 33.23)	(39.35, 41.56)
NKER	(45.87, 30.55)	(38.02, 38.00)	(50.58, 25.03)	(33.23, 42.68)	(38.62, 38.62)	(31.80, 43.63)
NUER	(45.87, 30.55)	(38.02, 38.00)	(45.43, 19.49)	(33.23, 42.68)	(33.56, 33.56)	(26.60, 38.67)
UKER	(43.59, 38.26)	(35.60, 46.28)	(48.07, 33.05)	(27.66, 47.40)	(33.19, 42.71)	(27.66, 47.40)
NUKER	(45.87, 30.55)	(38.02, 38.00)	(50.58, 25.03)	(33.23, 42.68)	(38.62, 38.62)	(31.80, 43.63)

	NKE	NKR	NER	UKE	UKR	UER
N	(50.57, 25.05)	(43.87, 34.31)	(45.03, 19.04)	(52.33, 23.52)	(49.83, 31.58)	(46.79, 17.19)
U	(37.90, 37.97)	(34.13, 47.62)	(32.13, 32.20)	(44.52, 36.70)	(43.27, 43.27)	(37.22, 29.28)
K	(39.98, 39.98)	(33.29, 49.99)	(39.98, 39.98)	(48.36, 37.48)	(45.86, 45.79)	(48.36, 37.48)
E	(25.03, 50.58)	(21.99, 54.47)	(19.02, 45.05)	(33.05, 48.07)	(31.64, 49.82)	(27.26, 40.22)
R	(37.99, 38.03)	(35.29, 43.47)	(32.20, 32.24)	(42.54, 38.49)	(43.34, 43.32)	(34.75, 31.15)
NU	(43.63, 31.80)	(39.08, 42.25)	(38.37, 26.25)	(47.40, 27.66)	(45.28, 35.38)	(42.17, 21.85)
NK	(45.87, 30.56)	(39.44, 39.44)	(45.87, 30.56)	(50.82, 24.90)	(48.37, 32.87)	(50.82, 24.90)
NE	(38.62, 38.62)	(30.56, 45.87)	(33.25, 33.25)	(42.71, 33.19)	(37.46, 37.39)	(38.39, 26.22)
NR	(50.57, 25.05)	(43.81, 34.25)	(45.35, 19.42)	(52.33, 23.52)	(49.83, 31.58)	(47.11, 17.61)
UK	(37.39, 37.46)	(32.87, 48.37)	(37.39, 37.46)	(45.28, 35.44)	(43.27, 43.27)	(45.28, 35.44)
UE	(31.77, 43.66)	(24.90, 50.82)	(26.22, 38.39)	(39.37, 41.58)	(35.44, 45.28)	(33.90, 33.90)
UR	(37.90, 37.97)	(34.13, 47.62)	(32.47, 32.54)	(44.52, 36.70)	(43.27, 43.27)	(37.60, 29.74)
KE	(30.55, 45.87)	(21.99, 54.47)	(30.55, 45.87)	(38.26, 43.59)	(32.95, 48.33)	(39.12, 42.25)
KR	(38.00, 38.02)	(31.52, 47.78)	(38.00, 38.02)	(46.28, 35.60)	(43.86, 43.82)	(46.28, 35.60)
ER	(25.03, 50.58)	(22.14, 54.31)	(19.40, 45.36)	(33.05, 48.07)	(31.64, 49.82)	(27.66, 40.57)
NUK	(42.68, 33.23)	(38.21, 43.60)	(42.68, 33.23)	(47.40, 27.66)	(45.28, 35.38)	(47.40, 27.66)
NUE	(38.62, 38.62)	(30.56, 45.87)	(33.43, 33.43)	(42.71, 33.19)	(37.46, 37.39)	(38.57, 26.42)
NUR	(43.63, 31.80)	(39.08, 42.25)	(38.54, 26.46)	(47.40, 27.66)	(45.28, 35.38)	(42.34, 22.08)
NKE	(38.62, 38.62)	(30.56, 45.87)	(38.62, 38.62)	(42.71, 33.19)	(37.46, 37.39)	(43.66, 31.77)
NKR	(45.87, 30.56)	(39.38, 39.38)	(45.87, 30.56)	(50.82, 24.90)	(48.37, 32.87)	(50.82, 24.90)
NER	(38.62, 38.62)	(30.56, 45.87)	(33.43, 33.43)	(42.71, 33.19)	(37.46, 37.39)	(38.57, 26.42)
UKE	(33.19, 42.71)	(24.90, 50.82)	(33.19, 42.71)	(40.72, 40.72)	(35.44, 45.28)	(41.58, 39.37)
UKR	(37.39, 37.46)	(32.87, 48.37)	(37.39, 37.46)	(45.28, 35.44)	(43.27, 43.27)	(45.28, 35.44)
UER	(31.77, 43.66)	(24.90, 50.82)	(26.42, 38.57)	(39.37, 41.58)	(35.44, 45.28)	(34.12, 34.12)
KER	(30.55, 45.87)	(22.13, 54.31)	(30.55, 45.87)	(38.26, 43.59)	(32.95, 48.33)	(39.12, 42.25)
NUKE	(38.62, 38.62)	(30.56, 45.87)	(38.62, 38.62)	(42.71, 33.19)	(37.46, 37.39)	(43.66, 31.77)
NUKR	(42.68, 33.23)	(38.21, 43.60)	(42.68, 33.23)	(47.40, 27.66)	(45.28, 35.38)	(47.40, 27.66)
NKER	(38.62, 38.62)	(30.56, 45.87)	(38.62, 38.62)	(42.71, 33.19)	(37.46, 37.39)	(43.66, 31.77)
NUER	(38.62, 38.62)	(30.56, 45.87)	(33.56, 33.56)	(42.71, 33.19)	(37.46, 37.39)	(38.68, 26.55)
UKER	(33.19, 42.71)	(24.90, 50.82)	(33.19, 42.71)	(40.72, 40.72)	(35.44, 45.28)	(41.58, 39.37)
NUKER	(38.62, 38.62)	(30.56, 45.87)	(38.62, 38.62)	(42.71, 33.19)	(37.46, 37.39)	(43.66, 31.77)

	KER	NUKE	NUKR	NKER	NUER	UKER	NUKER
N	(54.47, 21.99)	(50.57, 25.05)	(48.07, 33.01)	(50.57, 25.05)	(45.03, 19.04)	(52.33, 23.52)	(50.57, 25.05)
U	(47.58, 34.22)	(37.90, 37.97)	(36.64, 44.53)	(37.90, 37.97)	(32.13, 32.20)	(44.52, 36.70)	(37.90, 37.97)
K	(50.04, 33.31)	(39.98, 39.98)	(37.48, 48.30)	(39.98, 39.98)	(39.98, 39.98)	(48.36, 37.48)	(39.98, 39.98)
E	(34.33, 43.91)	(25.03, 50.58)	(23.53, 52.32)	(25.03, 50.58)	(19.02, 45.05)	(33.05, 48.07)	(25.03, 50.58)
R	(43.46, 35.34)	(37.99, 38.03)	(38.47, 42.53)	(37.99, 38.03)	(32.20, 32.24)	(42.54, 38.49)	(37.99, 38.03)
NU	(50.80, 24.93)	(43.63, 31.80)	(41.56, 39.35)	(43.63, 31.80)	(38.44, 26.33)	(47.40, 27.66)	(43.63, 31.80)
NK	(54.47, 21.99)	(45.87, 30.56)	(43.60, 38.21)	(45.87, 30.56)	(45.87, 30.56)	(50.82, 24.90)	(45.87, 30.56)
NE	(45.87, 30.55)	(38.62, 38.62)	(33.23, 42.68)	(38.62, 38.62)	(33.33, 33.33)	(42.71, 33.19)	(38.62, 38.62)
NR	(54.31, 22.14)	(50.57, 25.05)	(48.07, 33.01)	(50.57, 25.05)	(45.42, 19.50)	(52.33, 23.52)	(50.57, 25.05)
UK	(48.33, 32.95)	(37.39, 37.46)	(35.38, 45.28)	(37.39, 37.46)	(37.39, 37.46)	(45.28, 35.44)	(37.39, 37.46)
UE	(42.25, 39.12)	(31.77, 43.66)	(27.66, 47.40)	(31.77, 43.66)	(26.30, 38.46)	(39.37, 41.58)	(31.77, 43.66)
UR	(47.58, 34.22)	(37.90, 37.97)	(36.64, 44.53)	(37.90, 37.97)	(32.55, 32.62)	(44.52, 36.70)	(37.90, 37.97)
KE	(39.46, 39.46)	(30.55, 45.87)	(24.93, 50.80)	(30.55, 45.87)	(30.55, 45.87)	(38.26, 43.59)	(30.55, 45.87)
KR	(47.80, 31.57)	(38.00, 38.02)	(35.58, 46.24)	(38.00, 38.02)	(38.00, 38.02)	(46.28, 35.60)	(38.00, 38.02)
ER	(34.28, 43.85)	(25.03, 50.58)	(23.53, 52.32)	(25.03, 50.58)	(19.49, 45.43)	(33.05, 48.07)	(25.03, 50.58)
NUK	(50.80, 24.93)	(42.68, 33.23)	(40.70, 40.70)	(42.68, 33.23)	(42.68, 33.23)	(47.40, 27.66)	(42.68, 33.23)
NUE	(45.87, 30.55)	(38.62, 38.62)	(33.23, 42.68)	(38.62, 38.62)	(33.56, 33.56)	(42.71, 33.19)	(38.62, 38.62)
NUR	(50.80, 24.93)	(43.63, 31.80)	(41.56, 39.35)	(43.63, 31.80)	(38.67, 26.60)	(47.40, 27.66)	(43.63, 31.80)
NKE	(45.87, 30.55)	(38.62, 38.62)	(33.23, 42.68)	(38.62, 38.62)	(38.62, 38.62)	(42.71, 33.19)	(38.62, 38.62)
NKR	(54.31, 22.13)	(45.87, 30.56)	(43.60, 38.21)	(45.87, 30.56)	(45.87, 30.56)	(50.82, 24.90)	(45.87, 30.56)
NER	(45.87, 30.55)	(38.62, 38.62)	(33.23, 42.68)	(38.62, 38.62)	(33.56, 33.56)	(42.71, 33.19)	(38.62, 38.62)
UKE	(43.59, 38.26)	(33.19, 42.71)	(27.66, 47.40)	(33.19, 42.71)	(33.19, 42.71)	(40.72, 40.72)	(33.19, 42.71)
UKR	(48.33, 32.95)	(37.39, 37.46)	(35.38, 45.28)	(37.39, 37.46)	(37.39, 37.46)	(45.28, 35.44)	(37.39, 37.46)
UER	(42.25, 39.12)	(31.77, 43.66)	(27.66, 47.40)	(31.77, 43.66)	(26.55, 38.68)	(39.37, 41.58)	(31.77, 43.66)
KER	(39.40, 39.40)	(30.55, 45.87)	(24.93, 50.80)	(30.55, 45.87)	(30.55, 45.87)	(38.26, 43.59)	(30.55, 45.87)
NUKE	(45.87, 30.55)	(38.62, 38.62)	(33.23, 42.68)	(38.62, 38.62)	(38.62, 38.62)	(42.71, 33.19)	(38.62, 38.62)
NUKR	(50.80, 24.93)	(42.68, 33.23)	(40.70, 40.70)	(42.68, 33.23)	(42.68, 33.23)	(47.40, 27.66)	(42.68, 33.23)
NKER	(45.87, 30.55)	(38.62, 38.62)	(33.23, 42.68)	(38.62, 38.62)	(38.62, 38.62)	(42.71, 33.19)	(38.62, 38.62)
NUER	(45.87, 30.55)	(38.62, 38.62)	(33.23, 42.68)	(38.62, 38.62)	(33.73, 33.73)	(42.71, 33.19)	(38.62, 38.62)
UKER	(43.59, 38.26)	(33.19, 42.71)	(27.66, 47.40)	(33.19, 42.71)	(33.19, 42.71)	(40.72, 40.72)	(33.19, 42.71)
NUKER	(45.87, 30.55)	(38.62, 38.62)	(33.23, 42.68)	(38.62, 38.62)	(38.62, 38.62)	(42.71, 33.19)	(38.62, 38.62)

References

Buchanan, Allen and Russell Powell, *The Evolution of Moral Progress: A Biocultural Theory*, Oxford: Oxford University Press, 2018.

Curry, Oliver, Daniel Austin Mullins, and Harvey Whitehouse, "Is It Good to Cooperate: Testing the Theory of Morality-as-Cooperation in 60 Societies," *Current Anthropology*, 2019, *60* (1), 47–69.

de Waal, Frans, *The Bonobo and the Atheist: In Search of Humanism Among the Primates*, New York: W. W. Norton & Company, 2014. Accessed January 2, 2025.

Harms, William and Brian Skyrms, "Evolution of Morality," in "Oxford Handbook of Philosophy of Biology," Oxford New York: Oxford Univeristy Press, 2008, pp. 434–450.

Korsgaard, Christine M., "Morality and the Distinctiveness of Human Action," 2006. Accessed January 2, 2025.

Krebs, Dennis, "Why We Are Not Selfish By Nature," in Michael Ruse, ed., *The Philosophy of Biology: A Companion for Educators*, Oxford University Press, 2011, pp. 359–377. Accessed January 2, 2025.

Nagel, Thomas, *Mortal Questions*, Cambridge: Cambridge University Press, 1979.

Parfit, Derek, *On What Matters, Volumes 1 and 2*, Oxford University Press, 2011.

Rawls, John, *A Theory of Justice*, revised edition ed., Cambridge, MA: Harvard University Press, 1971. Revised edition published in 1999.

Schroder, Wilhelm, "Continental Ethics," in Hugh LaFollette, ed., *The Blackwell Guide to Ethical Theory*, Oxford: Blackwell Publishing, 2000, pp. 375–399.

Smyth, Nicholas, "The Function of Morality," *Philosophical Studies*, 2017, *174* (5), 1127–1144.

Suresh, Shyam Gouri and Paul Studtmann, "Angels and Devils on Our Shoulders: A Framework for Modeling Moral Agency," January 2023.