



Shiny Moral People

A Modeling Approach towards Understanding Moral Hypocrisy within a Virtual Society

Markus Christen,

University of Zurich & University of Notre Dame



Table of Contents

- Moral Hypocrisy: The Question(s)
- Conceptualization of Moral Hypocrisy in the Model
- Pre-Tests and Determination of Main Scenarios
- Three Main Results
- Conclusion



**University of
Zurich** ^{UZH}

Institute of Biomedical Ethics

**University of
Notre Dame**



Psychology Department

Moral Hypocrisy: The Question(s)



A definition of moral hypocrisy (Dan Batson):

“(...) avoid the cost of being moral while maintaining the appearance of morality (...).”



Ted Haggard



Eliot Spitzer



Karl-Theodor zu Guttenberg



Tiger Woods



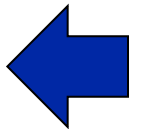
Issues that frame the concept of ‘moral hypocrisy’

- 1) **A normative condemnation:** Moral hypocrisy is the wrong way to fill the fundamental “is-ought-gap” that morality implies.
- 2) **Internalism/Externalism:** Moral hypocrisy questions the motivational force of moral standards/beliefs persons hold.
- 3) **A social understanding of morality:** A safeguard of society from behaviors profitable for individuals but damaging for the group.
- 4) **Moral reputation as a core factor in morality:** Reputation is considered to be an essential component for the development of morality in human foragers, where each individual is strongly aware that he or she must have a positive reputation in case of future need, and painfully guards it (Hrdy 2009).
- 5) **Temptations as gains:** Moral behavior is (partially) understood to involve disadvantages for an individual in the sense of “missed opportunities”.
- 6) **Cover is necessary:** Moral hypocrisy requires violating moral norms such that the violation is not detected (e.g. subtle cheating; Trivers 1971).



Research questions with respect to moral hypocrisy

- 1) **Philosophy:** To what extent is moral hypocrisy a necessary part of human life (to fight the “terror of morality”)? How is moral hypocrisy related to the concept of morality one holds?
- 2) **Personality Psychology:** How can individuals maintain a motivational state with the ultimate goal to appear moral while, if possible, avoiding the costs to self of actually being moral (Batson et al. 1997).
- 3) **Social Psychology:** Why do individuals’ evaluations of their own moral transgressions often differ substantially from their evaluations of the same transgressions enacted by others (Valdesolo & DeSteno 2007)?
- 4) **Sociology:** What is the effect of social strategies intended to avoid moral hypocrisy on the prevalence of moral hypocrisy?





**University of
Zurich** ^{UZH}

Institute of Biomedical Ethics

**University of
Notre Dame**



Psychology Department

Conceptualization of Moral Hypocrisy in the Model



Agent states and behavior types

The model implements the conceptual idea of moral hypocrisy by distinguishing two different types of agent-states:

- The **reputation** of the agent (either morally good, G, or bad, B)
- Its **disposition to act** toward temptations (either to be tempted, T, or to resist a temptation, R).

This offers four different behaviors to the agents:

- Appearing good and resist a temptation (GR; “good guys”)
- Appearing good but being tempted (GT; “hypocrites”)
- Appearing bad and being tempted (BT; “bad guys”)
- Appearing bad but resisting temptations (BR; “inconsistent guys”).



Moral hypocrisy as the “best behavior”

The payoff structure represents the basic idea of moral hypocrisy, i.e. an agent gains most if he takes the bait: the model *assumes* that moral hypocrisy is the optimal behavior for a single agent within a society.

		Disposition to act	
		Be tempted (T)	Resist temptation (R)
Reputation	Good (G):	GT (yellow): one point for each temptation and for each neighbor	GR (blue): one point for each neighbor
	Bad (B)	BT (red): one point for each temptation	BR (pink): 0

The goal of this study **is to assess the success of different strategies compared to a benchmark** (no strategy installed) in terms of changes in the population distribution of agents that follow one of the four behaviors.



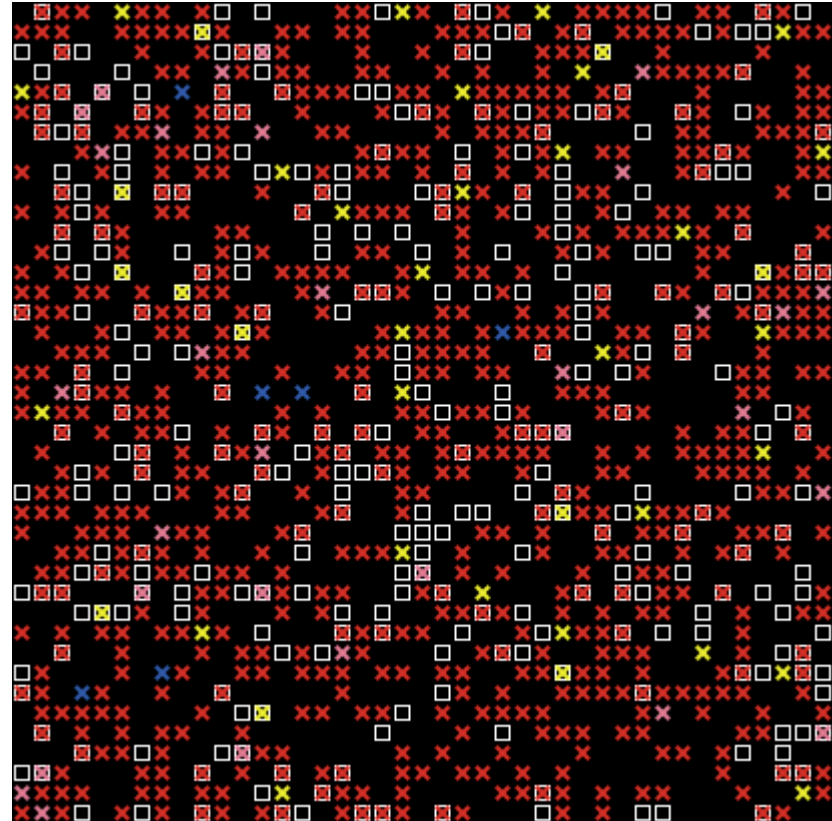
Social Strategies to overcome moral hypocrisy

#	Description of Strategy
1	Avoid agents that are tempted: Every agent that has either a yellow or red neighbor moves to the closest free cell on the lattice without such neighbors (if possible).
2	Seek agents with good reputation: Every agent that does not yet have either a blue or yellow neighbor moves to the closest free cell on the lattice with at least one such neighbor (if possible).
3	Disclose hypocrite (local version): Whenever the majority of agents in a two-degree Moore neighborhood of a yellow agent is non-yellow, the yellow agent changes his behavior to BT (red).
4	Disclose hypocrite (global version): Whenever the majority of agents of a specified yellow agent is non-yellow, the yellow agent changes his behavior to BT (red).
5	First strategy 2, then strategy 1
6	First strategy 1, then strategy 3
7	First strategy 1, then strategy 4
8	First strategy 2, then strategy 3
9	First strategy 2, then strategy 4
10	First strategy 3, then strategy 2, then strategy 1
11	First strategy 4, then strategy 2, then strategy 1



What the model does

1. It selects an agent A
2. It calculates the payoff of A
3. It does 1 and 2 for each agent chosen in a random order
4. It changes the behavior of each agent to the behavior of its best performing neighbor
5. It applies the strategy (1-11) and goes back to step 1.
6. It stops when the model reaches a quasi-stable state (no population size changes)





Model Parameters

- | | |
|---|----------------------|
| <ul style="list-style-type: none">- Initial probability of having a good reputation- Initial probability of being tempted if a temptation is present- Strategies | Full sampling |
| <ul style="list-style-type: none">- Population Density- Temptation Density- Sequences of single steps within strategies | Pre-Test |
| <ul style="list-style-type: none">- Neighborhood for strategy comparison- Payoffs- Initial spatial distribution of agents- Initial spatial distribution of agents- Agent-temptation interaction | No changes |



**University of
Zurich** ^{UZH}

Institute of Biomedical Ethics

**University of
Notre Dame**



Psychology Department

Pre-Tests and Determination of Main Scenarios



Paradigmatic scenarios – Description

Scenario A – Pre-Modern: Low population (10%) and low temptation density (5%). Pre-modern societies consist of small groups that have high social control minimizing the number of available temptations.

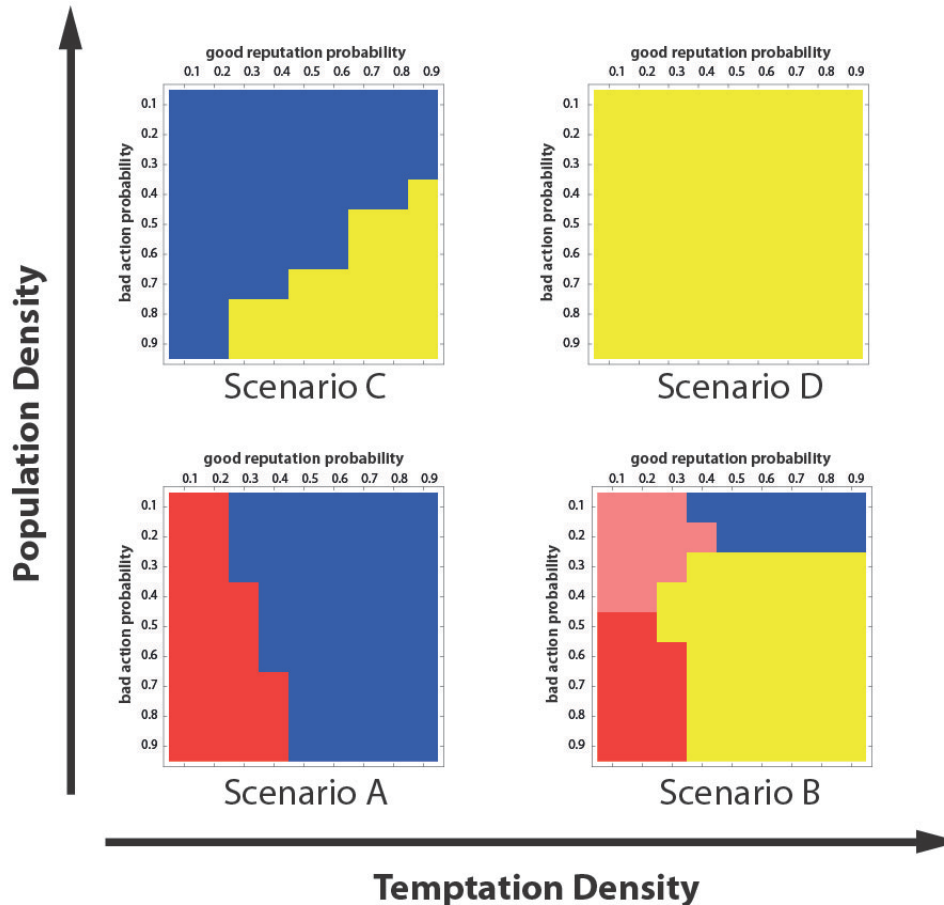
Scenario B – Modern Agricultural: Low population (10%) and high temptation density (50%). Modern agriculture consists of large farms (low population density) that have access to all means of modern societies in terms of mobility, communication etc. that tend to increase the “temptation space”.

Scenario C – Brave New World: High population (66%) and low temptation density (5%). A city state (Singapore?) with a tight control regime with respect to temptations.

Scenario D – Sin City: High population (66%) and high temptation density (50%). It implements the idea of a densely populated city full of temptations (Chicago?).



Paradigmatic scenarios – Majorities





**University of
Zurich** ^{UZH}

Institute of Biomedical Ethics

**University of
Notre Dame**

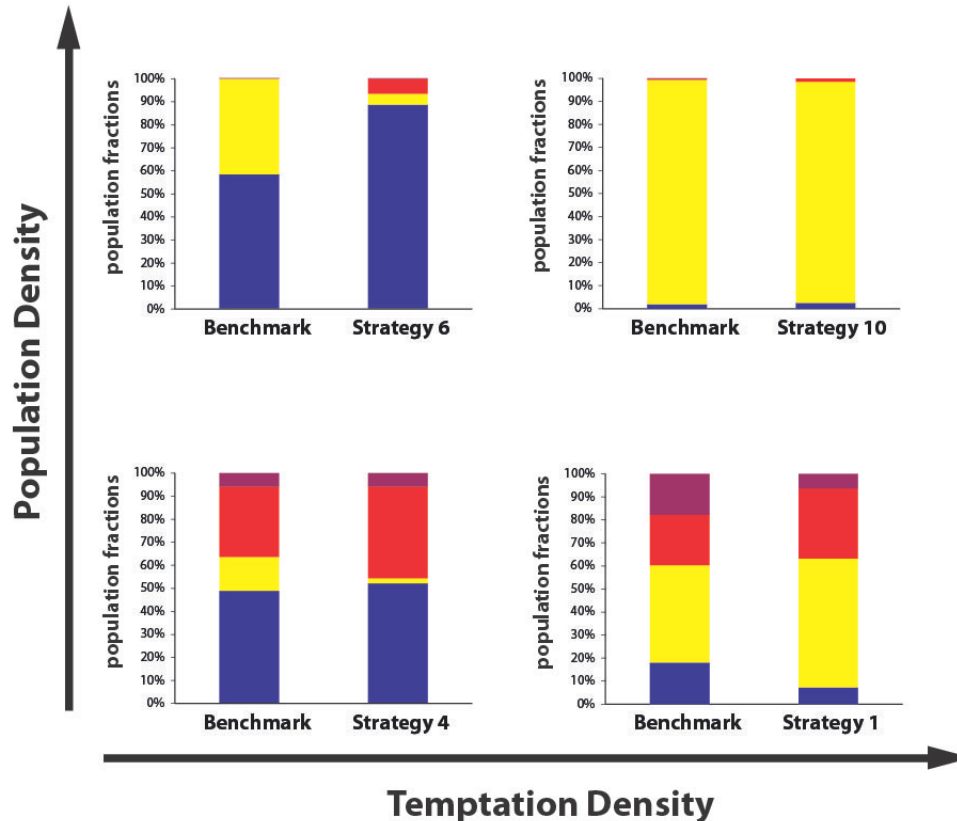


Psychology Department

Three Main Results



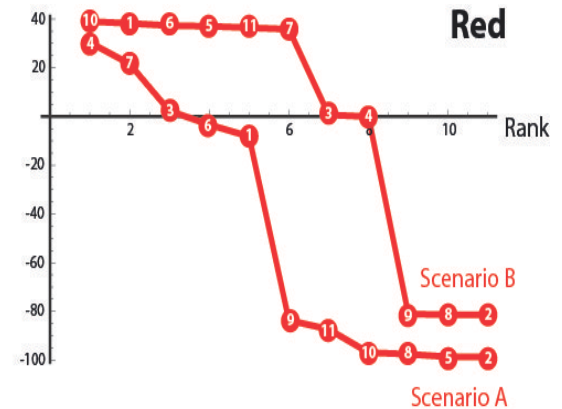
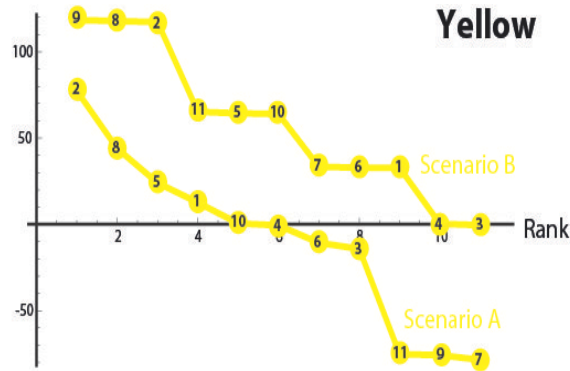
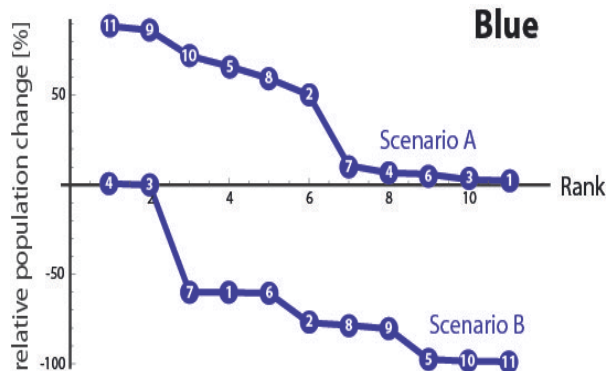
Result 1: Scenario parameters are the main determinants of population distributions





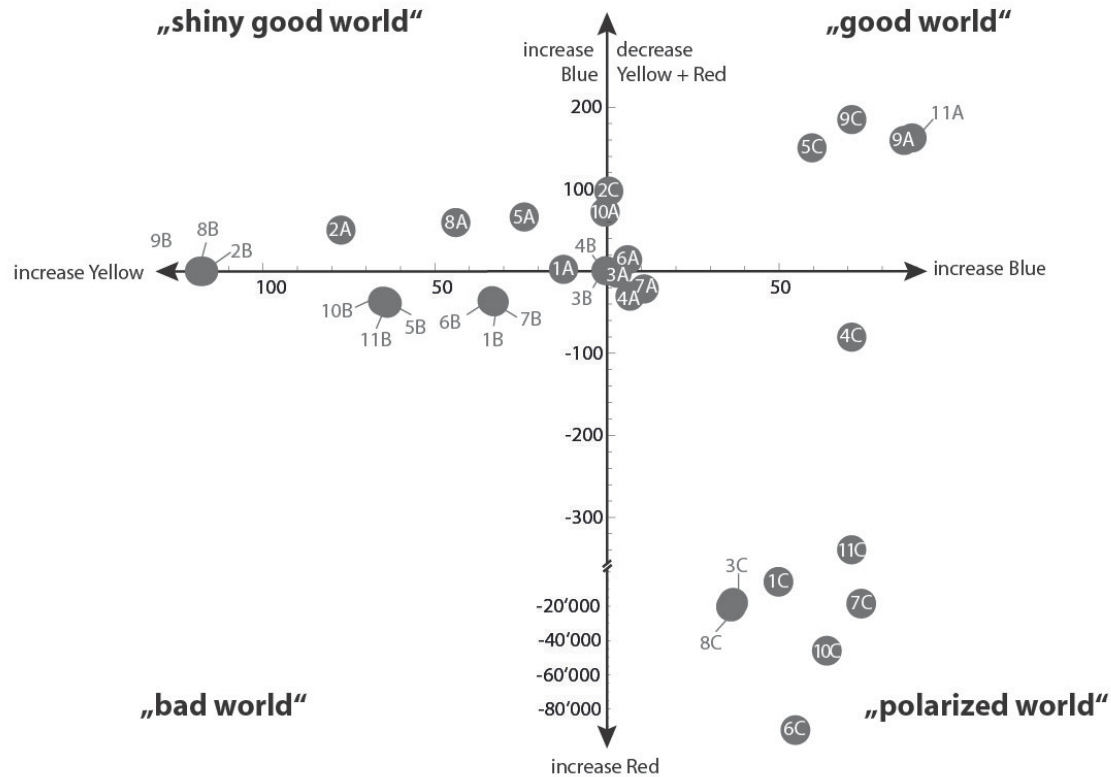
Result 2: Strategy rankings reveal conflicting effects of interventions

We ranked the strategies according to their ability to increase the population of a specific behavior type relative to the benchmark population size. We display the two sequences for each population with the highest dissimilarity measured by the the Kendall Rank correlation (a measure for the similarity of rankings).





Result 3: Strategy effects can be attributed to four "moral worlds"





How scenarios frame the effect of interventions

- The Modern Agricultural scenario creates a context that promotes bad world strategies – i.e. red and yellow can often increase their weight.
- The Brave New World scenario creates a context that promotes polarizing world strategies increasing both the blue and the red population.
- The Pre-Modern scenario creates a context that promotes shiny good world strategies.

Effect of strategies:

- Disclosing strategies in their pure form (3 and 4) tend to be polarizing, i.e. form strong minorities of red agents.
- Strategy 2 tends to be “shiny”, which is plausible as the yellow population profits from a strategy that benefits good reputation.



**University of
Zurich** ^{UZH}

Institute of Biomedical Ethics

Conclusion

**University of
Notre Dame**



Psychology Department



Main Points

We have shown that there is no optimal strategy to act against moral hypocrisy independent of population and temptation density – parameters that determine paradigmatic social scenarios

In fact, the most successful strategies in one scenario can have disastrous consequences in other scenarios.

The findings therefore support a context-sensitive approach towards the problem of avoiding moral hypocrisy.



Some thoughts for real world interventions

- Disclosing hypocrites, in a Brave New World scenario may have the unwanted effect to also increase the population of “bad guys”.
- Strategies against moral hypocrisy in a society with relatively low interaction density due to rather low population density (Modern Agricultural scenario) generally tend to be unsuccessful, indicating that not moral hypocrisy itself, but other aspects of social organization (e.g. interaction density) may be better targets for social interventions.



Reminder of shortcomings

The current model disregards the heterogeneity of modern societies with respect to the generic social scenarios used.

The current model does not take into account the psychological complexity of moral hypocrisy with respect, e.g., to the type of temptation.

The model does not account for real world handling of moral hypocrisy, e.g. forgiveness.

The current analysis does not involve all possible social strategies against moral hypocrisy.

Other model parameters may become object of further investigations (e.g. changes in the payoff-structure, non-Moorean interactions between agents).



**University of
Zurich** ^{UZH}

Institute of Biomedical Ethics

**University of
Notre Dame**



Psychology Department

Thank you!

**And thanks to the members of the Institute
“Computer Simulation in the Humanities”**