5-1-2023

God games: An experimental study of uncertainty, superstition, and cooperation

Aidin Hajikhameneh
San Jose State University, aidin.hajikhameneh@sjsu.edu

Laurence R. Iannaccone
Chapman University

Games and Economic Behavior 139 (2023) 88–116

Contents lists available at ScienceDirect

Games and Economic Behavior

journal homepage: www.elsevier.com/locate/geb

God games: An experimental study of uncertainty, superstition, and cooperation ☆

Aidin Hajikhameneh [a,*], Laurence R. Iannaccone [b]

[a] *Department of Economics, San Jose State University, Dudley Moorhead Hall, 1 Washington Square, San Jose, CA 95112, USA*
[b] *Institute for the Study of Religion, Economics, and Society, Chapman University, 338 N. Glassell, Orange, CA 92866, USA*

A R T I C L E I N F O

Article history:
Received 2 November 2021
Available online 27 January 2023

JEL classification:
C91
C92
Z12
Z13

Keywords:
Religion
Superstition
Uncertainty
Collective action
Experiments

A B S T R A C T

This paper uses a novel lab experiment to test claims about the origins and functions of religion. We modify the standard public goods game, adding a computer-based agent that adjusts earnings in ways that might depend on players' contributions. Our treatments employ three different descriptions of the adjustment process that loosely correspond to monotheistic, atheistic, and agnostic interpretations of the computer's role. The adjustments neither mask players' contributions nor magnify their impact. Yet players in all three adjustment treatments contribute much more than those who play the standard public goods game. Players' contributions and survey responses show that adjustments induce superstitions in all treatments, with the strongest superstitions appearing in the quasi-monotheistic treatment and the weakest in the quasi-atheistic treatment. Text-based communication raises contributions and strengthens coordination. But when paired with the quasi-monotheistic description, communication also promotes counterproductive quests for winning numbers.

1. Introduction

Can the methods of experimental economics create something approaching supernatural belief in the lab? And can this pseudo-supernaturalism shed light on religion in the real world? We tackled both questions with a simple but novel experiment designed to test classic claims about the origins and functions of religion. According to the first claim, people more readily attribute outcomes to supernatural powers when facing substantial risk or uncertainty. According to the second, the resulting attributions can raise welfare by boosting confidence and cooperation.

Our experiment modifies the standard public goods game, adding a computer-based agent that adjusts players' earnings in a way that might plausibly depend on their contributions. The experiment includes four core treatments, three of which introduce the adjustment process with narrative frames that loosely correspond to monotheistic, atheistic, and agnostic

☆ We thank James Andreoni, Christopher Bader, Gary Charness, Joseph Henrich, Daniel Houser, Hillard Kaplan, Michael McBride, Joshua Miller, Douglas Norton, Nathan Nunn, Ryan Oprea, Benjamin Priday, Jared Rubin, Eric Schniter, Roman Sheremeta, and Nathaniel Wilcox for their comments and suggestions, along with seminar participants at Chapman University, Florida State U., Harvard U., San Jose State U., and U.C. Irvine, and session participants at the 2018 ASREC, 2018 BABEEW, 2018 ESA, and 2019 SEA conferences. We gratefully acknowledge funding support from the John Templeton Foundation (award #59214), Chapman University's Argyros School of Business, and Chapman's Institute for the Study of Religion, Economics, and Society.

* Corresponding author.
E-mail addresses: aidin.hajikhameneh@sjsu.edu (A. Hajikhameneh), iannacco@chapman.edu (L.R. Iannaccone).

https://doi.org/10.1016/j.geb.2023.01.004

interpretations of the computer's role.[1] In the fourth treatment, which served as a baseline, subjects merely played the standard public goods game. The quasi-monotheistic frame emphasized the computer's attention to players' contributions; the quasi-atheistic frame emphasized the exogenous random nature of the adjustment process; the quasi-agnostic frame was uninformative, noting only that the adjustments were based on a complex sequence of calculations.

The computer's adjustments were simple and separately reported amounts – always equal to $+5$ or -5 tokens, and hence half the 10-token endowment received by each player at the start of each round. The adjustments neither masked players' contributions nor magnified their impact. Their arithmetic impact was equivalent to a post-play series of 20 coin flips. Nevertheless, players substantially altered their behavior in response to these arithmetically irrelevant adjustments, and they did so in ways that fit the classic claims noted above. Relative to the zero-adjustment baseline, average contributions in all three adjustment treatments increased by about 50%. We also observed strong evidence of superstitious responses to the computer's adjustments, with the largest responses in the quasi-monotheistic treatment and the smallest in the quasi-atheistic treatment.[2]

To learn more about players' thoughts and facilitate shared belief, we ran four communication treatments in which group members could converse by text. This raised contributions further still, but when paired with the monotheistic description it also promoted counterproductive quests for winning numbers.

We use the term *pseudo-supernaturalism* to distinguish our approach from those that precede it. In recent decades economists and other social scientists have run numerous well-designed lab and field experiments measuring the economic impact of religions.[3] Taken together with other studies based on wide range of methods and sources, these experiments have persuaded many researchers that religions have shaped economic development across cultures and over time. The experiments are compelling because they concern real religions and real settings. But these same strengths make it all but impossible to abstract from religion's complexity and context. We seek to complement existing work by creating something akin to religion and magic *in the lab*. Social scientists routinely define religion and magic as systems of belief and practice premised upon the existence of supernatural beings or forces (Stark and Bainbridge (1996); Iannaccone (2006); Norenzayan et al. (2016); Norenzayan (2013)). Our experiment adds a simple lab-based analog of supernaturalism to an otherwise standard finitely-repeated linear public goods game (PGG).

Fig. 1 provides a preview of what lies ahead. The solid black line toward the bottom tracks average contributions in the Baseline treatment. As in numerous replications of the standard public goods experiment, average contributions start high, drop rapidly in the first few rounds, decay more slowly in subsequent rounds, and plummet near the end. In contrast, the solid red line shows that the (quasi-monotheistic) *Watchful* treatment raises average contributions by more than 50%. The contributions also display less decay in the middle rounds and a less dramatic drop toward the end. But we cannot conclude that the narrative framing caused players to treat the computer like a god that rewards charity and punishes greed. The solid green and blue paths show that the (quasi-atheistic) *Random* and (quasi-agnostic) *Unknown* treatments are no less effective than their *Watchful* counterpart. As a mechanism for raising cooperation, earnings shocks worked surprisingly well no matter how we described them.

The communication treatments in dotted black, red, blue, and green provide a second surprise. As in past experiments, text-based chat dramatically raises contributions.[4] But in this case, chat also causes the paths to diverge, with the Watchful treatment now yielding the lowest overall rates of contribution, lower even than the Baseline treatment. Something is going on, and as we will show in sections 5.4, this "something" has more in common with *magic* than *religion*.[5]

Our results also suggest some general observations. First, lab experiments may provide an effective setting in which to study the causes and consequences of supernaturalism. Second, when designing such experiments, economists do well to start with the findings of other social scientists who have carefully studied religion since the late 1800's. Third, studies of pseudo-supernaturalism can enrich the growing body of economic research on real religions. Fourth, other standard experiments should probably be revisited to test whether seemingly irrelevant forms of risk and uncertainty affect their outcomes as strongly as those of the public goods game.[6] We may find that humans respond to cheap noise no less than cheap talk and that both religious and secular institutions adapt to channel these responses. Fifth, exogenous shocks may be viewed

[1] Our narrative frames correspond to different views of the supernatural only in the very limited sense that they suggest alternative views of the relationship between the players' contributions and the computer's adjustments. For more on our choice of terms and narratives, see section 3.

[2] We define superstitions as demonstrably false rules of conduct that attribute valued outcomes to feasible actions and persist within a group over time. By "demonstrably false," we mean that the group can disconfirm the superstition's implicit claims of cause and effect within reasonable time and cost using readily available knowledge, methods, and resources.

[3] Hoffmann (2013) reviews experiments in the economics of religion prior to 2010. For references to more recent experiments and related experiments in other fields, see Benjamin et al. (2016) and Norenzayan (2013). For examples of religious priming, see Benjamin et al. (2016) and Warner et al. (2015). For insights derived from comparisons among religions, see Ruffle and Sosis (2006) and Chaudhary and Rubin (2016).

[4] See Wilson and Sell (1997) for a summary of experimental findings regarding communication. Brandts et al. (2019) recently reviewed the various forms of communications implemented in laboratory environments.

[5] We follow Stark and Bainbridge (1996), Iannaccone and Berman (2006), and (Roberts and Yamane, 2015, p. 11) who distinguish religion and magic as alternative forms of supernaturalism, the former centered on interacting with supernatural beings and the latter centered on manipulating impersonal supernatural forces.

[6] Several published papers study how uncertainty in the return to public or private good investment affects contributions in public goods games (Dickinson, 1998; Cox and Stoddard, 2021; Stoddard, 2017; Théroude and Zylbersztejn, 2020; Dorner et al., 2020). In contrast to these and other papers, the additive risk and uncertainty in our experiments do not change the rate of return to public or private investments, endowments, or any other element of the standard linear public goods game profit function.

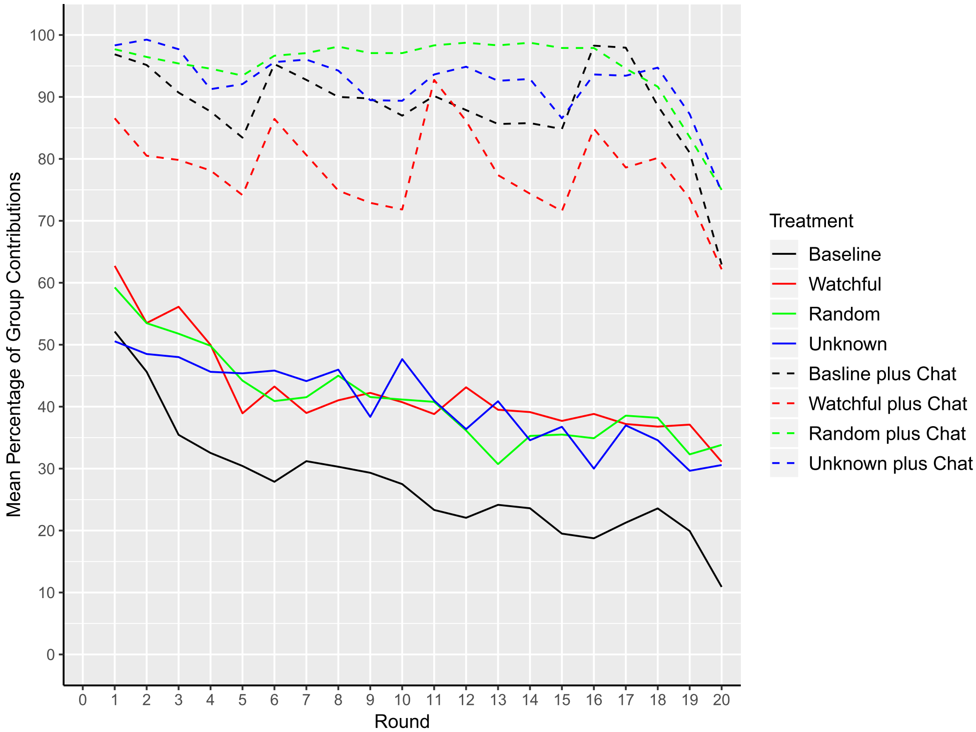

Fig. 1. Mean contributions, plotted by round

as one member within a family of mechanisms that promote collective action without recourse to reward, punishment, exclusion, or other previously studied methods that require information about individuals' actual rates of cooperation. The mechanisms also include communication, costly signaling (Iannaccone, 1992; Gintis et al., 2001; Ruffle and Sosis, 2006), distinctive identities, and some forms of endogenous sorting (Aimone et al., 2013). We think it no coincidence that all these mechanisms are prevalent among religions that foster high levels of commitment.

2. Supernaturalism

Supernaturalism provides a natural starting point for economic studies of religion because of its universal presence in societies and its historic role in legitimizing norms and customs (Henrich et al., 2010; Malefijt, 1968; Wade, 2009; Steadman and Palmer, 2015). Anthropologists have studied supernaturalism across hundreds of cultures, cataloging their findings together with numerous other measures in databases suitable for statistical analysis (Ember and Ember, 2011). Supernatural beliefs underpin moral precepts in Judaism, Christianity, Islam, Hinduism, and Buddhism. Evolutionary biologists, pyschologists, and anthropologists argue that these and other "big god" religions were key to the emergence of large scale civilizations and market economies (Norenzayan, 2013; Norenzayan et al., 2016; Purzycki et al., 2016, 2018; Wade, 2009; Wright, 2010; Stark, 2005). Recent contributions to economic history demonstrate that supernaturalism has played a significant role in promoting or inhibiting trust, property rights, literacy, equality, democracy, and growth; and recent contributions to development economics demonstrate that it continues to do so.[7] The study of supernaturalism also complements the shift in experimental economics from theories of rational calculation and pure self-interest to the messier world of hard-wired emotions, evolved predispositions, and social preferences (Chaudhuri, 2011, 2016).

For experimental studies of supernaturalism, there's no better starting point than the work of Bronislaw Malinowski, one of the most influential anthropologists of the 20th century and one of the few with training in economics. He demolished the claim that the religions of ancient and indigenous cultures were the products of limited mental powers, that by extension all religions rest on ignorance and irrationality, and that science and supernaturalism must therefore be fundamentally incompatible. Malinowski showed that the Trobriand Islanders of Melanesia pursued most of their activities in remarkably rational ways:

[7] See, for example, Rubin (2017), Kuran (2012), Nunn and Sanchez de la Sierra (2017), McCloskey (2006, 2010, 2016), and Gershman (2016). Hayek (1989) appears to be the first major economist to make the argument in evolutionary terms, noting that "mystical and religious beliefs" were key to the transmission of customs and beliefs that sustained "the evolution of moral orders through group selection" and that monotheism provided the pre-rational moral foundation for the "extended [market] order" (pp. 136 & 137).

"If you were to suggest to a native that he should make his garden mainly by magic and skimp his work, he would simply smile on your simplicity. He knows as well as you do that there are natural conditions and causes, and by his observations he knows also that he is able to control these natural forces by mental and physical effort".

[Malinowski, 1948, pp. 11-12]

But empiricism did not eliminate supernaturalism. Rather, as demonstrated most clearly in their fishing, the Islanders reserved the supernatural for what could be neither understood nor controlled:

While in the villages on the inner lagoon fishing is done in an easy and absolutely reliable manner by the method of poisoning, yielding abundant results without danger and uncertainty, there are on the shores of the open sea dangerous modes of fishing and also certain types in which the yield greatly varies according to whether shoals of fish appear beforehand or not. It is most significant that in the Lagoon fishing, where man can rely completely upon his knowledge and skill, magic does not exist, while in the open-sea fishing, full of danger and uncertainty, there is extensive magical ritual to secure safety and good results ... [Thus] primitive man recognizes both the natural and the supernatural forces and agencies, and he tries to use them both for his benefit ... He never relies on magic alone, while, on the contrary, he sometimes dispenses with it completely ... But he clings to it, whenever he has to recognize the impotence of his knowledge and of his rational technique.

[Malinowski, 1948, p. 14-15]

Subsequent ethnographic studies draw similar conclusions about the role of uncertainty in the formation of both superstitions and supernaturalism (e.g., Palmer, 1989), and history provides countless examples of new technologies displacing supernatural approaches to previously insoluble problems.

Though religions are cultural products with complex histories, supernatural beliefs and practices are supported in part by general features of cognition. People routinely "see" trends, turning points, and clusters in random phenomena (including the randomness of stock market returns, winning streaks, or spatial distributions from aerial bombing) – partly because our brains extract patterns from very limited information, and partly because the consequences of randomness can be sufficiently subtle and counterintuitive that even trained statisticians misinterpret its real-world manifestations (Palmer, 1989; Gilovich et al., 1985; Miller and Sanjurjo, 2018; Benjamin, 2019; Oskarsson et al., 2009). We are likewise predisposed to ascribe anthropomorphic agency and emotions to high-stakes outcomes that have purely natural causes (Boyer, 2007).

In a famous study of " 'Superstition' in the Pigeon," the psychologist B. F. Skinner (1948) identified a related phenomenon derived from nothing more than the conjunction of random actions and unrelated outcomes. When placed in separate cages with hoppers that supplied food at predetermined intervals, pigeons developed patterns of behavior that were distinctive, durable, and entirely idiosyncratic. "The bird behaves as if there were a causal relation between its behavior and the presentation of food, although such a relation is lacking." While claiming only that "the experiment might be said to demonstrate a sort of superstition," Skinner emphasized that "[t]here are many analogies in human behavior. Rituals for changing one's luck at cards are good examples. A few accidental connections between a ritual and favorable consequences suffice to set up and maintain the behavior in spite of many unreinforced instances" (Skinner, 1992, p. 274). Though human superstitions have many sources, subsequent studies show that the risks associated with sports, fishing, warfare, and many other endeavors routinely lead people to adopt rituals with little or no connection to their goals (Bleak and Frederick, 1998).

3. Experiment design

The public goods game provides a natural starting point for experiments on supernaturalism – partly because of its simple structure and trade-offs, but also because of its repeated play in stable groups. In the standard game subjects play in equal-sized groups, receive an equal number of tokens at the start of each round, divide their own tokens between their personal account and their group's account, earn payoffs equal to the number of tokens kept in their personal accounts plus a multiple of the total contributed to their group's account, and finish after a known number of rounds.

To make the game more like fishing in the Trobriand Islands we needed a source of earnings shocks. We therefore added a second stage to each round in which the computer adjusted group earnings, adding or subtracting five tokens to each member's payoff. Players were truthfully told that all members of their group receive the same adjustments but different groups may receive different adjustments.

Because faith is founded on social interactions, we also needed a substitute for cultural inheritance. Children acquire their beliefs from other people, starting with their parents and the friends and teachers chosen by their parents. Though people often stray from early indoctrination, they rarely migrate to entirely different faiths (Berger, 1969; Iannaccone, 1992; Stark, 1999). Our experiment required an analog of socialization that neither pushed players toward a single worldview nor left them with nothing but the personal views they carried into the lab.

Our solution was to run treatments that provided different descriptions of the computer's adjustment process. While containing no false information, these narrative frames emphasized different aspects of the adjustment algorithm. Within the toy world of the experiment the three frames correspond roughly (albeit *very* roughly) to monotheistic, atheistic, and agnostic views of the supernatural. These views encompass those professed by the vast majority of all Americans including more than 95% of college students (Wormald, 2015; Froese and Bader, 2010).

The computer was programmed to monitor what every player does and influence what every group earns. Insofar as the computer rewards some types of behavior or punishes others, it functions in the manner of a "moralistic High God" as defined by Norenzayan (2013); Stark (2003); Purzycki et al. (2016) and others. Hence, the narrative frame for our (quasi-monotheistic) *Watchful* treatment noted that the computer's adjustment calculations took account of players' contribution decisions.[8]

On the other hand, insofar as the adjustment process operates without regard to players' actions, the game lacks *any* sort of god-like agents, moralistic or otherwise. Hence, our (quasi-atheistic) *Random* treatment emphasized that each adjustment had a 50/50 chance of adding or subtracting five tokens from the earnings of each group member, making adjustments both unpredictable and independent of player actions. The (quasi-agnostic) *Unknown* treatment provided an uninformative description, which said only that the computer would choose its ± 5 adjustments based on a complex sequence of calculations. We ran a fourth treatment that omitted the adjustment stage and thereby provided a standard public goods *Baseline*. To provide more scope for social effects, we also ran four communication treatments. These were identical to the four core treatments except that each gave players opportunities to chat within their groups through a shared on-screen text box.

The narrative frames correspond to different views of the supernatural only in the narrow sense that they provide alternative views of the relationship between the players' contributions and the computer's adjustments.[9] Nothing in the frames, the game itself, or the overall experiment referenced the supernatural either directly or indirectly. We did not anticipate significant correlations between players' actions and their personal religious beliefs and practices, nor did our post-game questionnaire reveal any – see section 4.4.

Note also the boundaries within which our experiment operates. Social scientific studies of religion limit themselves to secular standards of inquiry and argument, an approach that Peter Berger (1969, pp. 100, 180) aptly described as "methodological atheism." Supernatural belief and behavior are explained as consequences of purely natural phenomena. Malinowski's theory is a case in point, with attributions of supernatural agency sustained by uncertainty, instruction, and indirect benefits. Experiments must likewise seek to induce the experimental analogs of supernatural belief and behavior *without* resort to experimental analogs of supernatural intervention.

3.1. Core treatments

The *Baseline* treatment employed parameters typical of past public goods game experiments: 4-person groups, 10-token endowments, a group contribution multiplier of 0.4, and 20 periods. At the end of each period, players received on-screen reports that listed their group's combined contribution, their own contribution, their personal account and group account earnings, and their total earnings. Each round consists of a single stage in which players divide their endowments between their group's joint account and their personal accounts. Each player i then receives earnings determined by the formula: $\pi_i = (E - g_i) + \mu G_i$ where E denotes the endowment, g_i is i's contribution in the group account, G_i is the total amount invested by i's group, $\mu = 0.4$, and tokens are worth 5¢. Experiments routinely refer to this as the *Voluntary Contribution Method* or VCM.

The three adjustment treatments added a second stage to each round in which players receive additional earnings of ± 5 tokens. In all three treatments players read and heard detailed instructions that differed in just one way: the phrasing of the paragraph that described the adjustment process. All other features of the experiment remained the same, including session size, group size, number of rounds, earnings formula, and adjustment algorithm. Appendix B specifies the complete instructions in each treatment.

Below, we provide the relevant paragraph with the key phrases highlighted in bold italics. For complete instructions see Appendix B.

The Watchful treatment narrative reads as follows:

> In the second stage of each period, ***a computer-based Artificial Intelligence (or "AI") will make an adjustment to your earnings.*** The adjustment will be either +5 or -5. The same adjustment will be made to the earnings of each member in your group. ***The AI will choose the adjustment for your group based on a complex series of calculations that take account of your investment decisions, the investment decisions of the other members in your group, and the investment decisions of the other groups in this experiment.*** The AI will do separate calculations and make separate adjustment decisions for the other three groups.

[8] Though it seems unlikely that any player ever thought about the computer in these terms, it would be natural for players to wonder how contributions factored into the computer's adjustments, and thoughtful players may well have considered this question in light of the experimenters' probable interests, assumptions, or goals.

[9] Note also that the frames by no means exhaust the ways in which a supernatural agent or natural environment might respond to players' actions. In Hajikhameneh and Iannaccone (2022) we explore the effects of additional narrative frames and adjustment processes, including one in which each player's random adjustments are separately calculated. The results help explain *why* shocks promote pro-social behavior, but they do not contradict the results or conclusions reported in the current study.

As seen in Appendix D, the adjustment algorithm is indeed based on a complex series of calculations and is parameterized to place more or less weight on the absolute size of the group's contributions, the ranking of the group's contributions relative to those of other groups, and random effects. To keep the focus on Malinowski's theory of supernaturalism, Skinnerian superstition, and the impact of (instruction-based) socialization, our Watchful treatments placed all weight on the random effects. Following the advice of reviewers, we later ran a second set of Watchful treatments with modified parameters that shifted some of the weight to the group's contribution relative to that of other groups. Depending on whether a group's contribution in period t ranked 1st, 2nd, 3rd, or 4th relative to those of the other groups, it received a positive contribution with probability 0.52, 0.51, 0.49 or 0.48. In all other respects, this *modified Watchful* treatment was identical to the original Watchful treatment: same instructions, same group size, same session sizes, same earnings function, and same university lab. Although participants in the original and modified sessions came from distinct student cohorts, separated by five years and the Covid crisis, results from the second sets of sessions largely mirrored those of the original treatments. Appendix E provides detailed comparisons of the original and modified treatment results. We focus hereafter on the eight original treatments graphed in 1, all of which were conducted between April 2017 and May 2018. Appendix A lists the date of every session.

The Random treatment narrative reads as follows:

> In the second stage of each period, the computer will make adjustments to the earnings of all the participants in the experiment. The adjustments will be either +5 or -5 tokens. The same adjustment will be made to the earnings of every member in your group, but the members of other groups may receive different adjustments. ***The computer will choose the adjustment for your group based on a random number calculation, so that there will be a 50% chance that 5 tokens are added to the earnings of everyone in your group and a 50% chance that 5 tokens are subtracted from every member's earnings.*** The computer will calculate a different random number for each of the other three groups in the experiment.

The Unknown treatment aimed for Knightian uncertainty and therefore provided minimal information about the adjustments, and therefore notes only that each was based on a complex sequence of calculations:[10]

> After the members of your group have made their investment decisions, the computer will make an adjustment to your earnings. The adjustment will be either +5 or -5 tokens. The same adjustment will be made to the earnings of every member in your group, but the members of other groups may receive different adjustments. ***The computer will choose the adjustment for your group based on a complex sequence of calculations***. The computer will choose a separate adjustment for each of the other three groups based on a separate sequence of calculations.

3.2. Communication treatments

The four communication treatments were identical to their core treatment counterparts except that prior to rounds 1, 6, 11, and 16 the members of each group could chat with one another by typing into a shared text box.[11] Section 5 provides detailed analyses of both the consequences and content of players' communications.

3.3. Other features

We ran all sessions at a medium-sized American university with subjects randomly drawn from the pool of students registered to participate in economics experiments. Each session included 16 subjects randomly assigned to anonymous 4-person groups, and no subject participated in more than one session. All groups played 20 rounds of just one treatment, with each player receiving 10 tokens at the start of each round, contributing a portion that could vary by increments of 0.01, and receiving the rest in their personal account. At the end of each period, each player received an on-screen report that listed the group's combined contribution, their own contribution, their personal account earnings, their group account earnings, their 5-token adjustment (if non-Baseline players), and their total earnings. As seen in Fig. B1, the reports clearly distinguish earnings from the player's personal account, the group's joint account, and the computer's adjustment.

Risk aversion: To assess the extent to which players' actions were influenced by their risk preferences, we followed the game with an incentivized risk elicitation task (Holt and Laury, 2002). For details, see Appendix B.4.

[10] Although contemporary economists sometimes use the terms ambiguity and uncertainly interchangeably, we prefer to reserve "ambiguity" for situations in which ambiguity aversion is clearly relevant – i.e., situations in which actors confront an Ellsberg-style choice between outcomes governed by a known probability distribution versus outcomes with unknown probabilities (Ellsberg, 1961; Machina and Siniscalchi, 2014). Depending on their treatment, our players confronted adjustment risk *or* adjustment uncertainty but never both.

[11] Communication has been added to public goods games in many forms. In their seminal paper Isaac and Walker (1988) used face-to-face communications after each round of play. Bochet et al. (2006) found that verbal anonymous chat raised cooperation as effectively as face-to-face communication. Koukoumelis et al. (2012) implemented a one-way communication from group leader to members prior to the first round of the game. Brandts et al. (2016) found that free-form, people oriented communications by leaders tended to revive cooperation in the public goods game. For more examples see Chaudhuri (2011).

Questionnaire: Each session concluded with an on-line questionnaire that asked subjects for their major field of study, gender, religious beliefs and behavior, and open-ended feedback. The Unknown and Watchful treatment questionnaires also asked subjects if they noticed any patterns in the adjustments. For details, see Appendix B.5.

Earnings: Overall earnings included a $7 show-up payment, the value of the tokens earned over all 20 rounds (converted at 5¢ per token), and money earned from the risk elicitation task. The experiment included 29 sessions, with 16 subjects per session, and 464 subjects overall. We ran 16 core sessions: four each of the Baseline, Watchful, Unknown, and Random treatments. The 256 core treatment subjects earned total payoffs that averaged $21.49 and ranged from $14.50 to $29.40. We ran 13 communication sessions: four Baseline and three of each adjustment treatment. The 208 communication treatment subjects earned total payoffs that averaged $23.79 and ranged from $16.45 to $30.12. As explained in Appendix C, we supplemented the experiment with three sessions that modified the Watchful communication treatment by expanding the set of adjustment values from ± 5 to $\{-5, -2, 0, +2, +5\}$. We programmed all versions of the experiment in z-Tree (Fischbacher, 2007), analyzed data with Stata, and created figures with R.

4. Core treatments

4.1. Predictions

Based on the research and reasoning described above, we made several predictions before running our core treatments:

Prediction 1. In the Baseline treatment, contribution rates will consistently exceed the (zero-contribution) Nash equilibrium, decay over time, and approach zero in the final rounds.

Prediction 2. Relative to the Baseline treatment, contribution rates in the Watchful treatment will be higher and decay more slowly.

Prediction 3. Players in the Random treatment will behave in same manner as players in the Baseline treatment.

Prediction 4. The Unknown treatment will yield intermediate behavior and responses relative to the Watchful and Random treatments.

Prediction 5a. Unconscious (Skinner-style) superstition will cause players' actions to display more inertia following positive adjustments. Hence, relative to a negative adjustment, a positive adjustment will raise the probability that players repeat their previous contributions.

Prediction 5b. Unconscious superstition will induce inertia effects of roughly equal magnitude in all three adjustment treatments.

Prediction 6. Conscious (Malinowski-style) superstition will operate most strongly among players who view adjustments as quasi-supernatural responses to their behavior. Watchful treatment players will therefore display more overall inertia than their Random treatment counterparts following positive adjustments. In the Unknown treatment overall inertia will fall between that of the Watchful and Random treatments.

Though we anticipated the direction of each effect, we were not at all sure that the magnitudes would be substantively large or statistically significant. There is, after all, no prior research on the sort of descriptive primes we used and no way to know how they might interact with subjects' priors. Moreover, the core treatments lacked any analog of the social interactions that are central to real-world supernaturalism.

4.2. Basic results

As previously noted many of key results are captured by Fig. 1, which plots the paths of average contributions for the eight core and communication treatments.[12]

Baseline: The solid black line at the bottom of Fig. 1 strongly confirms Prediction 1. Though by no means surprising, it's good to know that our Baseline results mirror those routinely obtained in standard VCM experiments. Apparently, there was nothing odd about our basic design or subject pool.

Watchful: Past experiments offer little guidance regarding the Watchful treatment effects, since they lack anything like an all-knowing *deus ex machina* that deals out punishment and reward. Game theory is likewise of limited help given our intentionally ambiguous description of the computer's adjustment process. Some players might assume that the computer will

[12] The core-treatment contribution paths look virtually the same whether plotted as round-by-round means or medians.

Table 1
Summary statistics - core treatments

Treatments	Group Contributions averaged over all rounds			
	Mean	SE	Median	N
Baseline	10.99	0.539	9.25	320
Random	16.50	0.445	16.00	320
Unknown	16.23	0.448	15.22	320
Watchful	16.94	0.483	16.25	320

punish groups whose members fail to contribute most of their endowments to their group account. Others might assume that the computer tolerates or even rewards groups whose members are clever enough to anticipate a zero contributions equilibrium. Still others might imagine that the computer rewards groups that invest more than other groups, or have the fewest free-riders, or maintain the highest rates of contribution over time.[13] As play proceeds and the computer makes its actual adjustments, players might begin to suspect that the process is, in fact, random. But they will be hard-pressed to confirm this guess given that no one observes the individual actions of other members nor anything at all about other groups.

On the other hand, the Watchful treatment seeks to mimic the world of Malinowski's ocean-going Trobriand Islanders where groups impute supernatural agency to environments characterized by uncertainty and collective action. Although the resulting beliefs and practices will vary greatly across cultures, the underlying function should be to promote group welfare through increased effort and solidarity.

This theory certainly fits Fig. 1. Compared to the Baseline VCM, the Watchful treatment yields a contribution path that is dramatically higher, decays more slowly, and tails off much less. Moreover, as we will see later in this section, regression analysis shows that these dramatic differences persist even after introducing a variety of other explanatory variables. At first glance, Prediction 2 stands up remarkably well.

Random: Game theory offers a clear prediction in the face of risk that is both additive and exogenous. To see why, rewrite a typical player's total earnings over all rounds of the game, collecting the adjustments at the end of the expression:

$$\pi_i = \sum_{t=1}^{20} \pi_{it} = \sum_{t=1}^{20} \left[(E - g_{it}) + 0.4 G_{it} \right] + \sum_{t=1}^{20} a_{it}$$

Recall also that we truthfully tell Random treatment players that each adjustment, a_{it}, has 50/50 odds of being -5 or +5, and we report each adjustment separately from the overall group contribution, G_{it}. Apart from an overall payoff shock that averages \$0 and falls between ±\$1.50 about 90% of the time, the Random treatment is equivalent to the Baseline PG game treatment.

Absent huge income or risk effects, we therefore expect no change in player behavior. Yet, as seen in Fig. 1, this seemingly straightforward prediction fails and does so quite dramatically. Not only are Random treatment contributions higher than those of the Baseline, they are virtually identical to those of the Watchful treatment.

Unknown: Prediction 4 succeeds insofar as the path of Unknown treatment contribution tracks those of the Watchful and Random treatments. But this predictive "success" merely underscores the overall puzzle. Why are the paths of all three adjustment treatments so similar to each other yet so different from the Baseline?

Statistical analysis confirms the visual story of Fig. 1. Table 1 lists the group-level summary statistics for each of the four core treatments. Group contributions across the three adjustment treatments are essentially equal and exceed those of the Baseline treatment by about 50%.[14] Despite being uninformative and unaffected by players' decisions, random shocks appear to boost cooperation as well as or better than many methods explored in the public goods literature.[15]

[13] Thoughtful players might conclude that zero-contribution remains the dominant strategy by reasoning as follows: The computer knows nothing about a group's players apart from their randomly-assigned IDs and contribution choices. Hence, members' actions will likely enter the adjustment function symmetrically. This probably means the marginal impact of my own current-round contribution averages no more than 1/4 of the 10-token adjustment range and less still if the adjustment function includes other inputs, such as past round contributions, the contributions of other groups, or random shocks.

[14] We ran two-sided Wilcoxon rank sum tests for each pair of treatments. The tests included 16 observations per treatment, one observation for each group, with each observation equal to the total contributions of the group's 4 members averaged over 20 rounds. The comparisons between adjustment treatments were all statistically insignificant – e.g., the rank sum test for equality of average Watchful and Random contributions yields a p-value of 0.68. The tests comparing Baseline contributions to those of each adjustment treatment are all statistically significant, with p-values less than 0.01. The robust rank order test of Feltovich (2003, 2005) yields similar results.

[15] Such methods include different forms of punishment and reward, approval and disapproval, exogenous sorting, endogenous group formation, and more (Chaudhuri, 2011, 2016; Villeval, 2020).

Table 2
GLS analysis of contributions

	(1) Contribution$_t$	(2) Contribution$_{t>1}$	(3) Contribution$_{t>1}$	(4) Contribution$_{t>1}$
Random	-0.109	-0.104	0.616	0.733
	(0.541)	(0.550)	(1.146)	(1.135)
Unknown	-0.178	-0.126	0.581	0.686
	(0.568)	(0.577)	(1.347)	(1.312)
Positive Adjustment$_{t-1}$		0.184**	0.181**	0.328**
		(0.089)	(0.084)	(0.140)
Risk			0.588	0.605
			(2.067)	(2.061)
Random × Risk			-1.626	-1.639
			(2.409)	(2.405)
Unknown × Risk			-1.530	-1.548
			(2.714)	(2.708)
Female			-0.700**	-0.698**
			(0.287)	(0.287)
Round			-0.096***	-0.096***
			(0.013)	(0.013)
Random × Positive Adjustment$_{t-1}$				-0.232
				(0.193)
Unknown × Positive Adjustment$_{t-1}$				-0.207
				(0.200)
Intercept	4.235***	4.042***	5.240***	5.170***
	(0.409)	(0.423)	(0.954)	(0.945)
Observations	3840	3648	3648	3648
R^2	0.00	0.00	0.04	0.04

Standard errors clustered by group in parentheses.
* p<0.1, ** p<0.05, *** p<0.01.
The Watchful treatment is the omitted category.

4.3. Regression results

There is, however, more at work than meets the eye. Despite the apparent equivalence of the Watchful, Random, and Unknown treatments in Fig. 1, and despite the fact that all three treatments employ the same adjustment algorithm, a closer look at the data show that behavior does differ across treatments. Moreover, the differences accord with both Malinkowski's theory of supernaturalism and Skinner's theory of superstition.

Contributions: Regression analysis illuminates how players responded to adjustments and how these responses differed across treatments, changed over time, and varied with gender, risk preferences, and other variables. Table 2 summarizes GLS regressions (with group-clustered standard errors and subject-specific random effects because subjects played multiple rounds). The dependent variable (Contribution) measures the number of tokens each player allocated to the group account in each of the 20 rounds. Because we are seeking to understand how players respond to adjustments, we omit the Baseline treatment and restrict our analysis to the three adjustment treatments, using Watchful as the omitted category.

Because contributions must fall in the zero-to-ten token range we also ran Tobit regressions with bootstrapped standard errors and the same random effects as in GLS. Because decisions to fully free-ride may reflect a qualitatively different view of the game, we also ran regressions that employed the double-hurdle estimation method developed by Cragg (1971) and applied to public goods game data by Engel and Moffat (2012) and Cheung (2014). See Tables F1 and F3 of Appendix F for details, and note that the key GLS results carry over to both the Tobit and Hurdle regressions.

The first column in Table 2 verifies that overall rates of contribution are nearly equal across the three adjustment treatments but the remaining columns reveal differences that cannot be inferred from simple plots and summary statistics. For example, Watchful treatment players become more public-spirited following a reward: they contribute *more* when the previous round's adjustment was positive rather than negative. This result has no clear basis in standard game theory nor in standard theories of superstition.[16]

Columns 3 and 4 show that contributions decline over time, that women contribute less than men, and that contributions are not significantly related to risk preferences as measured by our version of the Holt and Laury (2002) risk elicitation task.[17]

[16] Given that most players are contributing well below the maximum, there's no obvious reason why rewards should boost group contributions more effectively than punishments. Moreover, a simple subconscious response to positive reinforcement would show up in all three adjustment treatments, but column 4 shows that the net lagged adjustment effects are substantially smaller and statistically insignificant in the Random and Unknown treatments.

[17] When added to the regressions, the effect of the lagged average of other members' contributions is positive – the opposite of that which is often seen in standard public goods experiments (Ashley et al., 2010; Heap et al., 2016). This could reflect efforts to coax positive adjustments out of the computer. Whatever its cause, including this lagged effect does not change the statistical significance or substantive impact of the other variables in Table 2.

Table 3
GLS analysis of absolute changes and no changes in contribution decisions

	(1) Absolute Change	(2) No Change (all cases)	(3) No Change (0's & 10's omitted)
Random	-0.136	0.016	-0.001
	(0.221)	(0.046)	(0.045)
Unknown	-0.087	-0.036	0.016
	(0.217)	(0.041)	(0.043)
Positive Adjustment$_{t-1}$	-0.849***	0.162***	0.170***
	(0.128)	(0.033)	(0.033)
Random × Positive Adjustment$_{t-1}$	0.337**	-0.082**	-0.092**
	(0.169)	(0.036)	(0.040)
Unknown× Positive Adjustment$_{t-1}$	0.242	-0.048	-0.074*
	(0.181)	(0.040)	(0.041)
Intercept	2.187***	0.302***	0.193***
	(0.138)	(0.034)	(0.035)
Observations	3648	3648	3175
R^2	0.02	0.02	0.02

Standard errors clustered by group in parentheses.
* $p<0.1$, ** $p<0.05$, *** $p<0.01$.
The Watchful treatment is the omitted category.

Inertia: Consider what we would expect to see if our human subjects displayed the same sort of unconscious acquisition of superstitious behaviors that Skinner (1948) induced in his randomly-reinforced pigeons. Skinner's fundamental insight was that random rewards condition animals to associate the rewards with whatever happened to precede them. These associations are certainly not the product of conscious reflection, but they can be modeled as by-products of adaptive learning strategies or natural selection (Beck and Forstmeier, 2007; Foster and Kokko, 2009). Hence, Predictions 5a and 5b state that positive adjustments would lead to larger contributions and the magnitude of this unconscious inertia effect would be roughly equal across all three treatments.

If players also display *conscious* forms of superstition like those of Malinowski's Trobriand Islanders, then the overall inertia effect should differ across treatments. Watchful treatment players will more likely infer a relationship between their actions and the computer's subsequent adjustments. Hence, Prediction 6 states that overall inertia will be greatest in the Watchful treatment.

Table 3 tests Predictions 5a, 5b, and 6 by regressing two measures of contribution inertia on the previous round's adjustment. The first measure of inertia (*AbsChange*) equals the absolute change in each player's contribution from the previous round to the current round. The second measure (*NoChange*) equals 1 if contributions in the current and preceding round were equal and 0 otherwise. We then run GLS regressions of each inertia measure on *PosAdjust*, a 0/1 indicator of positive adjustment in the previous round interacted with Random and Unknown (versus Watchful) treatment indicators. As in our previous GLS regressions, we cluster standard errors by group and include player-specific random effects.

The regressions of Table 3 provide strong support for Predictions 5a and 6 and indirect support for 5b. The *AbsChange* coefficients in column (1) show that in period immediately following a negative adjustment Watchful treatment players change their contributions by an average of 2.19 tokens. But the average change is about 40% smaller (0.85 tokens less) immediately following a positive adjustment. As predicted, the corresponding estimates of inertia are smaller in the Random treatment: the change in contributions averages 2.05 tokens after a negative adjustment but only 1.53 (25% less) following a positive adjustment. The Unknown treatment's 29% reduction lies between those of the Watchful and Random treatments.

The *NoChange* regression in column (2) tells a similar story for the *NoChange* measures of inertia. Following a negative adjustment, Watchful treatment players repeat their previous contribution 30% of the time, but the probability of exact repetition rises to 46% following a positive adjustment. The corresponding probabilities are smaller but still significant in the Random treatment, where a positive adjustment raises the probability of exact repetition from 30% to 39%. Again, the Unknown treatment's effect lies between those of the Watchful and Random treatments. Because players might repeat zero- or ten-token contributions simply because they cannot contribute outside this range, column (3) restricts the *NoChange* regression to observations in which the player's previous contribution fell strictly within the (0,10) range. Although the coefficients differ somewhat from those of the unrestricted regression, the relative magnitudes and statistical significance persist.

In short, inertia effects in the Random treatment point to unconscious Skinner-style superstitions; the much stronger Watchful treatment effects suggest that conscious Malinowski-style superstitions are also at work; and the intermediate effects in the Unknown treatment are consistent with the hypothesis that unconscious superstitions arise with little or no explicit support whereas conscious superstitions require explicit narratives.

4.4. Questionnaire results

The results from our end-of-experiment survey provide some additional evidence for Malinowski-style superstitions. We asked subjects in the Watchful and Unknown treatments to tell us whether they observed any patterns in the computer's adjustments and, if so, what the patterns were. Only 26% of the subjects in the Watchful treatment and 30% of the subjects in the Unknown treatment claimed to have spotted patterns – shares that may have been depressed by the fact that saying "yes" required the subject to type out a description of the pattern. Nevertheless, among those who claimed to see a pattern, the most common conclusion (espoused by slightly more than half of all respondents) was that the computer was rewarding the group for high overall contributions and punishing low. Interestingly, no one claimed that the computer rewarded *low* contributions. Hence, insofar as our Watchful and Unknown treatments did induce something akin to supernaturalism, the imputed *deus ex machina* was always pro-social in character, rewarding benevolence and punishing selfishness.

The questionnaire provided no evidence that players' responses to the adjustment treatments were influenced by the real-world religious views they brought into the lab. Average contributions did not correlate with frequency of church attendance, childhood church attendance, belief that spirits exist, belief that some people can mentally heal others, or belief that some people can "see" into the future.[18] See Appendix B.5 for the survey questions and response categories.

5. Communication treatments

As we have noted, supernaturalism is a social phenomenon. To provide more scope for social effects, we ran a series of communication sessions. Specifically, we let group members chat with one another prior to rounds 1, 6, 11, and 16 through a on-screen text box. The communication treatments were in all other respects identical to their core treatment counterparts. The conversations reveal much about how superstitious inferences arise and how they are collectively reinforced.

5.1. Predictions

To the extent that readers can ignore what they saw about the communication treatments from Fig. 1, we invite predictions concerning the impact of communication on the four treatments' contribution paths. While asking you to forget what you may already have noted about communication in *this* experiment, we nevertheless encourage you to take account of what is generally known about the impact of communication on public goods games. After obtaining the core treatment results described above, we ourselves predicted the following:

Prediction 7. Compared to its non-communication counterpart, the path of average contributions in the Baseline-plus-chat treatment will be much higher and decline more slowly.

Prediction 8. Relative to the Baseline-plus-chat treatment, the path of contributions in the Watchful-plus-chat treatment will be higher still and decline even more slowly.

Chastened by our failed prediction that Random treatment contributions would equal those of the Baseline, we conjectured that the Random-plus-chat path would more nearly follow that of the Watchful treatment:

Prediction 9. Rates of group contribution in the Random-plus-chat treatment will exceed those of the Baseline-plus-chat treatment and approach but not exceed that of the Watchful-plus-chat treatment.

And once again, we predicted intermediate results for the Unknown treatment:

Prediction 10. The Unknown-plus-chat treatment will yield intermediate behavior and responses relative to the Watchful and Random treatments.

5.2. Basic results

Turning to Fig. 1, again, we see that our Baseline-plus-chat results are consistent with past studies and hence also consistent with Prediction 7. But we have again failed to predict the relative effects of the different treatments. This time our error concerns the Watchful treatment rather than the Random treatment.

Comparing the summary statistics in Tables 1 and 4, we see that communication does indeed send cooperation soaring. Whereas group contributions in the original (non-communication) Baseline averaged 10.99 tokens (27% of the 40-token maximum); group contributions in the Baseline-plus-chat treatment average 35.44 tokens (87% of the maximum); and group

[18] The two-sided Spearman's rank correlation test had *p-values* of 0.42, 0.97, 0.88, 0.42, and 0.50 respectively for Watchful treatment players and *p–values*= 0.73, 0.53, 0.17, 0.60, and 0.23 for the entire sample.

Table 4
Summary statistics – communication treatments

Treatments	Group contribution in the joint account Mean	SE	Median	N
Baseline-plus-chat	35.44	0.551	40	320
Random-plus-chat	37.94	0.332	40	240
Unknown-plus-chat	36.95	0.406	40	240
Watchful-plus-chat	31.35	0.713	38.75	240

Table 5
GLS analysis of contributions with communication

	(1) Contribution$_t$	(2) Contribution$_{t>1}$	(3) Contribution$_{t>1}$	(4) Contribution$_{t>1}$
Random plus Chat	1.647***	1.676***	1.288**	1.341**
	(0.590)	(0.595)	(0.605)	(0.681)
Unknown plus Chat	1.401**	1.413**	0.338	0.439
	(0.650)	(0.657)	(0.877)	(0.899)
Positive Adjustment$_{t-1}$		-0.018	0.007	0.124
		(0.132)	(0.120)	(0.299)
Risk			-1.852***	-1.862***
			(0.671)	(0.666)
Random plus Chat × Risk			0.987	0.997
			(0.793)	(0.786)
Unknown plus Chat × Risk			2.732***	2.744***
			(1.059)	(1.060)
Female			-0.263	-0.264
			(0.334)	(0.334)
Round			-0.046*	-0.046*
			(0.027)	(0.027)
Random plus Chat × Positive Adjustment$_{t-1}$				-0.126
				(0.354)
Unknown plus Chat × Positive Adjustment$_{t-1}$				-0.221
				(0.350)
Intercept	7.838***	7.804***	9.190***	9.140***
	(0.566)	(0.587)	(0.715)	(0.769)
Observations	2880	2736	2736	2736
R^2	0.18	0.18	0.21	0.21

Standard errors clustered by group in parentheses.
* p<0.1, ** p<0.05, *** p<0.01.
The Watchful plus Chat treatment is the omitted category.

contributions in the Random-plus-chat treatment average 37.94 tokens (95% of the maximum). Though the overall difference in group contributions is statistically insignificant in the rank sum test, the Random-plus-chat contributions exceed their Baseline-plus-chat counterparts in 18 of the 20 rounds.

But contributions in the Watchful-plus-chat sessions average just 31.35, 78% of the 40-token maximum. They fall below their Baseline counterparts in 19 of 20 rounds and below the Random counterparts in 20 of 20 rounds, differences that *are* significant in the rank sum tests (at p-values of 0.02 and 0.00, respectively). Though technically “intermediate,” contributions for the Unknown-plus-chat groups nearly equal those of their Random counterparts and substantially exceed those of the Watchful-plus-chat groups.

Table 5 summarizes GLS regression results for the chat treatments, and Table F2 of Appendix F reports on the corresponding Tobit regressions. As in the core treatment regressions of Table 2, the dependent variable measures individual contributions to the group account. The regressions again focus on the relative size on contributions across the three adjustment treatments, and specifications 2, 3, and 4 again control for the previous round’s adjustment, risk aversion, gender, and interaction terms. The first specification confirms the non-parametric conclusion that compared to the (omitted category) Watchful treatment, contributions are substantially (and significantly) higher when subjects are explicitly told that the adjustments are random. Thus, communication breaks the equality of contributions that characterized the Watchful versus Random treatments when players could *not* communicate.

The statistically significant coefficients associated with risk aversion, round, and gender effects are generally not the same for GLS and Tobit estimations, and the most consistently significant effect (that of risk aversion) is the opposite of what we would expect to see and also the opposite of what we observed in the non-communication treatments. The weak effect of subject-level attributes makes sense because (as shown in section 5.4) communication strongly encourages the emergence of group-level agreements on contribution rates.

The contribution paths, summary statistics, and regression results all raise the same question: *Why does communication cause Watchful-treatment contributions to fall so far behind those of all other communication treatments?* We argue below that the answer traces back to the particular form of superstitions induced by the Watchful description, superstitions best characterized as *numerology*. These quasi-magical superstitions are particularly counterproductive in linear games with corner solutions.

5.3. Counterproductive superstitions

Malinowski found that Trobriand Islanders resorted to magical thinking only when danger, uncertainty, and limited knowledge rendered rational empiricism nearly worthless. When they could more readily control and comprehend cause and effect they routinely relied upon experience, skill, and practical reasoning. Our Random treatments were designed to provide players with the analog of lagoon fishing – an environment characterized by the clear rules and the predictable risk of a computer-generated coin flip as opposed to the uncertainty of a god-like agent. Insofar as our Random-treatment subjects embraced this description of the adjustment process, they operated in a clear and quasi-atheistic environment within which they could readily identify the optimal group-level strategy.[19]

In combination with the contribution results, these considerations suggest a generalization:

Conjecture. *The search for illusory patterns and quasi-magical numbers lead the members of Watchful-plus-chat groups to agree on contribution strategies that differ from the full-contribution strategy more frequently than groups in other chat treatments.*

We subsequently found strong support for this conjecture. Virtually all groups *except* those of the Watchful treatment routinely agreed to fully contribute. For example, at the start of the game, members of Random-plus-chat group #21-3 wrote as follows:

Initial communication, before round 1:
Member # 10: *what is the move haha*
Member # 12: *ok so hear me out, if we all put in 10 into the group fund, we make 16 give or take 5 every round*
Member # 11: *Let's invest within the group*
Member # 12: *if one person screws us over, the maximum they make is 19 give or take five*
Member # 12: *we need trust here*
Member # 10: *AGREED*
Member # 12: *or else this doesnt work*
Member # 10: *SO 10 EVERYTIME?*
Member # 11: *Lets do that*

The reasoning is straightforward and collectively rational, and the members proceeded to put all their tokens into the group account. Ten rounds in, and faced with a series of negative adjustments, they exchanged the following texts:

Before round 11:
Member # 12: *man we keep getting hit with those -5*
Member # 11: *THEY DID US DIRTY*
Member # 10: *the -5 started*
Member # 12: *its alright at least we are all doing what would maximize our profit ...*
Member # 12: *just keep trusting eachother and this will all continue to work*
Member # 10: *keep the 10 because thats how we each get max profits*
Member # 12: *yeah*

They then continued to maximize their group contributions despite additional negative adjustments. As can be seen from Fig. 1 and Table 4, the vast majority of players in Random-plus-communication sessions invested their full endowment in the group account until the second to last round.

As we will show in section 5.4 the chat among Unknown treatment players followed a similar pattern. And as might be expected given the standard game's structure and absence of adjustments, Baseline groups were especially focused on the collective benefits of maximizing contribution in the group account.

But for the Watchful-adjustment groups chat turned out to be a mixed blessing. Though it did raise rates of contribution, it also promoted false and counterproductive beliefs. This might seem to contradict the quasi-evolutionary theories of superstition embraced by Malinowski and other anthropologists of his era. But our lab-based "tribes" are extremely small, experience no evolutionary pressures or inter-tribal conflicts, and have just 20 rounds of play and 4 rounds of text chat in which to fine tune their inferences. Searching for patterns in random rewards and punishments led to inefficient strategies which some groups maintained despite poor results.

[19] Of the 48 players in Random-plus-chat sessions, none questioned the adjustments' randomness or 50/50 probability in their texts, and only one player raised the issue when asked for open-ended feedback in the post-game survey.

Communication moved Watchful treatment players away from the false but functional notion that the computer was rewarding “good” behavior toward the equally false but counterproductive notion that the computer rewarded specific numbers or sequences. Consider, for example, how group 1 in session 13 arbitrarily agreed to each invest 8 tokens after receiving several negative adjustments in the first 5 rounds:

Before round 6:
Member # 4: *new plan?*
Member # 1: *Strategy?*
Member # 1: *Put 7 or 8*
Member # 4: *lets put 8*
Member # 2: *ok perfect*
Member # 4: *everyone in?*
Member # 3: *yeah*

Despite the fact that the 50/50 nature of adjustments quickly disconfirmed any simple rule for controlling outcomes, some groups convinced themselves that they had found the magic number. Group 3 of session 14 thus reasoned as follows:

Before round 6:
Member # 11: *okay so the AI calculations is just flipping off every time so if we assume that it is giving the plus 5 to the two lowest teams then lets just put in 9.5 every time and we'll still get over 15 plus the 5 hopefully*
Member # 10: *agreed*
Member # 12: *9.5 each time now?*
Member # 11: *yes*
Member # 12: *sounds good*
Member # 9: *yup sounds good*

After then receiving several positive adjustments, member #11 proclaimed that he had indeed divined that the secret was to each play 9.5, and the other members agreed:

Before round 11:
Member # 11: *wow i'm smart*
Member # 10: *lets stick to this game plan*
Member # 9: *9.5 is perf*
Member # 12: *agreed*

Having locked in on this superstition, the group members largely stayed with 9.5 through the rest of the game despite receiving several negative adjustments.

Inefficient numerology also arose among groups in three additional sessions that we ran using a variant of the Watchful-plus-chat treatment in which the computer algorithm chose from five equi-probable adjustments {-5,-2,0,+2,+5} rather than just {-5,+5}. See Appendix C for details.

Here again, group chat reveals counterproductive quests for magical numbers rather than potentially helpful moral inferences that the computer is rewarding group-oriented behavior. For example, group 3 of session 23 decided first to try an ascending sequence of contributions (5,6,7,8,9):

Initial communication, before round 1:
...
Member # 12: *not sure... maybe we can do 5,6,7,8,9 and then re discuss after the first five rounds?*
Member # 9: *okay*
Member # 12: *everyone good with that?*
Member # 10: *ok*
Member # 11: *yes!*
Member # 9: *yes*

After receiving +5 adjustments when the all invested 5 and 7, they narrowed their plan accordingly but also experimented with 10.

Before round 6:
Member # 12: *okay maybe we should try like 5,7,5,7,10*
Member # 9: *yes 5 and 7 were the best...*
Member # 11: *so should we do the 5,7,5,7,10?*
Member # 10: *so 5 and 7 were best. yes im down for 5,7,5,7,10*
Member # 12: *so should we try 5,7,5,7 and then try 10 at the end?*
Member # 10: *wait why 10*
Member # 10: *just to try*

Table 6
Agreements and adherence

Treatment	4-player groups	Total agreements	Suboptimal agreements	Sub-agree adherence	Opt-agree adherence
Baseline-plus-chat	16	64	3/64 (4.8%)	57/60 (95.0%)	1063/1220 (87.1%)
Random-plus-chat	12	48	4/48 (8.3%)	53/80 (65.2%)	822/880 (93.4%)
Unknown-plus-chat	12	48	4/48 (8.3%)	55/80 (68.7%)	807/880 (91.7%)
Watchful-plus-chat	12	48	13/48 (27.1%)	164/260 (63.1%)	537/700 (76.7%)

Member # 9: *sounds good*
Member # 10: *good plan*

They then received two +5's and two -5's in the two rounds in which they all invested 5 and likewise in the two rounds when they invested 7. But a -5 followed the one round of 10's, so they turned their backs on the (socially optimal) 10 and focused on 5 and 7:

Before round 11:
Member # 12: *should we just do 7,7,7,7,7 lol*
Member # 9: *it seems like the adjustment is completely random. so annoying haha*
Member # 12: *the adjustment might also not be random...*
Member # 11: *im down to do anything maybe stick with 5 and 7*
Member # 10: *our best outcomes were with 7 and 5*
Member # 9: *agreed*

They then received a -5 after each invested 5 but a -2 after 7, so they settled on magic 7 for the rest of the game[20]:

Before round 16:
Member # 12: *i think maybe we should stick with 7's*
Member # 10: *7 is best*
Member # 9: *good with that*
Member # 10: *so 7 throughout*
Member # 12: *perfect*

5.4. Statistical analysis of group chat

In evaluating transcripts, we ourselves run the risk of seeing false patterns. We therefore subjected the chat text to a quantitative review, coding the agreements reached by each group and comparing them to members' subsequent contributions. Groups could communicate prior to rounds 1, 6, 11, and 16, and every group did some chatting in each period. Each chat concluded with an agreement that could be expressed as a contribution number, sequence, set, or range. In all, the chat record comprised 2,886 lines of text, spread over 208 separate chats by 52 four-member groups.

After puzzling over the contribution paths of Fig. 1, we had conjectured that communication led Watchful groups to choose suboptimal strategies more frequently than their counterparts in other treatments. The third column of Table 6 confirms that this was indeed true. Watchful groups made more than three times as many suboptimal agreements as their counterparts. But columns four and five point to a second source of reduced giving: Even when Watchful players committed to contributing fully, their *subsequent* contributions adhered less frequently to their jointly optimal agreements. The combined impact of suboptimal agreements and substandard adherence accounts for 94% of the giving gap.[21]

Fig. 2 provides visual perspective on all eight treatments. The top panels plot the period-by-period contributions of each group and overall contribution path for in each core treatment. Again, we see that the core Baseline groups start lower and decay faster than their adjustment counterparts. Compared to the Baseline, each of the three adjustment treatments displays higher mean paths across the rounds and greater variability within rounds. But communication shatters this three-way equivalence. The bottom panels show that the mean path of Watchful group contributions is consistently lower than the paths of the Random and Unknown treatments. Moreover, the Watchful path displays a much stronger *reset* effect pulling

[20] This example illustrates another potentially counterproductive feature of group chat. Absent communication, group members can reasonably assume they are unlikely to make identical decisions, especially since they can contribute in increments of 0.01 tokens. As noted in footnote 13, players may therefore assume that their individual choices will decisively influence the computer no more than 1/4th of the time. But chat enables them to coordinate their actions and thereby turns each into a potentially decisive actor if the computer rewards group-level numbers or patterns.

[21] The differences in agreements and adherence are both substantively large and statistically significant. Under rank sum tests, the rate of suboptimal agreement among Watchful groups differs from that of their Baseline counterparts at a 0.02 p-level and from their Random and their Unknown counterparts at 0.05 levels. In GLS regressions, the gap between Watchful player contributions and those of other treatments drops from 1.32 tokens per contribution to 0.09 tokens simply by adding a 0/1 dummy variable that codes for suboptimal group agreements and a second dummy that codes for player adherence to the group's most recent agreement.

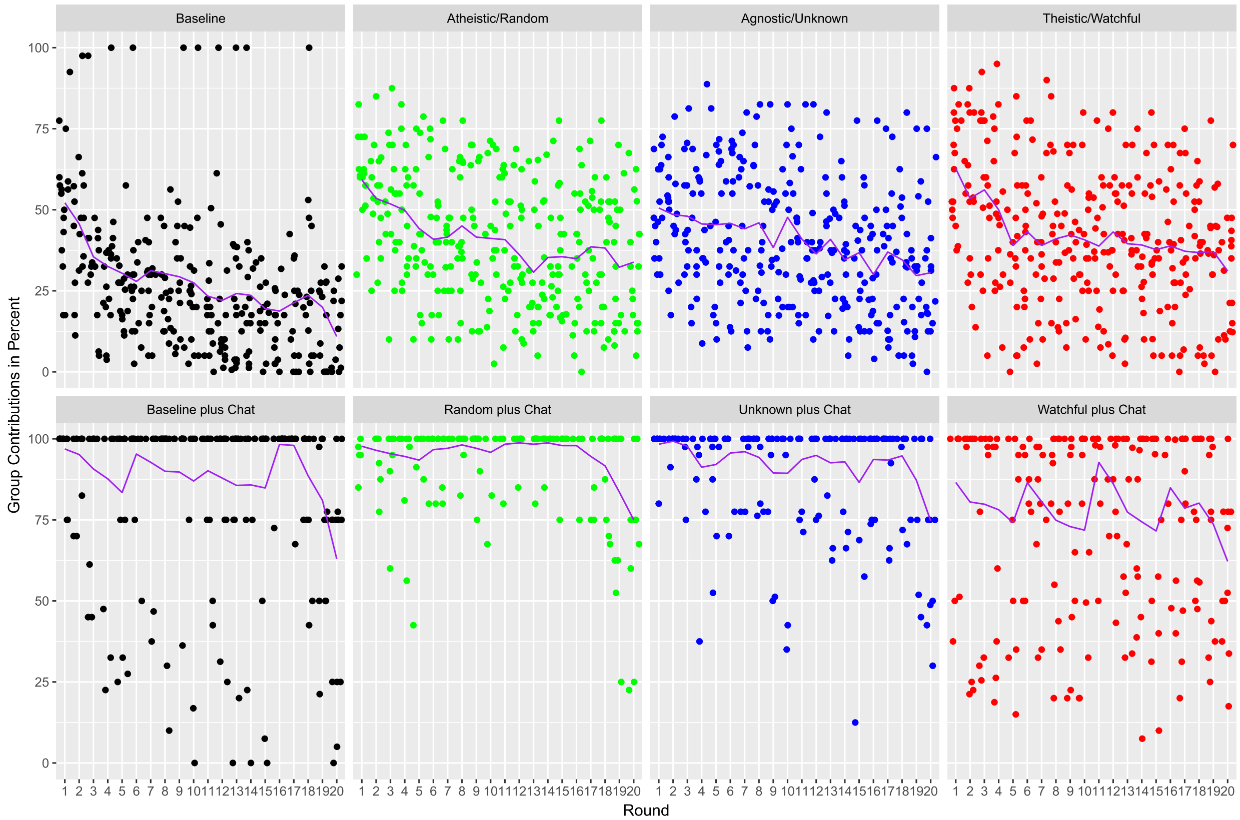

Fig. 2. Average and individual group contributions across all rounds

contributions back toward their agreed upon levels immediately after each chat but tailing off rapidly thereafter. And in all rounds, the variance in Watchful group contributions is larger than that of the Random or Unknown treatments. The net result is a visually striking spray of Watchful group contributions – with fewer groups making maximal contributions, more variation in group giving, and a stronger tendency to fall away from each agreement.

Except for the last few rounds, these patterns persist over time.[22] In other words, Watchful groups show little sign of losing faith. When a simple theory X fails, as it does with 50/50 odds in each round, the group may shift to other theories but rarely do they abandon all hope of influencing the computer. The quasi-magical quests for winning numbers persist even though most chats end in agreement to go all-in. And for every Watchful group that turns away from a suboptimal agreement, one or more other groups takes its place.[23] Individual contributions display a similar sort of churning. Frequent disappointment does not suffice to permanently disenchant the Watchful population. With instructions that provided little or no basis for faith, the players in other treatments make more productive use of their chats, forging more profitable agreements and maintaining greater solidarity.

6. Concluding remarks

We have described a novel exercise in pseudo-supernaturalism – a simple experiment that generates superstitions and channels exogenous shocks toward greatly increased cooperation. How players respond to their unpredictable environment depends on the narrative framing they inherit, with superstitions becoming more pronounced as the emphasis shifts from simple randomness, to unknown calculations, and finally to a Watchful computerized actor. As in studies of real-word

[22] The main features of Fig. 2 all stand up to statistical analysis, including the Watchful treatment's consistently greater variability in rates of optimal agreements and rates of actual contributions – between groups, within groups (over time), and between and within subjects. Regressions confirm that the Watchful treatment's distinctive features persist over time – the lower rates of optimal agreements, the lower rates of actual contributions, and the greater variability of both.

[23] Among the 12 Watchful groups, 7 different groups make at least one suboptimal agreement, with 2 groups making one such agreement, 2 making two, 3 making three, and none making four. Across the four chat periods, the sequence of agreements is 2, 3, 3, and 5. In contrast, among the 12 Random treatment groups (and also among the 12 Unknown groups) there are just 4 suboptimal agreements, with a single group making 3 of them, and a time sequence of 0, 2, 1, and 1.

supernaturalism, we find that social interactions are key to sustaining beliefs in the face of disconfirming evidence, including quasi-magical beliefs that reduce players' welfare.

Although we sought to explore the impact of laboratory worldviews analogous to atheism, agnosticism, and monotheism, our results to date look more like mechanistic superstition than moralistic faith. In retrospect, we suspect that most subjects played our experiment as they might a standard computer game, searching for hidden patterns rather than striving to win the computer's approval. Anthropologists and sociologists have often distinguished religion from magic based on whether the presumed interactions involve supernatural *beings* or impersonal supernatural *forces*. One might therefore saythat in seeking to simulate religion but ended up stimulating magic. The distinction is important, perhaps especially in economics, because it is religion rather than magic that routinely sustains collective action and moral imperatives.[24]

Moralistic religions versus mechanistic magic: A key question is whether small changes in the experiment might incline subjects to approach the computer as moralistic being rather than a mechanistic force. In Hajikhameneh and Iannaccone (2022), we report on follow-up experiments that truthfully tells players the adjustment takes one of several different forms that we describe in detail. The potential forms mirror different theological traditions, such as the moralistic God of the Abrahamic monotheism versus the capricious gods of the Greco-Roman pantheon. Some algorithms are purely random while others tend to reward higher absolute levels of group contribution or higher contributions relative to other groups. The results suggest that players routinely impute motives and propensities even when the underlying algorithm is random and do indeed gravitate toward interpretations that mirror the moralistic God of Judaism, Christianity, and Islam.

Cultural evolution: As noted in section 5, supernaturalism is more likely to prove functional where groups compete and beliefs evolve. We have begun to experiment with environments that offer more opportunities for adaptation, and we plan additional variations that incorporate different forms of group selection – through conversion (opportunities to switch groups), indoctrination (advising new players), proselytizing (advising members of other groups), competition (publicizing the earnings of other groups), and the like. Initial results from a first set of intergenerational treatments are reported in Hajikhameneh (2022).

Identity and mutual insurance: The contribution boost in the Random treatment merits special attention. The treatment differed from the standard public goods game only insofar as each round concluded with a positive or negative increment to earnings that neither masked nor magnified contributions. Players were clearly and correctly told that the adjustments were 50/50 random, yet their contributions averaged 50% more than their standard game counterparts. The contribution boost stood out in every period including those of the first period, which preceded any adjustments.

There are, of course, many potential explanations, ranging from subtle forms of experimenter demand to confusion and cognitive biases. In Hajikhameneh and Iannaccone (2022) we attempt to evaluate explanations in light of our results and those of other social dilemma experiments involving risk or uncertainty. In addition to superstition and supernaturalism, two other conjectures strike us as good candidates for further work.

The first is *common enemy* identity effect. Though a random number generator is a singularly bloodless opponent, the treatment did confront players with an electronic outsider that repeatedly meddled with their welfare – punishing the entire group 50% of the time while remaining immune to influence or persuasion. Few human traits have been demonstrated more frequently in lab settings than the tendency to adopt an "us versus them" outlook based on even the smallest and most arbitrary distinctions, including initial seating patterns, randomly distributed colors, and irrelevant preferences. In seeking to create a neutral alternative to our Watchful treatment, we may have stumbled onto cheap noise as a significant source of group solidarity.

Another possibility is a *mutual insurance* effect, a tendency to pull together in the face of risk. Food sharing among hunter-gatherers is clear example of risk-pooling that smooths individual consumption over time (Dyble et al., 2016). Sharing is less easily explained in our experiment, because all group members receive the same adjustments. Nevertheless, the historic value of risk-pooling has been so great, literally a matter of life and death, that the tendency may have evolved into a trait that readily arises in the face of risk. If so, it may well persist even when logic or direct experience implies its irrelevance.

Regardless of the underlying cause, the response to seemingly irrelevant risk merits careful study. In a world full of shocks and surprises, god game analogs arise at every turn.

Declaration of competing interest

None.

[24] Emile Durkheim, a founder of the sociology of religion, went so far as to distinguish religion from magic based on the former's collective character: "In all history we do not find a single religion without a Church...It is quite another matter with magic...The magician has a clientele and not a Church... A Church...is a moral community... But Magic lacks any such community" (Durkheim, 1965,[1912], pp. 59,44-45). As Hayek (1989), Norenzayan (2013) and others have emphasized, moral constraints and community are particularly central to the monotheistic traditions that appear to have played a major role in the emergence of large-scale society and economic development. For more on the contrasting character of religion and magic, see Iannaccone and Berman (2006, pp.11-12).

Data availability

Data will be made available on request.

Appendix A. Session dates and types

Date	Treatment	Session
2017-04-10	Baseline	1
2017-04-11	Baseline	2
2017-04-11	Random	1
2017-04-12	Baseline	3
2017-04-12	Random	2
2017-04-12	Random	3
2017-04-18	Baseline	4
2017-04-18	Watchful	1
2017-04-18	Watchful	2
2017-04-19	Random	4
2017-04-19	Watchful	3
2017-04-19	Watchful	4
2017-04-20	Watchful+chat	1
2017-04-20	Watchful+chat	2
2017-04-24	Baseline+chat	1
2017-04-24	Random+chat	1
2017-04-24	Random+chat	2
2017-04-26	Baseline+chat	2
2017-04-26	Watchful+chat	1
2017-04-28	Baseline+chat	3
2017-04-28	Random+chat	3
2017-05-01	5-value Watchful+chat	1
2017-05-03	5-value Watchful+chat	2
2017-05-04	5-value Watchful+chat	3
2017-05-11	Baseline+chat	4
2017-12-04	Unknown	1
2017-12-04	Unknown	2
2017-12-06	Unknown	3
2017-12-06	Unknown	4
2018-05-03	Unknown+chat	1
2018-05-04	Unknown+chat	2
2018-05-09	Unknown+chat	3
2022-04-20	mod. Watchful+chat	1
2022-04-20	mod. Watchful+chat	2
2022-04-21	mod. Watchful	1
2022-04-21	mod. Watchful+chat	3
2022-05-10	mod. Watchful	2
2022-05-11	mod. Watchful	3

Appendix B. Experiment instructions

This appendix documents the instructions used in the core and communication treatments' instructions. The core treatments' instructions mainly differ in the description of the adjustment algorithm.[25] Hence, **Stage 2**'s narration is the main variation across treatments.

B.1. Core treatment instructions

Instructions: This experiment is a study of group and individual behavior. Everyone in this experiment is receiving exactly the same instructions. You will earn $7 simply for participating in this experiment. But if you follow the instructions and make careful decisions, you can earn a significant amount of additional money. You will be paid at the end of the experiment, privately and in cash.

During the experiment your earnings will be measured in tokens. Each token is worth 5 cents, so at the end of the experiment your tokens will be converted to dollars at the rate of 20 tokens per dollar.

[25] Since we describe the adjustment algorithm differently across treatments, the earning stage's instruction and examples given at the end of the instruction are modified accordingly. Note that there is no adjustment and therefore no Stage 2 in the Baseline treatment.

It is important that you remain silent and not look at other people's work during the experiment. If you talk, laugh, or exclaim out loud you will be asked to leave and will not be paid. If you have questions or need assistance, raise your hand and an experimenter will come to you.

Your group: Each participant will be **randomly** assigned an ID number and randomly assigned to a four-person group. The groups will remain the same throughout the entire experiment. There are 16 participants in today's experiment, so there will be 4 groups in all with 4 members in each group. Because you will be known only by your random ID number, there will be no way for anyone (including experimenters) to know which group you are in or what decisions you make.

Your decisions: The experiment includes 20 decision-making periods, and each period consists of 2 stages. At the start of each period you will receive 10 new tokens.

Stage 1: In the first stage of each period, you must decide how many of your 10 new tokens you want to invest in your group's joint account versus how many you want to invest in your own personal account.

The tokens you invest in your personal account will be directly added to your earnings. (For example, if you invest 4 tokens in your personal account then you will earn 4 tokens from your personal account.)

Earnings from your group's joint account will depend on the total number of tokens that you and the other members of your group invest in the joint account. Each member will earn 0.4 times the total number of tokens invested in the joint account. (For example, if each member of your group invested 6 tokens in the joint account, then the total joint investment would be 24 tokens, and each member would earn 0.4 x 24 tokens, which equals 9.6 tokens.)

To make your investment decisions, simply enter a number between 0.00 and 10.00 indicating the amount you want to invest in your group's joint account. The computer will automatically invest the rest of your 10 tokens in your personal account. (For example, if you enter the number 6.0, then 6 of your tokens will go into your joint account and the computer will put your remaining 4 tokens into your personal account.)

Stage 2 *[Random treatment]*: In the second stage of each period, the computer will make adjustments to the earnings of all the participants in the experiment. The adjustments will be either +5 or -5 tokens. The same adjustment will be made to the earnings of every member in your group, but the members of other groups may receive different adjustments. The computer will choose the adjustment for your group based on a random number calculation, so that there will be a 50% chance that 5 tokens are added to the earnings of everyone in your group and a 50% chance that 5 tokens are subtracted from every member's earnings. The computer will calculate a different random number for each of the other three groups in the experiment.

Stage 2 *[Watchful treatment]*: In the second stage of each period, a computer-based Artificial Intelligence (or "AI") will make an adjustment to your earnings. The adjustment will be either +5 or -5. The same adjustment will be made to the earnings of each member in your group. The AI will choose the adjustment for your group based on a complex series of calculations that take account of your investment decisions, the investment decisions of the other members in your group, and the investment decisions of the other groups in this experiment. The AI will do separate calculations and make separate adjustment decisions for the other three groups.

Stage 2 *[Unknown treatment]*: After the members of your group have made their investment decisions, the computer will make an adjustment to your earnings. The adjustment will be either +5 or -5 tokens. The same adjustment will be made to the earnings of every member in your group, but the members of other groups may receive different adjustments. The computer will choose the adjustment for your group based on a complex sequence of calculations. The computer will choose a separate adjustment for each of the other three groups based on a separate sequence of calculations.

Your earnings: At the end of each period the computer will show you how much you invested in each account, the total amount that was invested in your group's joint account, and how much you earned from each account. You will not see how each of the other members in your group divided their tokens between their own personal accounts and the group's joint account.

Your overall earnings in each period will be your earnings from the joint account *plus* your earnings from your personal account, *plus* the computer's adjustment. *[The Baseline instructions omit "plus the computer's adjustment."]* Fig. B1 provides an example of the report that each subject receives at the end of each period.

Examples:

Note: The following examples appear in the Random and Unknown treatment instructions. The Watchful treatment instructions substitute the word "AI" for "computer," and the Baseline instructions omit all references to adjustments.

Table 1 shows what happens if you invest 6 tokens in your group's joint account, while the second member of your group invests 3 tokens, the third member invests 1, and the fourth invests 10, and the computer randomly makes a +5 adjustment for your entire group. The total number of tokens invested in the joint account is 6+3+1+10 = 20, and $0.4\times20 = 8$. So each member earns 8 tokens from the joint account. You yourself earn an additional (10-6) = 4 tokens from your personal account, the second member earns an additional 7 tokens, the third member earns an additional 9 tokens, and the fourth member earns an additional 0 tokens. The computer's randomly adjustment adds another 5 token to each member's earnings. So you earn 17 tokens overall, 8 from the group's account, 4 from your personal account, and 5 from the adjustment. As you can see in the bottom row of the table, the other members earn 20, 22, and 13.

Table 1

	You	Mbr 2	Mbr 3	Mbr 4	Explanations
Investment in joint acct:	6	3	1	10	$TOTAL = 6 + 3 + 1 + 10 = 20$
Investment in personal acct:	4	7	9	0	10− Investment in joint account
Earnings from joint acct:	8	8	8	8	Each member earns 0.4 $\times TOTAL$
Earnings from personal acct:	4	7	9	0	Same as personal acct. investment
Adjustment:	+5	+5	+5	+5	Computer chooses adjustment = +5
Overall earnings for period:	17	20	22	13	= Earnings from both accounts plus computer's adjustment

Table 2 shows what happens if you invest 3.50 tokens in your group's account, while the second member of invests 6.50 tokens in the group account, the third invests 0, and the fourth invests 9, and the computer randomly makes a -5 adjustment for your entire group. The total amount invested in the group's account is 3.5+6.5+0+9 = 19, and 0.4×19 = 7.6. So each member earns 7.6 tokens from the group account. You yourself earn an additional 6.5 tokens from your personal account. So you earn 9.1 tokens in all, 7.6 from the group's account, 6.5 from your personal account, and -5 from the random adjustment. The other members earn 6.1, 12.6, and 3.6.

Table 2

	You	Mbr 2	Mbr 3	Mbr 4	Explanations
Investment in joint acct:	3.5	6.5	0	9	$TOTAL = 3.5 + 6.5 + 0 + 9 = 19$
Investment in personal acct:	6.5	3.5	10	1	10− Investment in joint account
Earnings from joint acct:	7.6	7.6	7.6	7.6	Each member earns 0.4 $\times TOTAL$
Earnings from personal acct:	6.5	3.5	10	1	Same as personal acct. investment
Adjustment:	-5	-5	-5	-5	Computer chooses adjustment = -5
Overall earnings for period:	9.1	6.1	12.6	3.6	= Earnings from both accounts plus computer's adjustment

Table 3 shows what happens if you invest all 10 of your tokens in your group's account, while the other three members do the same, and the computer randomly makes a +5 adjustment. Because the total group investment is 40 and 0.4×40 = 16, you earn 21 overall, and the other members earn the same.

Table 3

	You	Mbr 2	Mbr 3	Mbr 4	Explanations
Investment in joint acct:	10	10	10	10	$TOTAL = 10 + 10 + 10 + 10 = 40$
Investment in personal acct:	0	0	0	0	10− Investment in joint account
Earnings from joint acct:	16	16	16	16	Each member earns 0.4 $\times TOTAL$
Earnings from personal acct:	0	0	0	0	Same as personal acct. investment
Adjustment:	+5	+5	+5	+5	Computer chooses adjustment = +5
Overall earnings for period:	21	21	21	21	= Earnings from both accounts plus computer's adjustment

B.2. Communication treatment instructions

Note: The communication treatments added the following instructions immediately before the "Your Earnings" section. They were otherwise identical to their core treatment counterparts.

Group chat: You will have four opportunities to communicate by text with the other members of your group. The opportunities will occur just before period 1 and immediately after periods 5, 10, and 15. At these times, you and the other members of your group will have two minutes to type whatever you wish, and your screen will list what each group member types. You will not see anything typed by the members of the other groups.

When the chat screen opens, you'll be able to type in the blue field at the bottom of the screen. Do not type anything that identifies who you are or where you're sitting. When you hit $< enter >$, your member number and text will appear in the large white chat area. The text entries of other members will appear there as well. So, for example, if you were group member 2 and you typed *“Hello group. Got any suggestions? < enter >”* the chat box of every member in your group would display:

Member Number 2: Hello group. Got any suggestions?

If member 4 then typed *“Pay attention to what happens. < enter >”* and member 1 typed *“And make money! < enter >”* then each member’s chat box would display:

Member Number 2: Hello group! Got any suggestions?
Member Number 4: Pay attention to what happens.
Member Number 1: And make money!

Your chat box will remain open until you click on the red “Exit Chat” button at the bottom of the screen. Every participant must click “Exit Chat” before the experiment moves on to the next period.

B.3. Earnings reports

Fig. B1 shows the screenshot text from a typical end-of-period earnings report.

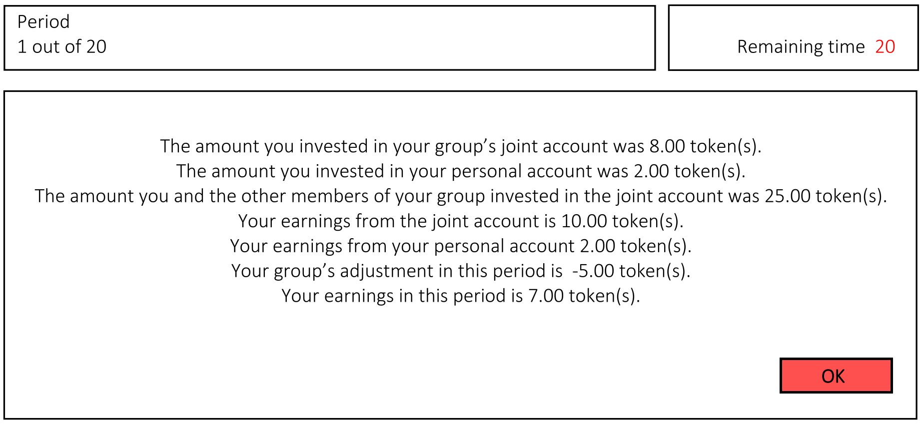

Fig. B1. Typical end-of-period earnings report

B.4. Risk preference task instructions

Note: Subjects completed the following risk elicitation task on their computers. The task is based on the method of Holt and Laury (2002) and includes two sequences of paired lotteries, A and B. In each A-B pairing, the sequence-A lottery offers $1 versus $3 with constant 1/2 probabilities. Playing lottery A is analogous to playing the zero-contribution Nash strategy in the AI and Random treatments, which yield total earnings each round of $1 (= 10 − 5 tokens) versus $3 (= 10 + 5 tokens) with 1/2 probabilities. The sequence-B lotteries offer prizes of $0.10 versus $4 with the probability of $4 rising in steps of 1/10 from 1/10 to 10/10. The number of times that a subject chooses B over A provides a measure of risk aversion.

In the questions that follow, you are going to be asked to make ten decisions. Each decision will be between Option A and Option B. One of the ten choices you make will be randomly selected to determine your earnings for this part of the experiment.

Options		Your Choice
A	B	
$1 or $3 each with probability 1/2	$0.1 with probability 9/10 or $4 with probability 1/10	A or B
$1 or $3 each with probability 1/2	$0.1 with probability 8/10 or $4 with probability 2/10	A or B
$1 or $3 each with probability 1/2	$0.1 with probability 7/10 or $4 with probability 3/10	A or B
$1 or $3 each with probability 1/2	$0.1 with probability 6/10 or $4 with probability 4/10	A or B
$1 or $3 each with probability 1/2	$0.1 with probability 5/10 or $4 with probability 5/10	A or B
$1 or $3 each with probability 1/2	$0.1 with probability 4/10 or $4 with probability 6/10	A or B
$1 or $3 each with probability 1/2	$0.1 with probability 3/10 or $4 with probability 7/10	A or B
$1 or $3 each with probability 1/2	$0.1 with probability 2/10 or $4 with probability 8/10	A or B
$1 or $3 each with probability 1/2	$0.1 with probability 1/10 or $4 with probability 9/10	A or B
$1 or $3 each with probability 1/2	$0.1 with probability 0/10 or $4 with probability 10/10	A or B

B.5. Questionnaire

Note: The following questions were asked of all subjects (except the questions in bold, which only were asked only in the AI treatments). Response categories follow the question in italics or in itemized lists.

- What is your major?
  - Math, Engineering, or the Physical Sciences
  - Business or Economics
  - English, Foreign Languages, or Classics
  - Humanities
  - Other
- What is your Gender? *[Male, Female]*
- **Did you notice any patterns in the AI's adjustments** *[Yes, No]*
- **If you think you noticed any patterns in the AI's adjustments, please describe them briefly. (Type "no pattern" if you did NOT notice any patterns.)**
- How well do the following statements fit your own decision-making style?
  - When making important decisions I focus on facts and logic *[Always, Never]*
  - When making important decisions I trust my feelings and intuition *[Always, Never]*
  - When making important decisions I consult with religious or spiritual leaders *[Always, Never]*
- How strongly do you agree or disagree with the following statements?
  - Some places really are haunted by spirits *[Strongly agree, Agree, Disagree, Strongly Disagree]*
  - Some people can use the power of their minds to heal other people *[Strongly agree, Agree, Disagree, Strongly Disagree]*
  - Some people can use the power of their minds to "see" into the future. *[Strongly agree, Agree, Disagree, Strongly Disagree]*
  - Would you describe yourself as a "spiritual" person? *[Very spiritual, Somewhat spiritual, Not at all spiritual]*
  - Would you describe yourself as a "religious" person? *[Very religious, Somewhat religious, Not at all religious]*
- Which of the following best describes your current religion? *[Christian, Jewish, Muslim, Buddhist, Hindu, Other, No religion]*
- In the past year, about how often have you attended religious services?
  - Almost every week.
  - About two or three times per month.
  - About once each month.
  - Several times each year.
  - Once or twice each year
  - Never or almost never.
- When you were growing up, around age 11 or 12, how often did you usually attend religious services? *[Same response categories as previous question]*
- The instructions for the experiment were clear and easy to follow *[Strongly agree, Agree, Disagree, Strongly disagree]*
- Thank you for completing this experiment. We value your feedback, so please use the following text box for comments or suggestions.

Appendix C. Adjustments with five values

Surprised by the nearly equal rates of contribution across treatments, we speculated that two adjustment outcomes might have made the random nature of the adjustment process too easy to spot. Perhaps more outcomes would induce stronger superstitions and less similar behavior across treatments. We therefore added three sessions that employed a modified version the Watchful-plus-chat treatment, enlarging the set of adjustments from two equally probable values {-5,+5} to five {-5,-2,0,+2,+5}. All other features of the experiment remained unchanged, including group size and number of subjects per session.

The five-value results were, however, almost identical to their two-value counterparts. Earnings averaged $24.25 in the 5-value sessions versus $24.18 in the 2-value sessions. Group contributions averaged 32.56 tokens versus 31.25, and medians averaged 40 versus 38.75. Neither means nor medians of the two treatments differed significantly under the rank sum comparison test (with p-values of 0.73 and 0.57, respectively). And as in the 2-value sessions, the mean and median group contributions in the 5-value sessions remained significantly lower than those of the two-value Random treatment (with p-values 0.03 for both means and medians). As seen in Fig. E1, even the two treatments' period-by-period plots of individual group contributions and their period-by-period paths of total contributions are quite similar to each other and much less like those of the Random-plus-chat and Unknown-plus-chat treatments.

Given that the shift from {-5,+5} to {-5,-2,0,+2,+5} made inference more difficult while simultaneously reducing risk, it is possible that these effects just happened to cancel each other. But such a coincidence strikes us as unlikely, especially given the weak and statistically insignificant responses to risk found in the regressions of Tables 5 and F2.

Appendix D. The adjustment algorithm

D.1. Original adjustment algorithm

In the second stage of each round t, the computer adjusts player earnings by adding either a high or low value, $H = +5$ or $L = -5$. Adjustments may vary across groups but are equal within groups. Groups with the highest contributions are ranked 1, and all other groups are ranked 2:

$$Rank_{jt} \equiv \begin{cases} 1 \;\; iff \;\; G_{jt} = max\{G_{.t}\} \\ 2 \;\; otherwise \end{cases} \tag{D1}$$

Groups also receive scores between 0 and 1 based on their contribution G_{jt} and a random number u_{jt} distributed uniformly over the unit interval:

$$Score_{jt} \equiv (G_{jt} + u_{jt}) - \lfloor G_{jt} + u_{jt} \rfloor \tag{D2}$$

The groups' adjustments are then based on their ranks and scores:

$$Adjustment_{jt} \equiv \begin{cases} H \; if \; (Rank_{jt} = 1 \; and \; Score_{jt} \geq \alpha) \; or \; (Rank_{jt} = 2 \; and \; Score_{jt} \leq \beta) \\ L \;\; otherwise \end{cases} \tag{D3}$$

The net effect is that groups with the highest contributions receive the high-value adjustment with probability $(1 - \alpha)$ **and the other groups receive the high-value adjustment with probability** β**.** For $(1 - \alpha) > \beta$, the computer operates more and more like a "god" that rewards generosity or commitment to the group as $(1 - \alpha)$ increases relative to β. For $\beta > (1 - \alpha)$, the computer operates more and more like a "god" that rewards selfishness or personal commitment as $1 - \alpha$ increases relative to β. The "god's" conduct becomes more and more nearly capricious (i.e., random with respect to group commitment) as $(1 - \alpha)$ approaches β. **To keep the focus on Malinowski's theory of supernaturalism, Skinnerian superstition, and the impact of socialization, we set** $(1 - \alpha) = \beta = 0.5$**.**

D.2. The modified adjustment algorithm

Groups were ranked based on their group contributions (G_{jt}): the group with the highest contributions is ranked 1, the group with the second most contributions is ranked 2, and so on. Any ties were randomly broken. As in the original adjustment algorithm, groups also received $Score_{jt}$ between 0 and 1 based on their contribution G_{jt} and a random number u_{jt} uniformly distributed over the unit interval:

$$Score_{jt} \equiv (G_{jt} + u_{jt}) - \lfloor G_{jt} + u_{jt} \rfloor \tag{D4}$$

The group adjustment of $H = 5$ or $L = -5$ was then determined based on the following algorithm:

$$Adjustment_{jt} \equiv \begin{cases} H \; if \; Score_{jt} \geq \alpha_{jrt} \\ L \;\; otherwise \end{cases} \tag{D5}$$

where $(1 - \alpha_{jrt})$ was the rank-based probability of receiving the H adjustment. The group with the highest contribution rank received the high (+5 token) adjustment with probability 0.52, the 2nd group with probability 0.51, the 3rd with p = 0.49, and the lowest-ranked group with p = 0.48.

Appendix E. Results from the modified Watchful treatment

Although experimental economists share a commitment to avoiding deception, there is as Ortmann (2019, p.29) states "considerable variance in what economists define as deception." There are, in particular, "gray areas with respect to practices that omit information or are misleading without an explicit lie being told" (Charness et al., 2022, p.385). The instructions for our original Watchful treatment fell within this gray area. Subjects were truthfully told that the in each period the computer would choose a ± 5 "adjustment for your group based on a complex series of calculations that take account of your investment decisions, the investment decisions of the other members in your group, and the investment decisions of the other groups," but they were *not* told that algorithm's parameter values ensured that the probability of a $+5$ adjustment equaled 0.50.

After consulting with colleagues and receiving reviewers' comments, we decided to run additional Watchful treatment sessions that employed the modified adjustment algorithm described in Appendix D.2 – three without communication and three with communication. The modified algorithm ensured that players now received positive adjustments with probability 0.52, 0.51, 0.49 or 0.48 depending on whether their group's total contribution in the current period ranked 1st, 2nd, 3rd, or 4th relative to those of the other groups. In all other respects, these modified Watchful treatments were identical to their original Watchful counterparts: same instructions, same group size, same session sizes, same earnings function, and same university lab.

The participants in the original and modified sessions came from distinct student cohorts, separated by five years and the Covid crisis. Knowing that results in public goods experiments can be sensitive to small changes in design, conditions, or expectations (Chaudhuri, 2018; Cookson, 2000), and knowing also that some experimentalists have noted significant differences in subject behavior pre- versus post-Covid, we restricted our analyses in the body of the paper to our original sessions. As detailed below, all results from the non-chat sessions were virtually unchanged. Contributions in the chat sessions increased, but all other results continued to support the conclusion that the Watchful instructions led to greater levels of superstition than those of the Random or Unknown treatments.

E.1. Core treatment comparisons

Average contributions: Table E1 reports contribution summary statistics for all five non-communication treatments. As in the original Watchful treatment, average contributions in the modified Watchful treatment greatly exceed those of the Baseline treatment; the 49% difference has a p-value of 0.00. Moreover, original versus modified Watchful contributions differ by just 3% (p-value = 0.77) and trace out very similar paths across the 20 periods.

Table E1
Summary statistics - original versus modified core treatments

	Group Contributions averaged over all rounds			
Treatments	Mean	SE	Median	N
Baseline	10.99	0.539	9.25	320
Random	16.50	0.445	16.00	320
Unknown	16.23	0.448	15.22	320
Watchful	16.94	0.483	16.25	320
Modified Watchful	16.41	0.463	16.00	240

Superstition and inertia: Predictions 5a, 5b, and 6, stated that superstitious responses to the adjustment process and instructions would induce contribution inertia (i.e., less change and more repetition relative to negative adjustments) following positive adjustments, and this inertia would be greatest in the Watchful treatment. Table E2 extends Table 3 by adding the modified Watchful treatment to the regression analysis, and here again the modified Watchful results mirror those of the original Watchful treatments.

Table E2
GLS analysis of contribution inertia

	(1) Absolute Change	(2) No Change (all cases)	(3) No Change (0's & 10's omitted)
Random	-0.136	0.016	-0.001
	(0.221)	(0.046)	(0.045)
Unknown	-0.087	-0.036	0.016
	(0.217)	(0.041)	(0.043)
Modified Watchful	0.430*	-0.058	-0.029
	(0.226)	(0.054)	(0.048)
Positive Adjustment$_{t-1}$	-0.849***	0.162***	0.170***
	(0.127)	(0.033)	(0.033)
Random × Positive Adjustment$_{t-1}$	0.337**	-0.082**	-0.091**
	(0.169)	(0.036)	(0.040)
Unknown × Positive Adjustment$_{t-1}$	0.241	-0.048	-0.074*
	(0.181)	(0.040)	(0.041)
Modified Watchful × Positive Adjustment$_{t-1}$	0.006	-0.026	-0.038
	(0.229)	(0.042)	(0.043)
Intercept	2.187***	0.302***	0.193***
	(0.137)	(0.034)	(0.035)
Observations	4560	4560	3993
R^2	0.03	0.02	0.02

Standard errors clustered by group in parentheses.
* $p<0.1$, ** $p<0.05$, *** $p<0.01$.
The Watchful treatment is the omitted category.

Recall that the first measure of inertia (*AbsChange*) equals the absolute change in each player's contribution from the previous round to the current round. The second measure (*NoChange*) equals 1 if contributions in the current and preceding round were equal and 0 otherwise. The original Watchful treatment is the omitted category. Hence, the "Intercept" row reports average contribution inertia among the original Watchful treatment players when their groups received a negative adjustment, and the "Random," "Unknown," and "Modified Watchful" rows report how much average inertia after negative adjustments differed from that of the original Watchful treatment. The "Positive Adjustment$_{t-1}$" row shows that for all three measures, contribution inertia is substantially and significantly greater following positive adjustments in the original Watchful treatment (e.g., the average absolute change in contributions (column 1) drops by 0.849 from 2.187 to 1.338, and the rate of repeat contributions (in column 2) increases from 30.2% to 46.4%. The "Modified Watchful x Positive Adjustment$_{t-1}$" row shows that for all three measures of inertial the (presumably superstitious) response to positive versus negative adjustments is virtually identical in the modified versus original Watchful treatments. The "Modified Watchful" row shows that inertia effects following *negative* adjustments were smaller in the modified treatment, but this only means that the *relative* effect of a positive adjustment is slightly smaller (e.g., whereas a positive adjustment reduced the Absolute Change in contributions by 39% from 2.187 to 1.338 in the original Watchful treatment, the reduction in the modified Watchful treatment is 32% from 2.617 to 1.774).

E.2. Chat treatment comparisons

Average contributions: Turning to the chat treatments, Table E3 reports contribution summary statistics for all five communication treatments. Note that mean contributions in the modified Watchful sessions exceed those of the mean contributions in the original Watchful treatments. Although this difference is not statistically significant (the two-sided Wilcoxon rank sum test $p - value = 0.14$), and the modified Watchful contributions are still less than the Random and Unknown treatment contributions, overall contributions in the modified Watchful treatment nearly equal those of the Random and Unknown treatments. In other words, the modified Watchful treatment does not display the clear evidence of *collectively counterproductive* superstitions seen in the original Watchful treatment.

Table E3
Summary statistics – original versus modified chat treatments

	Group contribution in the join account			
Treatments	Mean	SE	Median	N
Baseline-plus-chat	35.44	0.551	40.00	320
Random-plus-chat	37.94	0.332	40.00	240
Unknown-plus-chat	36.95	0.406	40.00	240
Watchful-plus-chat	31.35	0.713	38.75	240
Modified Watchful-plus-chat	35.88	0.492	40.00	240

Note, however, that group contributions in the 5-value Watchful-plus-chat treatment discussed in Appendix C averaged 32.6/40 tokens - an amount ths is neither substantively nor significantly greater than the 31.6/40 tokens in the original Watchful-plus-chat treatment (with a rank-sum comparison p-value of 0.73) but both substantively and significantly less than the Random treatment (p-value 0.03). Getting such similar contributions across two treatments with substantially different adjustment algorithms and correspondingly different instructions suggests that the Watchful framing did indeed induce more counterproductive superstitions than the Random framing.

Suboptimal agreements: Our content analyses clearly showed that the subjects in the modified Watchful communication treatment engaged in the same sort of counterproductive searches for specific (quasi-magical) numbers that we observed in our original treatments. Table E4 shows that 11 of the 48 modified Watchful chat sessions (22.9%) ended in suboptimal agreements - a rate not significantly less than the original Watchful treatment's 13-of-48 (27.1%) but far greater than the Random and Unknown treatments' 4-of-48 each and Baseline treatment's 3-of-64.

Table E4
Agreements and adherence including the modified Watchful

Treatment	4-player groups	Total agreements	Suboptimal agreements	Sub-agree adherence	Opt-agree adherence
Baseline-plus-chat	16	64	3/64 (4.8%)	57/60 (95.0%)	1063/1220 (87.1%)
Random-plus-chat	12	48	4/48 (8.3%)	53/80 (65.2%)	822/880 (93.4%)
Unknown-plus-chat	12	48	4/48 (8.3%)	55/80 (68.7%)	807/880 (91.7%)
Watchful-plus-chat	12	48	13/48 (27.1%)	164/260 (63.1%)	537/700 (76.7%)
Modified Watchful-plus-chat	12	48	11/48 (22.9%)	167/220 (75.9%)	676/740 (91.4%)

Adherence to agreements: The higher overall contributions in the modified treatment are largely due to four facts: (1) the bulk of all agreements were optimal in both treatments, (2) optimal agreements were slightly more common in the modified treatment, (3) modified treatment players adhered to their optimal agreements far more often than their original treatment counterparts, and (4) every failure to fulfill an optimal agreement reduced overall contributions, potentially by as much as 10 tokens. Modified treatment players were also more likely to adhere to their suboptimal agreements, but this had far less impact on total contributions because relatively few agreements were suboptimal and failure to fulfill a suboptimal agreement could either raise or lower overall contributions.

Further evidence of superstition: The two sets of Watchful treatments displayed additional similarities pointing to greater rates of superstition relative to the Random, Unknown, and Baseline treatments. For example, the words "pattern" or "patterns" appeared 16 times in the modified Watchful chat and 5 times in the original Watchful chat but just once in the Unknown chat and never in the Random chat. Also, as can be seen in Fig. E1 the scatter plots of average group contributions in both original and modified Watchful-plus-chat treatments displayed similar patterns of dispersion (with a substantial number of group contributions below 60% of maximum), whereas the dispersion and number of low contributions was smaller for in Unknown-plus-chat treatment and smaller still in the Random-plus-chat treatment.

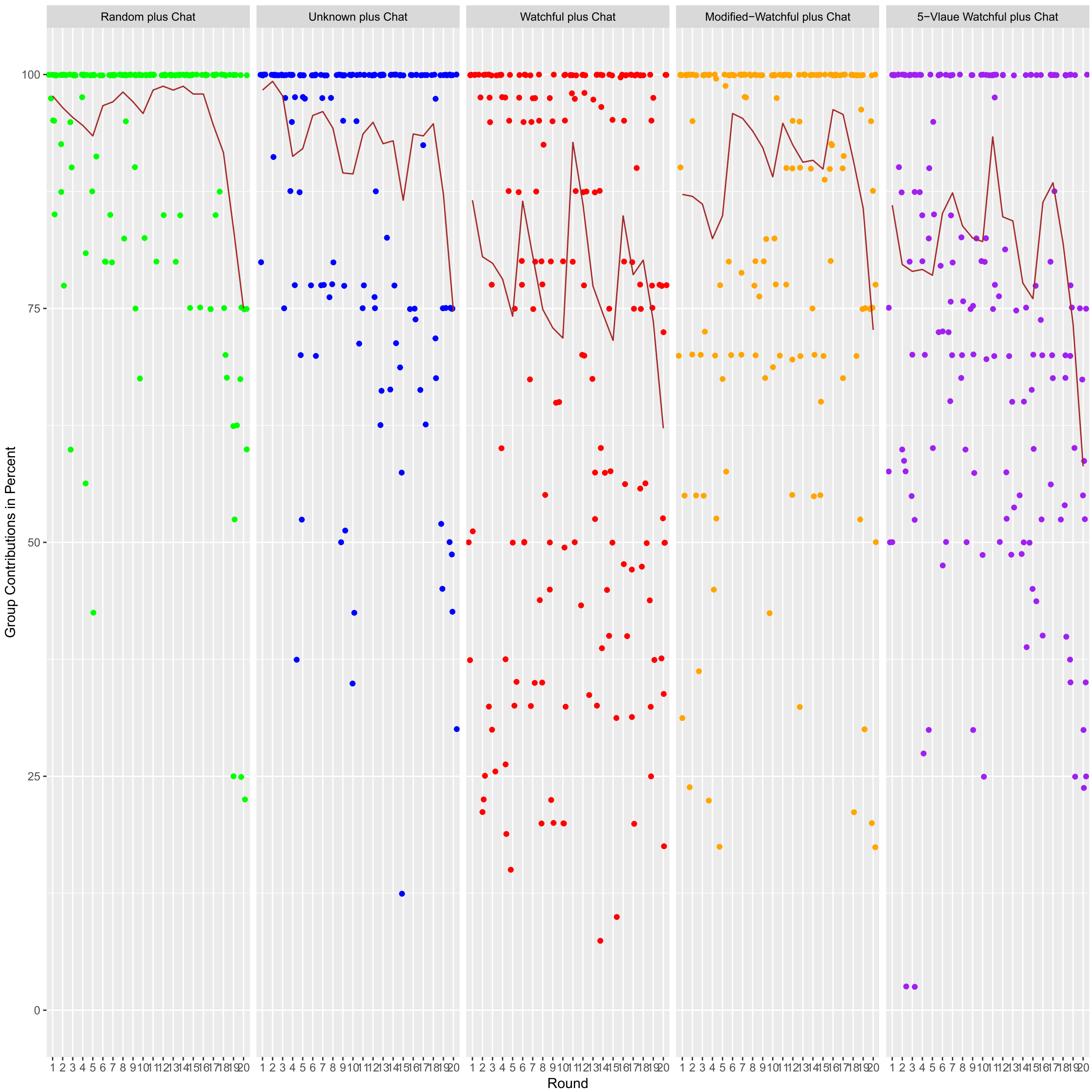

Fig. E1. Average and individual group contributions across all rounds

Appendix F. Tobit and hurdle model regressions

As noted in the text, we supplemented our GLS regression with Tobits because contributions must fall within the zero-to-ten token range, and we bootstrapped the Tobit standard errors and included the same random effects as in GLS. See Tables F1 and F2 below, and note that the parameter values and significance largely follow those of their GLS counterparts in Tables 2 and 5.

Table F1
Tobit analysis of contributions - core treatments

	(1) $Contribution_t$	(2) $Contribution_{t>1}$	(3) $Contribution_{t>1}$	(4) $Contribution_{t>1}$
Random	-0.133	-0.121	0.592	0.731
	(0.569)	(0.576)	(1.280)	(1.296)
Unknown	-0.164	-0.089	0.688	0.855
	(0.464)	(0.486)	(1.378)	(1.398)
Positive $Adjustment_{t-1}$		0.211**	0.208**	0.410**
		(0.104)	(0.103)	(0.188)
Risk			0.414	0.437
			(2.297)	(2.244)
Random × Risk			-1.642	-1.661
			(2.901)	(2.837)
Unknown × Risk			-1.724	-1.748
			(2.991)	(3.045)
Female			-0.881**	-0.878*
			(0.431)	(0.481)
Round			-0.124***	-0.125***
			(0.015)	(0.015)
Random × Positive $Adjustment_{t-1}$				-0.272
				(0.252)
Unknown × Positive $Adjustment_{t-1}$				-0.332
				(0.253)
Intercept	4.118***	3.870***	5.577***	5.476***
	(0.370)	(0.354)	(1.048)	(1.091)
Observations	3840	3648	3648	3648
Wald Chi-Sq	0.127	4.291	70.57	80.18

Bootstrapped standard errors in parentheses.
* $p<0.1$, ** $p<0.05$, *** $p<0.01$.
The Watchful treatment is the omitted category.

Table F2
Tobit analysis of contributions - chat treatments

	(1) $Contribution_t$	(2) $Contribution_{t>1}$	(3) $Contribution_{t>1}$	(4) $Contribution_{t>1}$
Random plus Chat	7.751***	7.763***	6.318*	6.526*
	(1.659)	(1.645)	(3.274)	(3.330)
Unknown plus Chat	8.413***	8.314***	4.908	5.322
	(1.900)	(1.820)	(3.768)	(3.745)
Positive $Adjustment_{t-1}$		0.279	0.456	0.798
		(0.438)	(0.412)	(0.595)
Risk			-7.577	-7.615
			(4.811)	(4.680)
Random plus Chat × Risk			3.523	3.580
			(7.311)	(7.113)
Unknown plus Chat × Risk			8.294	8.304
			(7.065)	(6.886)
Female			-0.563	-0.564
			(1.566)	(1.503)
Round			-0.212**	-0.211***
			(0.082)	(0.079)
Random plus Chat × Positive $Adjustment_{t-1}$				-0.514
				(1.133)
Unknown plus Chat × Positive $Adjustment_{t-1}$				-0.879
				(1.104)
Intercept	13.151***	12.802***	18.400***	18.258***
	(1.295)	(1.301)	(3.021)	(3.015)
Observations	2880	2736	2736	2736
Wald Chi-Sq.	27.63	28.68	44.74	52.97

Bootstrapped standard errors in parentheses.
* $p<0.1$, ** $p<0.05$, *** $p<0.01$.
The Watchful plus Chat treatment is the omitted category.

Because decisions to fully free-ride may reflect a qualitatively different view of the game, we also ran regressions that employed the double-hurdle estimation method developed by Cragg (1971) and applied to public goods game data by Engel and Moffat (2012) and Cheung (2014). See Table F3 below, and note that lagged adjustments remain a significant predictor of positive contribution decisions in the Watchful treatment.

Table F3
Hurdle analysis of contributions

	(1) Contribution$_{t>1}$	(2) Contribution$_{t>1}$	(3) Contribution$_{t>1}$
Main Model (Positive Contribution)			
Unknown	-0.142	0.046	0 .121
	(0.690)	(0.717)	(0.639)
Random	0.006	0.085	0.083
	(0.649)	(0.681)	(0.609)
Positive Adjustment$_{t-1}$	0.322**	0.512**	0.475**
	(0.155)	(0.295)	(0.260)
Round	-0.109***	-0.109***	-0.104***
	(0.016)	(0.019)	(0.014)
Unknown × Positive Adjustment$_{t-1}$		-0.393	-0.315
		(0.392)	(0.358)
Random × Positive Adjustment$_{t-1}$		-0.169	-0.139
		(0.358)	(0.324)
Risk			0.732
			(1.118)
Female			-1.374***
			(0.350)
Intercept	5.044***	4.951***	5.556***
	(0.586)	(0.617)	(0.695)
Hurdle Model (Zero Contribution)			
Unknown	0.117	0.108	-0.003
	(2.827)	(3.525)	(1.558)
Random	-0.507	-0.511	-0.309
	(2.014)	(2.107)	(0.803)
Risk	-5.460	-5.465	-4.432
	(6.919)	(6.824)	(2.137)
Female	2.605*	2.602*	2.531*
	(1.803)	(1.946)	(1.603)
Intercept	4.537	4.544	3.580
	(5.337)	(5.415)	(1.704)
Observations	3648	3648	3648

Block bootstrapped standard errors clustered by group in parentheses.
* $p<0.1$, ** $p<0.05$, *** $p<0.01$.
The Watchful treatment is the omitted category.

References

Aimone, J.A., Iannaccone, L.R., Makowsky, M.D., Rubin, J., 2013. Endogenous group formation via unproductive costs. Rev. Econ. Stud. 80 (4), 1215–1236.
Ashley, R., Ball, S., Eckel, C., 2010. Motives for giving: a reanalysis of two classic public goods experiments. South. Econ. J. 77 (1), 15–26.
Beck, J., Forstmeier, W., 2007. Superstition and belief as inevitable by-products of an adaptive learning strategy. Hum. Nat. 18 (1), 35–46.
Benjamin, D.J., 2019. Errors in probabilistic reasoning and judgment biases. In: Handbook of Behavioral Economics: Applications and Foundations 1, vol. 2, pp. 69–186.
Benjamin, D.J., Choi, J.J., Fisher, G., 2016. Religious identity and economic behavior. Rev. Econ. Stat. 98 (4), 617–637.
Berger, P.L., 1969. The Social Reality of Religion. Faber, London.
Bleak, J.L., Frederick, C.M., 1998. Superstitious behavior in sport: levels of effectiveness and determinants of use in three collegiate sports. J. Sport Behav. 21 (1), 1.
Bochet, O., Page, T., Putterman, L., 2006. Communication and punishment in voluntary contribution experiments. J. Econ. Behav. Organ. 60 (1), 11–26.
Boyer, P., 2007. Religion Explained: The Evolutionary Origins of Religious Thought. Basic Books.
Brandts, J., Cooper, D.J., Rott, C., 2019. Communication in laboratory experiments. In: Handbook of Research Methods and Applications in Experimental Economics. Edward Elgar Publishing.
Brandts, J., Rott, C., Solà, C., 2016. Not just like starting over-leadership and revivification of cooperation in groups. Exp. Econ. 19 (4), 792–818.
Charness, G., Samek, A., van de Ven, J., 2022. What is considered deception in experimental economics? Exp. Econ. 25 (2), 385–412.
Chaudhary, L., Rubin, J., 2016. Religious identity and the provision of public goods: evidence from the Indian princely states. J. Comp. Econ. 44 (3), 461–483.
Chaudhuri, A., 2011. Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. Exp. Econ. 14 (1), 47–83.
Chaudhuri, A., 2016. Recent advances in experimental studies of social dilemma games.
Chaudhuri, A., 2018. Belief heterogeneity and the restart effect in a public goods game. Games 9 (4), 96.
Cheung, S.L., 2014. New insights into conditional cooperation and punishment from a strategy method experiment. Exp. Econ. 17 (1), 129–153.
Cookson, R., 2000. Framing effects in public goods experiments. Exp. Econ. 3 (1), 55–79.
Cox, C.A., Stoddard, B., 2021. Common-value public goods and informational social dilemmas. Am. Econ. J. Microecon. 13 (2), 343–369.

Cragg, J.G., 1971. Some statistical models for limited dependent variables with application to the demand for durable goods. Econometrica, 829–844.
Dickinson, D.L., 1998. The voluntary contributions mechanism with uncertain group payoffs. J. Econ. Behav. Organ. 35 (4), 517–533.
Dorner, Z., Tucker, S., Hassan, G., et al., 2020. A veil of ignorance: uncertain and ambiguous individual productivity supports stable contributions to a public good. Tech. Rep. University of Waikato.
Durkheim, E., 1965,[1912]. The elementary forms of the religious life. Trans. J.W. Swain (1965).
Dyble, M., Thompson, J., Smith, D., Salali, G.D., Chaudhary, N., Page, A.E., Vinicuis, L., Mace, R., Migliano, A.B., 2016. Networks of food sharing reveal the functional significance of multilevel sociality in two hunter-gatherer groups. Curr. Biol. 26 (15), 2017–2021.
Ellsberg, D., 1961. Risk, ambiguity, and the Savage axioms. Q. J. Econ., 643–669.
Ember, C.R., Ember, M., 2011. A basic guide to cross-cultural research. In: User's Guide: HRAF Collection of Ethnography.
Engel, C., Moffat, P.G., 2012. Estimation of the House Money Effect Using Hurdle Models. Discussion Paper Series of the Max Planck Institute for Research on Collective Goods 2012-13. Max Planck Institute for Research on Collective Goods.
Feltovich, N., 2003. Nonparametric tests of differences in medians: comparison of the Wilcoxon–Mann–Whitney and robust rank-order tests. Exp. Econ. 6 (3), 273–297.
Feltovich, N., 2005. Critical values for the robust rank-order test. Commun. Stat., Simul. Comput. 34 (3), 525–547.
Fischbacher, U., 2007. z-tree: Zurich toolbox for ready-made economic experiments. Exp. Econ. 10 (2), 171–178.
Foster, K.R., Kokko, H., 2009. The evolution of superstitious and superstition-like behaviour. Proc. R. Soc. Lond. B, Biol. Sci. 276 (1654), 31–37.
Froese, P., Bader, C., 2010. America's Four Gods: What We Say about God–and What That Says about Us. Oxford University Press.
Gershman, B., 2016. Witchcraft beliefs and the erosion of social capital: evidence from sub-Saharan Africa and beyond. J. Dev. Econ. 120, 182–208.
Gilovich, T., Vallone, R., Tversky, A., 1985. The hot hand in basketball: on the misperception of random sequences. Cogn. Psychol. 17 (3), 295–314.
Gintis, H., Smith, E.A., Bowles, S., 2001. Costly signaling and cooperation. J. Theor. Biol. 213 (1), 103–119.
Hajikhameneh, A., 2022. Growing Gods: Superstition, Sacrifice, and Prosocial Behavior in Intergenerational Public Goods Games. Tech. Rep., Working Paper. Institute for the Study of Religion, Economics, and Society, Chapman University.
Hajikhameneh, A., Iannaccone, L.R., 2022. Shocks and Solidarity: Exploring the Foundations of Faith. Tech. Rep., Working Paper. Institute for the Study of Religion, Economics, and Society, Chapman University.
Hayek, F.A., 1989. The Fatal Conceit, edited by W.W. Bartley, III. University of Chicago Press.
Heap, S.P.H., Ramalingam, A., Stoddard, B.V., 2016. Endowment inequality in public goods games: a re-examination. Econ. Lett. 146, 4–7.
Henrich, J., Ensminger, J., McElreath, R., Barr, A., Barrett, C., Bolyanatz, A., Cardenas, J.C., Gurven, M., Gwako, E., Henrich, N., et al., 2010. Markets, religion, community size, and the evolution of fairness and punishment. Science 327 (5972), 1480–1484.
Hoffmann, R., 2013. The experimental economics of religion. J. Econ. Surv. 27 (5), 813–845.
Holt, C., Laury, S., 2002. Risk aversion and incentive effects. Am. Econ. Rev. 92 (5), 1644–1655.
Iannaccone, L.R., 1992. Sacrifice and stigma: reducing free-riding in cults, communes, and other collectives. J. Polit. Econ. 100 (2), 271–291.
Iannaccone, L.R., 2006. The market for martyrs. Interdiscip. J. Res. Relig. 2.
Iannaccone, L.R., Berman, E., 2006. Religious extremism: the good, the bad, and the deadly. Public Choice 128 (1), 109–129.
Isaac, R.M., Walker, J.M., 1988. Communication and free-riding behavior: the voluntary contribution mechanism. Econ. Inq. 26 (4), 585–608.
Koukoumelis, A., Levati, M.V., Weisser, J., 2012. Leading by words: a voluntary contribution experiment with one-way communication. J. Econ. Behav. Organ. 81 (2), 379–390.
Kuran, T., 2012. The Long Divergence: How Islamic Law Held Back the Middle East. Princeton University Press.
Machina, M.J., Siniscalchi, M., 2014. Ambiguity and ambiguity aversion. In: Handbook of the Economics of Risk and Uncertainty, vol. 1. Elsevier, pp. 729–807.
Malefijt, A.d.W., 1968. Religion and Culture: An Introduction to Anthropology of Religion. Macmillan College.
Malinowski, B., 1948. Magic, Science and Religion, vol. 23. Beacon Press, Boston.
McCloskey, D.N., 2006. Bourgeois Virtue. Wiley Online Library.
McCloskey, D.N., 2010. Bourgeois dignity: Why economics can't explain the modern world. University of Chicago Press.
McCloskey, D.N., 2016. Bourgeois Equality: How Ideas, Not Capital or Institutions, Enriched the World. University of Chicago Press.
Miller, J.B., Sanjurjo, A., 2018. Surprised by the hot hand fallacy? A truth in the law of small numbers. Econometrica 86 (6), 2019–2047.
Norenzayan, A., 2013. Big Gods: How Religion Transformed Cooperation and Conflict. Princeton University Press.
Norenzayan, A., Shariff, A.F., Gervais, W.M., Willard, A.K., McNamara, R.A., Slingerland, E., Henrich, J., 2016. The cultural evolution of prosocial religions. Behav. Brain Sci. 39.
Nunn, N., Sanchez de la Sierra, R., 2017. Why being wrong can be right: magical warfare technologies and the persistence of false beliefs. Am. Econ. Rev. 107 (5), 582–587.
Ortmann, A., 2019. Deception. In: Handbook of Research Methods and Applications in Experimental Economics.
Oskarsson, A.T., Van Boven, L., McClelland, G.H., Hastie, R., 2009. What's next? Judging sequences of binary events. Psychol. Bull. 135 (2), 262.
Palmer, C.T., 1989. The ritual taboos of fishermen: an alternative explanation. MAST, Marit. Anthropol. Stud. 2 (1), 59–68.
Purzycki, B.G., Apicella, C., Atkinson, Q.D., Cohen, E., McNamara, R.A., Willard, A.K., Xygalatas, D., Norenzayan, A., Henrich, J., 2016. Moralistic gods, supernatural punishment and the expansion of human sociality. Nature 530 (7590), 327–330.
Purzycki, B.G., Henrich, J., Apicella, C., Atkinson, Q.D., Baimel, A., Cohen, E., McNamara, R.A., Willard, A.K., Xygalatas, D., Norenzayan, A., 2018. The evolution of religion and morality: a synthesis of ethnographic and experimental evidence from eight societies. Relig. Brain Behav. 8 (2), 101–132.
Roberts, K.A., Yamane, D., 2015. Religion in Sociological Perspective. Sage Publications.
Rubin, J., 2017. Rulers, Religion, and Riches: Why the West Got Rich and the Middle East Did Not. Cambridge University Press.
Ruffle, B.J., Sosis, R., 2006. Cooperation and the in-group-out-group bias: a field test on Israeli kibbutz members and city residents. J. Econ. Behav. Organ. 60 (2), 147–163.
Skinner, B.F., 1948. "Superstition" in the pigeon. J. Exp. Psychol. Gen. 38, 168–272.
Skinner, B.F., 1992. "Superstition" in the pigeon. J. Exp. Psychol. Gen. 121 (3), 273.
Stark, R., 1999. A theory of revelations. J. Sci. Study Relig., 287–308.
Stark, R., 2003. One True God: Historical Consequences of Monotheism. Princeton University Press.
Stark, R., 2005. The Victory of Reason. How Christianity Led to Freedom, Capitalism, and Western Success. New York.
Stark, R., Bainbridge, W.S., 1996. A Theory of Religion. Rutgers Univ Pr.
Steadman, L.B., Palmer, C.T., 2015. Supernatural and Natural Selection: Religion and Evolutionary Success. Routledge.
Stoddard, B., 2017. Risk in payoff-equivalent appropriation and provision games. J. Behav. Exp. Econ. 69, 78–82.
Théroude, V., Zylbersztejn, A., 2020. Cooperation in a risky world. J. Public Econ. Theory 22 (2), 388–407.
Villeval, M.C., 2020. Public goods, norms and cooperation. In: Handbook of Experimental Game Theory. Edward Elgar Publishing.
Wade, N., 2009. The Faith Instinct: How Religion Evolved and Why It Endures. Penguin.
Warner, C.M., Kılınç, R., Hale, C.W., Cohen, A.B., Johnson, K.A., 2015. Religion and public goods provision: experimental and interview evidence from catholicism and Islam in Europe. Comp. Polit. 47 (2), 189–209.
Wilson, R.K., Sell, J., 1997. "Liar, liar..." cheap talk and reputation in repeated public goods settings. J. Confl. Resolut. 41 (5), 695–717.
Wormald, B., 2015. Religious landscape study. Pew Research Center's Religion & Public Life Project, 11.
Wright, R., 2010. The Evolution of God: The Origins of Our Beliefs. Hachette, UK.