God Games: An Experimental Study of Uncertainty, Superstition, and Cooperation*

Aidin Hajikhameneh† **Laurence Iannaccone**‡

December 10, 2017

Abstract

This paper tests classic claims about the origins and functions of religion and superstitions. We do so by modifying the standard VCM public goods game, adding a god-like agent that adjusts group earnings in a manner that, though effectively random, might plausibly depend on rates of cooperation. Although players' earnings and the agent's adjustments are reported separately, the mere presence of adjustments induces radically higher rates of group investment – whether the adjustments are described as "random" or "chosen" by an AI who monitors investments. Investment patterns, survey responses, and group chat witness to superstitions that arise in response to risk and (especially) to uncertainty. Although some superstitions enhance group welfare, text-based chat turns encourages a counterproductive quest for magical numbers.

JEL classifications: *C7, C9, Z1*
Keywords: *religion, superstition, uncertainty, collective action, experiments*

*We thank James Andreoni, Christopher Bader, Gary Charness, Hillard Kaplan, Michael McBride, Douglas Norton, Ryan Oprea, Jared Rubin, Eric Schniter, Roman Sheremeta, Nathaniel Wilcox, and participants in the Economic Science Association North American Meeting at Virginia Commonwealth University for their excellent feedback. We gratefully acknowledge funding support from the John Templeton Foundation and Chapman University's Argyros School of Business and Institute for the Study of Religion, Economics, and Culture.

†Institute for the Study of Religion, Economics and Society, Chapman University, 338 N. Glassell, Orange, CA 92866, USA, e-mail: hajikhameneh@chapman.edu

‡Institute for the Study of Religion, Economics and Society, Chapman University, 338 N. Glassell, Orange, CA 92866, USA, e-mail: iannacco@chapman.edu

1 Introduction

This paper seeks to test a pair of classic claims about the origins and functions of religion. According to the first claim, people more readily attribute outcomes to the influence of supernatural agents when those outcomes involve substantial risk. According to the second, the resulting supernatural attributions can raise both individual and collective welfare by boosting confidence and cooperation. We test these claims with a lab experiment that incorporates a god-like agent in an otherwise standard public goods game. Though our experiment generates nothing approaching real-world religions, it does demonstrate the ease with which superstitions arise under conditions of uncertainty. More importantly, we find that cooperation increases dramatically in the face of external shocks to group earnings, whether the source of those shocks is described as an "artificial intelligence" that "adjusts" earnings up or down based on players' contributions or a "random number generator" that mechanically raises or lowers group earnings by a fixed amount. The increased cooperation is all the more surprising given that our earnings shocks are simple and separately reported increments that neither mask members contributions nor magnify their value. Altering the game so that group members can communicate via text increases investment further still but also promotes counter-productive quests for magical numbers.[1]

Our results point to four major conclusions, all tentative but potentially important. First, economists should seriously study the ways in which attributions of supernatural agency may promote cooperation in the face of risk and uncertainty. Second, experimental methods can and should be used to induce and analyze superstition and supernaturalism. Third, standard experiments should be reworked to see whether seemingly irrelevant forms of risk and uncertainty

[1] In our opinion, real-world superstitions are best defined as "demonstrably false rules of conduct that relate feasible actions to valued outcomes and that persist within a group over time." By "demonstrably false," we mean that group members can disconfirm the superstition's implicit claims of cause and effect within reasonable cost and time using readily available knowledge, methods, and resources. In light of our experiment's short duration and limited interactions, we will not require the last condition in this paper. Even so, we are using "superstition" in a much more narrow and precise sense than is typical of dictionaries, which focus on false beliefs about cause and effect and thereby gloss over the fact that in every time and place, much of what people believe concerning cause and effect turns out to be false. Standard usage also conflates durable, costly, and widely held superstitions with the incorrect or ineffectual ideas and actions that people (or even animals) happen to adopt for purely idiosyncratic reasons.

routinely alter individual behavior and collective outcomes.[2] We may discover that humans respond to "cheap noise" no less than "cheap talk" and that religions adapt and evolve in ways that channel these responses and promote collective action. Fourth and finally, experimentalists should focus much more attention on mechanisms that can reduce free-riding and sustain collective action even when the actions of individuals *cannot* be observed (and hence cannot be modified by means of the systems of reward, punishment, exclusion, and group formation emphasized in recent research).

Related Research: Ours is by no means the first experiment to study religion's capacity to sustain cooperation. But to the best of our knowledge all previous experiments concern *pre-existing* religiosity.[3] They thus illuminate the impact of real-world religions, but the underlying causes of religiosity remain beyond their reach.[4] We instead seek to create something akin to supernaturalism in the lab. To do so, we abstract from the complexity and diversity of real-world religions and focus instead on the defining feature of supernaturalism: the presumed existence of beings or forces that can alter natural outcomes and respond to human action (Iannaccone, 2006; Norenzayan, 2013; Stark and Bainbridge, 1996).[5] There are many reasons to study the effects of supernatural motivations. Supernaturalism is a salient feature all historic cultures, and nearly all those cultures invoke supernaturalism to justify beliefs, practices, and institutions that shape individual conduct and group identity (Malefijt, 1968; Wade, 2009; Steadman and Palmer, 2015). Anthropologists have carefully studied supernaturalism across more than a thousand societies, cataloging their finding together with numerous other measures in mas-

[2]Work along these lines has already begun, most notably in a recent working paper by Butera and List (2017) that adds "Knightian uncertainty" concerning the group contribution multiplier in otherwise standard VCM experiments. Our results agree with theirs insofar as rates of contribution increase for no obvious reason, but in our experiment increased contributions persists even when the uncertainty is replaced by an utterly transparent source of risk equivalent to flipping coins. We pursue this topic further in sections 4 and 5 below.

[3]For a review of experiments in the economics of religion prior to 2010, see Hoffmann (2013). For references to more recent experiments and related experiments in other fields, see Benjamin *et al.* (2016) and Norenzayan (2013).

[4]For excellent examples of religious priming, see Benjamin *et al.* (2016) and Warner *et al.* (2015). For insights derived from comparisons or interactions among religions, see Ruffle and Sosis (2006), Chaudhary and Rubin (2016),and Brooks *et al.* (2016).

[5]Iannaccone and Berman (2006) define religion and magic as alternative forms supernaturalism, differentiated by the former's focus on interactions with supernatural *beings* and the latter's focus on manipulating impersonal supernatural forces. In doing so, they offer an economic approach to a key distinction that traces back to anthropologists Tyler and Frazier and the sociologist Durkheim.

sive cross-cultural databases suitable for statistical analysis.[6] The links between supernatural beliefs, moral precepts, and group identity are especially salient in Judaism, Christianity, Islam, Hinduism, and Buddhism. Recent contributions to evolutionary biology, anthropology, and psychology argue these "big god" religions were key to the emergence of large scale civilizations and market economies (Norenzayan, 2013; Wade, 2009; Wright, 2010; Stark, 2005).[7] 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.[8] Finally, we note that developments with the fields of experimental and behavioral economics provide yet another rationale for seeking to better understand the causes and consequences of supernaturalism and superstition. Forced by their own results to place ever more emphasis on non-rational, quasi-rational, and downright irrational motivations, these fields can ill afford to ignore humanities most ancient and enduring expressions of faith.[9]

***E*xperimental framework:**The classic public goods game (known as the voluntary contribution mechanism or "linear VCM") provides a natural starting point for experiments on supernaturalism partly because it has been studied so extensively – partly because it provides a simple and standard setting for testing determinants of cooperation, and partly because subjects play in groups. This last feature is important insofar as real religions are always and everywhere a

[6]Anthropologists have devoted tremendous effort to cataloging the features of cultures, most notably as summarized in Mudock's Ethnographic Atlas (http://eclectic.ss.uci.edu/ drwhite/worldcul/atlas.htm) and the Human Relations Area Files (http://hraf.yale.edu/).

[7]Hayek (1989) appears to be one of the first major scholars to argue 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, in particular, that it was "the main monotheistic ones" that provide the pre-rational moral foundation that underpins the "extended [market] order" upon which civilization depends for its very existence (pp. 136 & 137).

[8]See, for example, Rubin (2017), Kuran (2012), Nunn and de la Sierra (2017), McCloskey (2006, 2010, 2016), and Gershman (2016).

[9]Experimentalists have been drawn (and often dragged) from austere models of pure self-interest and rational calculation to the much messier world of hard-wired emotions, evolved predispositions, and interpersonal effects. Thus, although Isaac *et al.* (1984) voluntary contribution mechanism ("linear VCM") remains the workhorse game for experimental studies of public goods provision, publications of the past decade overwhelmingly focus motivations, behavior, and policies outside the traditional neo-classical framework. For a survey of the VCM literature from the mid-1990s through 2010, see Chaudhuri (2011) and note the focus on non-maximizing player types (such as "conditional cooperators"), non-standard motivations (such as fairness, envy, anger, reciprocity, and altruism), and policies predicated upon players willingness to undertake costly punishments or that can only benefit others or be deterred by expressions of disapproval that impose no costs. For more recent work, see Chaudhuri (2016).

social phenomenon (Berger, 1969; Iannaccone, 1992; Stark, 1999). It takes a village, or at least a tribe, to sustain supernaturalism in forms more durable and sophisticated than idiosyncratic superstitions.[10]

Our "god game" extends the classic VCM in a seemingly trivial way. As in the standard repeated game, all subjects are assigned to anonymous 4-player groups, all receive a fixed number of tokens at the start of each round, all divide their tokens between their personal account and their group's account, and all then earn a payoff equal to the number of tokens they kept in their own account plus a multiple of the total number of tokens invested in their group's account. The game ends after a pre-announced number of rounds. In our "baseline" treatment, subjects merely play the standard VCM, but in our "artificial intelligence" treatment, each round has two stages. Stage one is a standard VCM round, identical to those of the baseline treatment. But stage two consists of a positive or negative "adjustment" determined by the AI. We truthfully inform subjects that "the AI makes a separate adjustment calculation for each group based on a complex series of calculations that take account players' investment decisions." The AI calculations are indeed complex and do indeed take account of investments but they also incorporate random factors and weightings that yield adjustments that are nearly random, independent, and identically distributed (appendix B). By characterizing the adjustments as the consequence of an AI's decisions we are, of course, priming subjects to view them as the actions of an quasi-supernatural agent who monitors all activity and alters outcomes at will.

To distinguish the impact of quasi-supernaturalism from mere risk, we also ran a "*random*" treatment, with a third set of subjects. This game has exactly the same structure as the *AI* treatment and the adjustment are computed in *nearly* the same way, but subjects are truthfully informed that the adjustments determined by random number calculations with 50/50 odds of being positive or negative. Finally, we re-ran the entire experiment, giving all players opportunities to communicate within their groups through an on-screen text box.

[10] On the other hand, idiosyncratic behavior of a quasi-superstitious nature can be induced in individual pigeons and rats simply by putting them on random reinforcement schedules (Skinner, 1948; Morse and Skinner, 1957; Reberg *et al.*, 1977; Timberlake and Lucas, 1985).

Motivation: Our experiment was motivated by a theory of supernaturalism that traces back to field work by the renowned anthropologist, Bronislaw Malinowski. In his classic essay on "Magic, Science, and Religion," Malinowski (1948) disputed the claim that irrationality was characteristic of "primitive" minds and cultures, that the supernatural views of such people were the product of anti-scientific reasoning, and that by extension all religions were grounded on irrationality. After spending several years among the Trobriand Islanders of Melanesia, Malinowski could confirm that these supposedly primitive performed most of their day-to-day activities in remarkably rational, empirical, and effective ways:

> *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.*
>
> (Malinowski, 1948, p. 11)

Supernaturalism was reserved for high-staked activities and outcomes that involved great uncertainty and danger. Thus Trobriand Islanders approached inland fishing and ocean fishing in radically different ways:

> *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.*
>
> (Malinowski, 1948, p. 14)

The natives thus distinguished the realms of empiricism and magic based largely on what they could understand and control.[11]

[11] Subsequent studies have drawn similar conclusions in cultures "primitive" and "modern", present and past, and large and small (Palmer, 1989).

Psychological findings help us understand how durable superstitions and supernaturalism might arise. Faced with nearly random outcomes, such as daily stock market returns, winning streak in sports and gambling, or spatial distributions from aerial bombing, people "see" trends, turning points, and clusters (Gilovich *et al.*, 1985; Wikipedia) in part because our brains are adapted to perceive patterns even when faced with very limited or noisy information, and in part because the consequences of randomness can be sufficiently subtle and counter-intuitive that even trained statisticians misinterpret its real-world manifestations. Moreover, as Skinner (1948) reported in a classic study of caged pigeons, random reinforcement (in the form of food dropped randomly into a pigeon's cadge) induces "a sort of superstition" , whereby the pigeon ends up repeating the irrelevant and idiosyncratic behavior that it happened to be doing when reinforced.[12] Whether or not humans acquire superstitions in a similar manner, there's little doubt that in sports and many other risky endeavors, people do become remarkably attached to beliefs and rituals that almost certainly have no capacity to directly influence the presumed outcomes (Bleak and Frederick, 1998).

***K*ey results:** Although we carefully review the experiment in sections to follow, Figure 1 captures many key results. The figure plots the trends in average contributions for each experimental treatment.[13] The solid blue line near the base of the figure plots average contributions for our standard VCM baseline treatment. As one might expect, the plot mirrors those obtained in numerous VCM experiments run since the early-1980s. Average contributions begin high, drop quite rapidly in the first few rounds, decay more slowly in subsequent rounds, and rapidly approach zero in the final rounds. The pattern is quite different in the AI treatment plotted in solid red. Overall contributions are 54% higher than those in the baseline treatment (averaging 17 out of a possible 40 tokens versus 11 in the baseline). Despite an initial contributions decline in periods 2 through 5 paralleling that of the baseline treatment, subsequent contributions display

[12]Skinner (1948) concludes that "[t]he bird behaves as if there were a causal relation between its behavior and the presentation of food, although such a relation is lacking 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."

[13]We created our figures with R (R Development Core Team, 2012), analyzed our data with Stata, and programmed the experiment in z-Tree.

almost no ongoing decay or last-period effect. The (solid green) random treatment mirrors the AI treatment, which is truly remarkable given that the 20 random treatment "adjustments" amount to nothing more than a *single* increment to overall earnings, which players know to be randomly determined without regard to their actions, and which in no way masks the information players receive about their group's investments and earnings.

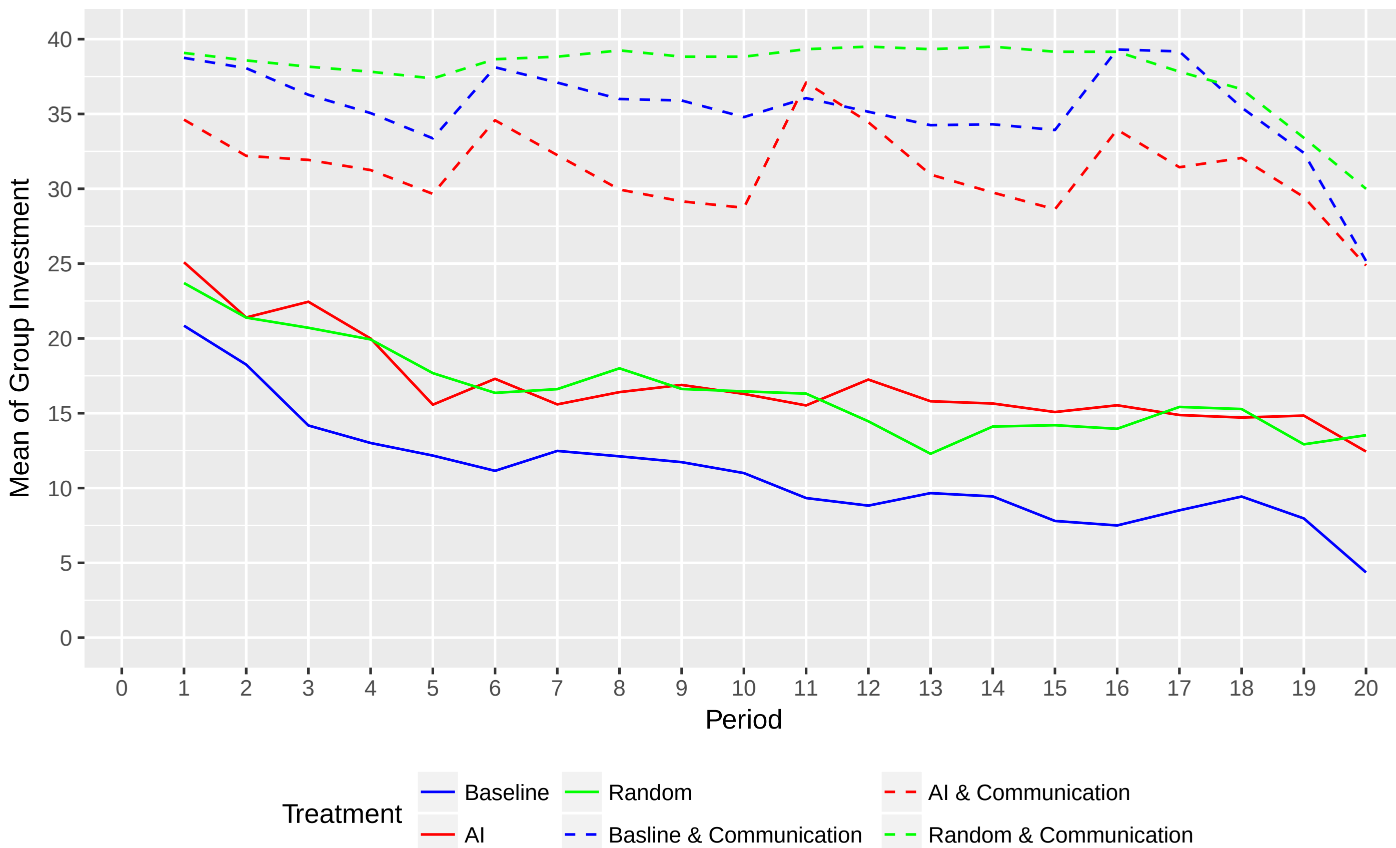

Figure 1: **Mean Group Investments Plotted by Period.**

The communication treatments (in dotted blue, red, and green) provide yet another surprise. Though overall rates of contribution display the dramatic rise that we would expect based on previous experiments,[14] the relative rates of contribution across the three treatments change. The AI treatment now yields the lowest overall rate of contribution; contributions in the baseline

[14]See Wilson and Sell (1997) for a summary of experimental findings regarding communication.

treatment substantially exceed those of the AI treatment; and the random treatment yields the highest contributions of all.

The following sections describe our experiment in greater detail, summarize the theories of superstition and supernaturalism that guide our procedure and predictions, analyze our results, and suggest some broad conclusions and recommendations.

2 Design and Procedure

As noted above, our "god game" extends the classic public goods game so as to incorporate uncertainty and quasi-supernatural agency. We do so by running three different treatments, each of which is built upon the standard VCM public good game popularized by Isaac *et al.* (1984). At the start of each game, subjects are assigned to anonymous 4-person groups, and at the start of every round, each subject (i) receives an endowment (E) and then chooses how much of this endowment will be invested in the group's joint account versus how much will remain in his/her personal account. Each player then receives initial earnings determined by the standard VCM formula: $\pi_i = (E - g_i) + \mu G_i$ where g_i denotes i's investment in the club good, G_i denotes the total amount invested by the members of i's group, and μ denotes the marginal per capita return (MPCR) on group investments.

***B*aseline treatment**: Our baseline treatment is a standard linear VCM, with each round consisting of the one stage described above.

***A*I treatment**: This treatment introduces a god-like agent that observes investments and adjusts earnings. In this second version of the game, each round has two stages. In stage one, play proceeds exactly as in the *baseline* treatment. But in the second stage, each player receives a high or low earnings adjustment of $H = +5$ or $L = -5$ tokens. The stage 2 instructions read as follows:

> **Stage 2**: 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.[15]

The ambiguous description of the AI's calculations creates Knightian uncertainty and accommodates different adjustment schemes. In the present version, the parameters of the adjustment algorithm yield adjustments that are, for all practical purposes, equally likely to be positive or negative.

***R*andom treatment**: In this treatment, a third set of subjects play a game that is identical to the game played in the AI treatment in all respects except for the last two sentences of the stage 2 instructions, which now read as follows:

> ... 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.

Table 1 summarizes the parameters and information that distinguish the three core treatments.

***R*isk aversion**: Since player's responses to the AI and random adjustment schemes might be influenced by their tastes for risk, we followed the game with an incentivized risk elicitation task (Holt and Laury, 2002). This task 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

[15] Appendix A documents the complete instructions associated with each treatment and appendix B describes the adjustment algorithm.

Treatments	Parameters μ	H	L	E	Information available to the subjects Adjustment
Standard VCM/Baseline	0.4	NA	NA	10	NA
Random adjustment	0.4	5	-5	10	Random
AI adjustment	0.4	5	-5	10	Unknown algorithm

Table 1: **Summary of Core Treatments**

lottery A is analogous to playing the zero-investment 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 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. For details, see Appendix A.3.

Questionnaire: Each session concluded with an on-line questionnaire that asks subjects about their major field of study and their religious beliefs and behavior, seeks open-ended feedback, and for the AI treatment only asks if they perceived any patterns in the AI's adjustments (Appendix A.4).

Communication treatments: As we have already noted, religion is a social phenomenon. Religious beliefs and behavior are shaped and sustained through talk, writings, song, art, rituals, and other forms of interaction that communicate assertions about supernatural cause and effect. Hence, we ran a second set of treatments that added communication to the core treatments described above. These communication treatments were identical to their core treatment counterparts except that prior to rounds 1, 6, 11, and 16 the members of each group were given time to chat with each by typing into a shared text box. We review the content and consequences of added communication in section 4 below.

General features of the experiment: We ran all treatments at a medium-sized American university, with subjects drawn randomly from the pool of students registered to participate in economics experiments. Each session involved 16 subjects randomly assigned to anonymous 4-person groups, and no subject participated in more than one session. In each session all groups played 20 rounds of just one treatment. In every round, each player chose how much to invest in the group account (with choices that could range from 0.00 to 10.00 in increments of 0.01). The rest of the endowment was automatically placed in the player's personal account. Overall earnings consisted of a $7 show-up payment, the money value of the tokens earned over all 20 rounds (converted at 5¢ per token), and money earned from the risk elicitation task. The entire experiment was programmed and run in z-Tree Fischbacher (2007).

The experiment included 25 sessions, with 16 subjects per session, and 400 subjects overall. We devoted 12 sessions to the core treatments: 4 sessions each to baseline, AI, and random without communication. The 192 core treatment subjects earned total payoffs that averaged $21.50 and ranged from $14.75 to $29.50. We devoted 13 additional sessions to communication treatments: 4 sessions to the baseline with communication, 3 sessions each to AI, and random with communication, and 3 additional sessions to a modified version of the AI communication treatment.[16]

3 Results from the Core Treatments

As noted in the introduction, many of our key results are captured by Figure 1, which plots the trends in average contributions for the six core and communication treatments. Figure 2 shows that core-treatment results look virtually the same whether plotted as period-by-period *median* investments or as means.

***B*aseline Treatment**: The solid blue line at the bottom of each figure confirms that our baseline treatment results mirror those routinely obtained in standard VCM experiments. There appears to be nothing odd about our subject pool or experimental design. For what it's worth, we've validated the following unremarkable conjecture:

> **Conjecture 1**: Group investments in the baseline treatment will (a) consistently average more than the Nash equilibrium rate of zero, (b) decline throughout the game,

[16]The only modification was a change in the set of potential adjustments, which were enlarged from $\{-5, +5\}$ to $\{-5, -2, 0, +2, +5\}$. Section 4 and Appendix C explain why we ran this modification and what we found. We are currently running additional AI sessions in which probability of high adjustments increases significantly in the group's absolute and relative investments.

Figure 2: **Median Group Investments Plotted by Period.**

particularly during the early and late portions of the game, and (c) approach zero in the final period.

***AI* Treatment**: Past experiments offer little guidance regarding the AI treatment effects, since no past experiment includes anything approaching a *deus ex machina* that sees all and inflicts punishment or reward. Game theory is likewise of limited help given our intentionally ambiguous description of the AI's adjustment process. Some players might assume that the AI will harshly punish groups whose members fail to invest most (or all) of their endowments in their group accounts. Others might assume that the AI tolerates or even rewards groups whose members are clever enough to recognize that investing zero is a dominate strategy in a finitely repeated VCM. Still others might imagine that the AI rewards groups that invest the most, or have the fewest free-riders, or maintain high levels of investment over time, and so forth.[17] As play proceeds, and the AI makes its actual adjustments, players might begin to infer that the process is, in fact, random. But they will be hard-pressed to confirm this inference given that no player observes the *individual* actions of his/her fellow group members, only total group investment, and no player observes anything at all regarding the other groups.

Recall, however, that that our AI treatment seeks to mimic the world of Malinowski's ocean-going Trobriand Islanders. If he is right, groups tend to impute supernatural agency to environments characterized by danger, uncertainty, and collective action. Although the resulting systems of beliefs and practices will vary greatly across cultures, the underlying function will be to promote group welfare through increased group effort and identity. We therefore conjecture as follows:

> **Conjecture 2**: Relative to the baseline treatment, group investments in the AI treatment will be higher, decline more slowly over time, and drop less sharply in the final periods.

As we show below, these predictions are strongly confirmed, not merely by in graphs like Figure 1, but also through regression analysis, even after introducing a variety of other explanatory variables.

***R*andom Treatment**: Even if our AI effects fit the patterns predicted by Malinowski's theory, the cause might not be imputed quasi-supernatural agency in the face of (Knightian) uncertainty. Our third treatment allows us to compare outcomes when players are confronted with *explicitly random* adjustments rather than the AI's effectively random but unspecified adjustments. Game theory offers a clear prediction in the face of risk that is both additive and exogenous:

[17] A highly rational/analytical player might well conclude that zero-investment remains the dominant strategy by reasoning as follows: The AI knows nothing about players apart from their (randomly assigned) personal ID, their group's (randomly assigned) group ID, and their investment choices. Hence, the actions of a group's members are likely to enter the AI's adjustment function in a symmetric manner. Such symmetry will likely mean that the marginal influence of any single member's current-period investment averages no more than 1/4 of the 10-token range in potential adjustments (and possibly much less if the adjustment function includes other inputs, such as past period investments, the investments of other groups, or random inputs).

Conjecture 3: Players in the Random treatment will behave in same manner as players in the baseline treatment.

To see why, rewrite a typical player's total earnings over all rounds of the game, grouping the adjustments at the end of the expression:

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

Recall also that we truthfully tell players that each adjustment, a_{it}, is randomly determined (with 50/50 odds of being -5 or +5) and we display each adjustment separately from the overall group investment, G_{it}. Hence, apart from an additional payoff that averages $0 (and falls between plus-or-minus $1.50 about 90% of the time), the Random treatment is *exactly* equivalent to the baseline linear VCM treatment (in which player earnings averaged about $21.08 with a standard deviation of $0.26). Absent some very strong and strange risk effects, we therefore expect no change in player behavior.

But as we have already seen in figures 1 and 2, this seemingly straightforward prediction fails, and does so quite spectacularly. Not only are group investments higher in the Random treatment, they are virtually identical to those of the AI treatment.

Statistical analysis tells the same story. Table 2 lists the mean, median, standard error, number of observations for each of the three core treatments. Overall rates of group investment in the AI and Random treatments exceed those of the baseline treatment by more than 50% differences that are not only statistically significant[18] but also quite large compared to other VCM extensions aimed at boosting cooperation.[19] On the other hand, the overall group investments in the AI treatment exceed those of the Random treatment by less than 3% - a difference that is substantively trivial and statistically insignificant.[20]

[18]We obtain p-values less than 0.01 using the two-sided Mann-Whitney-Wilcoxon rank sum tests (hereafter abbreviated as MWW). We applied the test to investments measured at the group level and thus have 16 observations per treatment.

[19]Such extensions include different forms of punishment and reward, approval and disapproval, exogenous sorting, endogenous group formation, and more (Chaudhuri, 2011, 2016)

[20]In this case, the MWW test's p-value is 0.68.

	Group investment in the join account			
Treatments	Mean	SE	Median	N
Standard VCM/Baseline	10.99	0.539	9.25	320
Random adjustment	16.50	0.445	16	320
AI adjustment	16.94	0.483	16.25	320

Table 2: **Summary statistics of the core treatments.**

Regression Results: Regression statistics help us understand how players respond to adjustments, how responses to the AI and Random treatments differ, and how gender, risk preferences, and other control variables influence behavior. Despite the graphical equivalence of the AI and Random treatments in figures 1 and 2, and despite the fact that both adjustment processes are essentially equivalent, it turns out that players respond quite differently when the process is described in a way that facilitates imputing agency. Moreover, those responses fit Malinkowski's theory of supernaturalism *and* Skinner's (1948) theory of superstition.

We estimate two sets of models. The first set uses GLS estimation with group-clustered standard errors and subject-specific random effects to control for the fact that subjects play multiple rounds. The second uses Tobit estimation to account for the fact that players' investments in the group account must fall between 0 and 10 tokens. We bootstrap the Tobit standard errors, and as with GLS we include random effects. All the results discussed below are statistically significant in both sets of regressions. The dependent variable (Investment) measures the number of tokens each player allocated to the group account in each of the 20 periods. Because we are seeking to understand how players respond to adjustments, we restrict our analysis to the AI and Random treatments, using AI as the omitted category.

In both tables, column 1 re-confirms that the AI and Random rates of overall investment are essentially equal. But column 2 reveals something that cannot be inferred from simple plots and summary statistics, nor from game theory or Malinowsky: the tendency for players to contribute

	(1) Investment$_t$	(2) Investment$_{t>1}$	(3) Investment$_{t>1}$	(4) Investment$_{t>1}$
Random	-0.109	-0.108	0.612	0.723
	(0.543)	(0.551)	(1.156)	(1.145)
Adjustment$_{t-1}$		0.261**	0.214**	0.325**
		(0.103)	(0.105)	(0.138)
Risk			0.608	0.621
			(2.066)	(2.062)
Random × Risk			-1.618	-1.627
			(2.418)	(2.417)
Female			-0.614*	-0.612*
			(0.371)	(0.371)
Period			-0.091***	-0.092***
			(0.016)	(0.015)
Random × Adjustment$_{t-1}$				-0.222
				(0.186)
Intercept	4.235***	4.006***	5.117***	5.072***
	(0.411)	(0.428)	(0.959)	(0.958)
Observations	2560	2432	2432	2432
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 AI adjustment treatment is the omitted category.

Table 3: **GLS Analysis of the Investment Decision**.

more to the group when the previous period's adjustment was positive rather than negative. It would appear that players become more public-spirited following a reward.[21]

Columns 3 and 4 show that players' investments are not significantly related to risk preferences (as measured by our version of the Holt and Laury (2002) risk elicitation task). Women invest less than men, but the difference is not statistically significant in the Tobit estimations of Table 4. The statistically significant "Period" effect shows that group investment declines over time.

Consider what we would expect to see if our (randomly reinforced) 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 is that random rewards condition animals to associate the rewards with whatever happened to precede them.

[21]Given that most players are investing well below the maximum, there's no obvious reason why rewards should boost group investment more effectively than punishments.

	(1) Investment$_t$	(2) Investment$_{t>1}$	(3) Investment$_{t>1}$	(4) Investment$_{t>1}$
Random	-0.133	-0.126	0.591	0.725
	(0.637)	(0.479)	(1.189)	(1.047)
Adjustment$_{t-1}$		0.343***	0.279**	0.411**
		(0.119)	(0.113)	(0.203)
Risk			0.409	0.424
			(1.846)	(2.223)
Random $\times$ Risk			-1.647	-1.656
			(2.426)	(2.347)
Female			-0.897	-0.895
			(0.606)	(0.627)
Period			-0.121***	-0.123***
			(0.019)	(0.020)
Random $\times$ Adjustment$_{t-1}$				-0.267
				(0.272)
Intercept	4.111***	3.801***	5.516***	5.463***
	(0.417)	(0.362)	(1.094)	(1.200)
Observations	2560	2432	2432	2432
Wald Chi-Sq.	0.0438	9.615	53.62	45.93

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

Table 4: **Tobit Analysis of the Investment Decision**.

These associations are certainly not the product of conscious reflection, but they can be modeled as by-products of adaptive learning strategies and natural selection (Beck and Forstmeier, 2007; Foster and Kokko, 2009). Hence,

Conjecture 4(a): Insofar as random reinforcement induces Skinnerian-style "superstition" in people as well as animals, the actions of god-game players will display *more inertia* following a positive adjustment than a negative adjustment. Hence, a positive adjustment raises the probability that a player *repeats* the previous investment.

Conjecture 4(b): Because the Skinnerian-style of superstition in unconsciously motivated, positive adjustments will induce inertia in *both* the AI and Random treatments.

But what if players also engage in *conscious* forms of superstition more like that which Malinowski attributed to the Trobriand Islanders? If our subjects *consciously* infer a relationship between their actions and the AI's subsequent adjustments, then we would expect more inertia in the AI treatment than the Random treatment:

> **Conjecture 5:**: Insofar as random reinforcement induces Malinowski-style superstition in players who perceive adjustments as a quasi-supernatural responses to their behavior, the AI-treatment players will display more inertia than their Random-treatment counterparts following positive adjustments.

We test conjectures 4 and 5 by regressing two measures of investment inertia on the previous period's adjustment. The first is the absolute change in investment from the previous period to the current period: $AbsChange_t = |Investment_t - Investment_{t-1}|$. The second is a variable that indicates whether investment in the current and preceding period are exactly equal: $NoChange = 1$ if $(Investment_t = Investment_{t-1})$ and 0 otherwise. We then run GLS regressions of each inertia measure on a 0/1 indicator of positive previous adjustment interacted with a Random (versus AI) treatment indicator. As in our previous GLS regressions, we cluster standard errors by group and include player-specific random effects.

The regression results strongly support all three conjectures. The first fitted equation is $AbsChange = 2.18 - 0.14\,Random - 0.85\,PositiveAdjustment^{***} + 0.34\,Random \times PosAdjustment^{*}$. Hence, after a negative adjustment players typically change their investment by 2.18 tokens, but positive adjustment in the AI treatment reduces the magnitude of this change by 39% (0.85 tokens). The corresponding effect in the Random treatment is smaller but still significant, reducing the magnitude of the typical period-to-period investment by 25% (from 2.04 tokens to 1.53). The second fitted equation is $NoChange = 0.30 + 0.015\,Random + 0.16\,PosAdjust^{***} - 0.08\,Random \times PosAdjust^{*}$. In words: after a negative adjustment players typically repeat their previous investment 30% of the time, but a positive adjustment in the AI treatment raises the probability of exact repetition to 46%. Again, this effect is smaller but still significant in the Random treatment, where a positive adjustment raises the probability of exact repetition from 29% to 37%.[22] Table D2 of Appendix D provide the full output from all three regressions.

Questionnaire Results: The results from our end-of-experiment survey provide additional evidence for Malinowski-style superstition. We asked subjects in the AI treatment to tell us whether they observed any patterns in the AI's adjustments and, if so, what the patterns were. Only 25% of the subjects claimed to have spotted patterns a fraction that was probably deflated by the fact that saying "yes" required the subject to expend additional time and effort typing 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 AI was rewarding the group for high overall investments and punishing low. The next most common interpretation was a variant on the "hot hands" cognitive error, which is known to support Skinnerian-style superstition in athletes and gamblers. I.e., subjects noted that the AI's positive and negative adjustments clustered or ran in streaks. Interestingly, no one claimed that the AI rewarded low investment. Hence, insofar as our AI treatment really did induce something like quasi-supernaturalism, the imputed supernatural agent was seen as acting like an Abrahamic god, rewarding good, public-spirited behavior and punishing evil.

As we will show in the next section, the tendency to repeat investments after positive adjustments persists when we introduce text-based chat to our AI and Random treatments with text-based communication. Moreover, the *content* of players' texts illustrate how superstitious inferences arise and how they are collectively reinforced.

4 Communication Treatments

The fact that adjustments influence cooperation and evince a form of quasi-supernatural inference is somewhat surprising given the extent to which both superstition and supernaturalism

[22]Because exact repetitions at the 0-token and 10-token extremes might simply represent "corner solution" behavior at the limits of the expenditure constraint, we also ran the second regression on a restricted sample consisting only of observations for which the previous period's investment was greater than 0 and less than 10. The coefficients change somewhat, but their relative magnitudes and statistical significance remain.

are *social* phenomena. To give group members more scope for social interactions and to test the claim that real-world faiths draw strength from shared claims about supernatural cause and effect, we add text-based communication to each of our core treatments. Specifically, we give members a chance chat with each prior to rounds 1, 6, 11, and 16 through a shared text box. The communication treatments are in all other respects identical to their core treatment counterparts. For full details and subject instructions, see Appendix A.2.

Predictions: To the extent that readers can ignore what they may already have learned about the communication treatments from figures 1 and 2, we invite predictions concerning the impact of communication on the three treatment's investment profiles. Though we ask you to "forget" what we've already noted about communication in this experiment, we urge you to recall what is generally known about the impact of communication on public goods games and what specifically emerged in our three non-communication treatments.

> **Conjecture 6**: Compared to its non-communication counterpart, group investments in the baseline-with-communication treatment (a) will be much larger and (b) will decline less quickly over time, but (c) will be just as likely to drop sharply in the final periods.

> **Conjecture 7**: Relative to the baseline-with-communication treatment, group investments in the AI-plus-communication treatment will be higher still, decline yet more slowly, and drop less sharply in the final periods.

And having been previously burned by our seemingly straightforward but utterly false assumption that random treatment would yield investments that mirror the standard VCM baseline, we apply what saw in the core treatments to the following prediction:

> **Conjecture 8**: Rates of group investment in the Random-with-communication treatment will mirror those of the AI-with-communication treatment and hence exceed those of the baseline-with-communication treatment.

These are the predictions we confidently made after reviewing the results of our our non-communication sessions.

4.1 Basic Results

Turning to again to Figure 1, we see that our baseline predictions (6) are entirely correct, but we have again anticipated only one of the adjustment treatment paths. This time the error concerns the AI treatment rather than Random treatment.

Communication does indeed send cooperation soaring in all three treatments. And even though baseline group investments leave little room for improvement (since they average 87% of the 40-token maximum), group investments across the random sessions are higher in 18 of 20 periods and 95% of the 40-token maximum overall.[23] On the other hand, group investments in the AI-with-communication sessions average just 78% of the 40-token maximum, and the AI investment rate falls below the corresponding baseline rate in 19 of 20 periods and less than random rate in 20 of 20 periods.[24] Table 4 lists the overall summary statistics for all three communication treatments.

	Group investment in the join account			
Treatments	Mean	SE	Median	N
Standard VCM/Baseline with Communication	35.44	0.551	40	320
Random adjustment with Communication	37.94	0.332	40	240
AI adjustment with Communication	31.35	0.713	38.75	240

Table 5: **Summary Statistics for Communication Treatments.**

[23] The difference between overall mean investments in the baseline versus random treatments is not "statistically significant" in the two-sided MWW rank sum tests (p-value = 0.76) which ignore the time structure of the data. But when the group means are separated by period, we see that the random treatment's 18 out of 20 period edge would be judged extremely significant (with non-parametric binomial distribution probability of Random mean exceeding baseline mean in 18 of 20 cases = 0.0002 if the true probability of random investment exceeding baseline in any given period was only 0.50).

[24] We can again compare the difference in overall rates of investment using the MWW test, which leads us to conclude that AI rate is "significantly" differs from both the baseline rate (p-value=0.02 and 0.00, respectively), but the period-by-period differences in group investments are even more striking.

To state the question in crude but semi-supernatural terms, *what the hell is going on in the AI communication treatment*? We argue below that the answer traces back to the way superstitious reasoning distorts VCM decisions when baseline contributions near 100%, but it helps to first review the results from GLS and Tobit regressions corresponding to those we ran for the non-communication treatments.

	(1) Investment$_t$	(2) Investment$_{t>1}$	(3) Investment$_{t>1}$	(4) Investment$_{t>1}$
Random with Communication	1.647***	1.673***	1.310**	1.366**
	(0.594)	(0.599)	(0.612)	(0.691)
Adjustment$_{t-1}$		0.047	0.060	0.123
		(0.181)	(0.174)	(0.297)
Risk			-2.061***	-2.066***
			(0.726)	(0.723)
Rand with Comm × Risk			0.854	0.862
			(0.841)	(0.833)
Female			-0.703*	-0.702*
			(0.405)	(0.405)
Period			-0.042	-0.042
			(0.039)	(0.038)
Rand with Comm × Adjustment$_{t-1}$				-0.124
				(0.354)
Intercept	7.838***	7.774***	9.469***	9.443***
	(0.570)	(0.600)	(0.817)	(0.853)
Observations	1920	1824	1824	1824
R^2	0.09	0.09	0.12	0.12

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

Table 6: **GLS Analysis of Investments with Communication**.

Tables 6 and 7 report the results of GLS and Tobit estimations. Again, the dependent variable is each investment in the group account, and again the data are restricted to the AI and Random treatments. The first specification in each table confirms the non-parametric conclusion that investments are substantially (and significantly) higher when subjects are explicitly told that the adjustments are random. Thus, communication "breaks" the AI and Random equivalence that arose when players faced exactly the same pair of adjustment processes but could *not* communicate.

	(1) Investment$_t$	(2) Investment$_{t>1}$	(3) Investment$_{t>1}$	(4) Investment$_{t>1}$
Random with Communication	7.205***	7.208***	5.974*	6.168**
	(1.659)	(1.582)	(3.535)	(2.666)
Adjustment$_{t-1}$		0.471	0.559	0.711
		(0.468)	(0.525)	(0.609)
Risk			-7.854	-7.866**
			(5.327)	(3.733)
Rand with Comm$\times$ risk			2.676	2.721
			(7.711)	(5.821)
Female			-2.428	-2.426
			(1.527)	(1.632)
Period			-0.164	-0.165*
			(0.103)	(0.088)
Rand with Comm$\times$ Adjustment$_{t-1}$				-0.459
				(0.946)
Intercept	12.675***	12.255***	18.585***	18.531***
	(1.483)	(1.122)	(3.407)	(2.638)
Observations	1920	1824	1824	1824
Wald Chi-Sq.	18.86	21.78	33.72	51.26

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

Table 7: **Tobit Analysis of Investments with Communication**.

Specifications 2, 3, and 4 control for the previous round's adjustment, risk aversion, gender, and interaction terms. As in the non-communication treatments, positive adjustments predict greater investments in all specification, but the estimated effects in the GLS regressions are smaller than their non-communication counterparts (probably because so many players' investments not hit the 10-token maximum). In the Tobit regressions, which control for the limits on investment, the estimated effects are larger than their non-communication counterparts but still insignificant.

The significant coefficients associated with risk aversion, period, 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. A possible explanation is that group chat shifts individual investment decisions toward the expressed recommendations (and underlying risk preferences) of the majority of group members. We explore this possibility in great detail below.

4.2 Counterproductive Superstitions

Malinowski's key insight was that the Trobriand Islanders resorted to magical thinking only when (as in the case of ocean fishing) danger, uncertainty, and technological limits rendered their usual, rational empiricism nearly worthless. Where they could more readily control and comprehend cause and effect (as in lagoon fishing), they relied upon knowledge, skill, and practical reasoning. Our Random treatments were designed to replace the uncertainty of a quasi-supernatural AI agent with the predictable risk of a computer-generated coin flip. Insofar as random-treatment subjects accepted our description of the random adjustment process, they faced a clear and comprehensible environment.[25] Within that environment, the sharing of ideas and information among group members could well lead to better decisions, or at least a better understanding of the expected consequences of their actions.

This, in fact, is what our analysis of the random treatment chat reveals. For example, at the start of the game, members of random-adjustment group 3 of session 21, exchanged the following texts:[26]

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*

[25] Of the hundred-plus subjects who played the Random treatments only one expressed any doubt that the adjustments were really random when asked for open-ended feedback regarding the experiment.

[26] In the beginning of each session subjects were randomly assigned numbers between 1 and 16. Member # refers to this randomly assigned number. Members # 1 to 4 were assigned to group 1, 5 to 8 were assigned to group 2, etc.

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 period 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 investments whether they received a negative adjustment. As can be seen from Figures 1 and 2 and Table 5, the vast majority of players in random-with-communication sessions invested their full endowment in the group account until the second to last period. Figures 1 and 2 and table 4 likewise affirm that subjects in the (no-adjustment) baseline-with-communication sessions do nearly as well. As might be expected given standard VCM structure and absence of any adjustment, players' texts center on the collective benefits of maximizing investment in the group account.

But for the AI groups communication is a mixed blessing. Group chat does lead to higher rates of investment, but it also sustains false and counterproductive beliefs. This might seem to contradict "functional" theories of superstition. But our lab-based "tribes" are ridiculously small, experience no evolutionary pressures or inter-tribal conflicts, and have just 20 rounds of play and 4 periods of text chat in which to fine tune their inferences. Seeking to see patterns in rewards and punishments that are essentially random, many groups explore inefficient strategies and some groups stick with these strategies even in when they perform poorly.

The fundamental problem is that communication moved players away from the false but functional notion that the AI was rewarding "good" (group-oriented) behavior toward the equally false but counterproductive notion that the AI paid off when players chose special numbers or sequences. This sort of thinking mirrors numerology lucky 7's, unlucky 13's, and the like. But it also fits the numerous video games that feature special numbers and winning patterns. Whatever the cause, communication seems to have led many of our AI groups to search for winning numbers. For example, after receiving some negative adjustments in the first 5 rounds, members of group 1 of session 13 decided to each invest 8 in the joint account:

Before period 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 the AI, some groups convinced themselves that they had found the magic number. Group 3 of session 14 thus reasoned as follows:

Before period 6:
Member # 11: *okay so the AI calculations is just flipping off overy 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 period 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 an additional three sessions that we ran using a slightly modified AI-with-communication treatment discussed in Appendix C. This treatment differed only in that the AI selected its adjustments (with nearly uniform probability) from {-5,-2,0,+2,+5} rather than just {-5,+5}. On the one hand, sophisticated players should give less weight to the adjustment process since the likely risk is now much lower (and the actual standard deviation has dropped from 5.0 to 3.4). On the other hand, five values make it harder to statistically disconfirm superstitious theories, particularly if players view -5 (which now appears only 20% of the time) as the only clear sign of failure/punishment. In practice, the greater scope for superstition seems to offset the reduced scope for effective reward and punishment. Mean rates of overall group investment are 32.56, only slightly higher than the 31.35 mean for the original {-5,5} treatment and significantly different from the original mean under the MWW test.[27]

And again, group chat reveals continuing counterproductive quests for magical numbers rather than potentially-helpful inferences that the AI is seeking to stimulate group investment. For example, group 3 of session 23 decided first to try an ascending sequence of investments (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 period 6:

Member # 12: *okay maybe we should try like 5,7,5,7,10*
Member # 9: *yes 5 and 7 were the best...*

[27]See Appendix C for additional information and statistics concerning this treatment.

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*
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 the 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 period 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:[28]

Before period 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*

[28]This example illustrates another potentially detrimental features of group chat. Absent communication, group members can reasonably assume there are not likely to invest equally (unless they all gravitate to 10's or 0's), especially since they could choose decimal values. Thus, as noted in footnote 3, non-communicating players can reasonably assume that their personal choices will decisively influence the AI no more than 1/4th of the time. But with chat, players can agree on any number, thereby making each of them the decisive"marginal" investor if the AI really is rewarding special numbers.

4.3 Statistical Evidence: AI-with-communication encourages counterproductive superstitions

There is, of course, a danger that in reviewing transcripts we ourselves are focusing on false patterns. Hence, we also subjected the free-form text to a simple form of statistical content analysis. Our test is based on the following conjecture:

> **Conjecture 8**: The search for illusory patterns and "magic" numbers will lead AI-with-communication groups to mention of numbers other than the collectively optimal 10 more frequently than their Random counterparts.

A MWW rank sum test confirms this conjecture. Groups in the AI-with-communication treatment mention numbers other than 10 much more frequently that their random treatment counterparts (p-value = 0.08).[29] So, even by this remarkably simple measure, it appears that chatting leads AI treatment players to get caught up in quests for magic numbers. Though asking why takes us beyond the limits of the experiment, we cannot help but wonder what it would have taken to get them talking about instead about the adjustments as rewards or punishments for good and bad behavior. We turn to this question in our concluding comments.

5 Discussion and Conclusion

Rather than restate points already emphasized in the abstract and introduction, we conclude with plans for work that builds upon our initial "god game".

Religion versus Magic: Scholars routinely distinguish "religion" and "magic" based on whether they emphasize supernatural *beings* or impersonal supernatural *forces*, and theres a good deal of evidence (from anthropology and sociology) that its only religion, not magic, that sustain strong bonds of group identity and commitment.[30]

[29] We used the package TXTTOOL in Stata created by Unislawa Williams to analyze the free-form text in this paper.

[30] The *supernatural* refers to forces or beings beyond or outside this natural order, which can suspend, alter, or ignore the normal flow of events. *Religion* consists of beliefs, practices, and institutions that relate to one or more *su-*

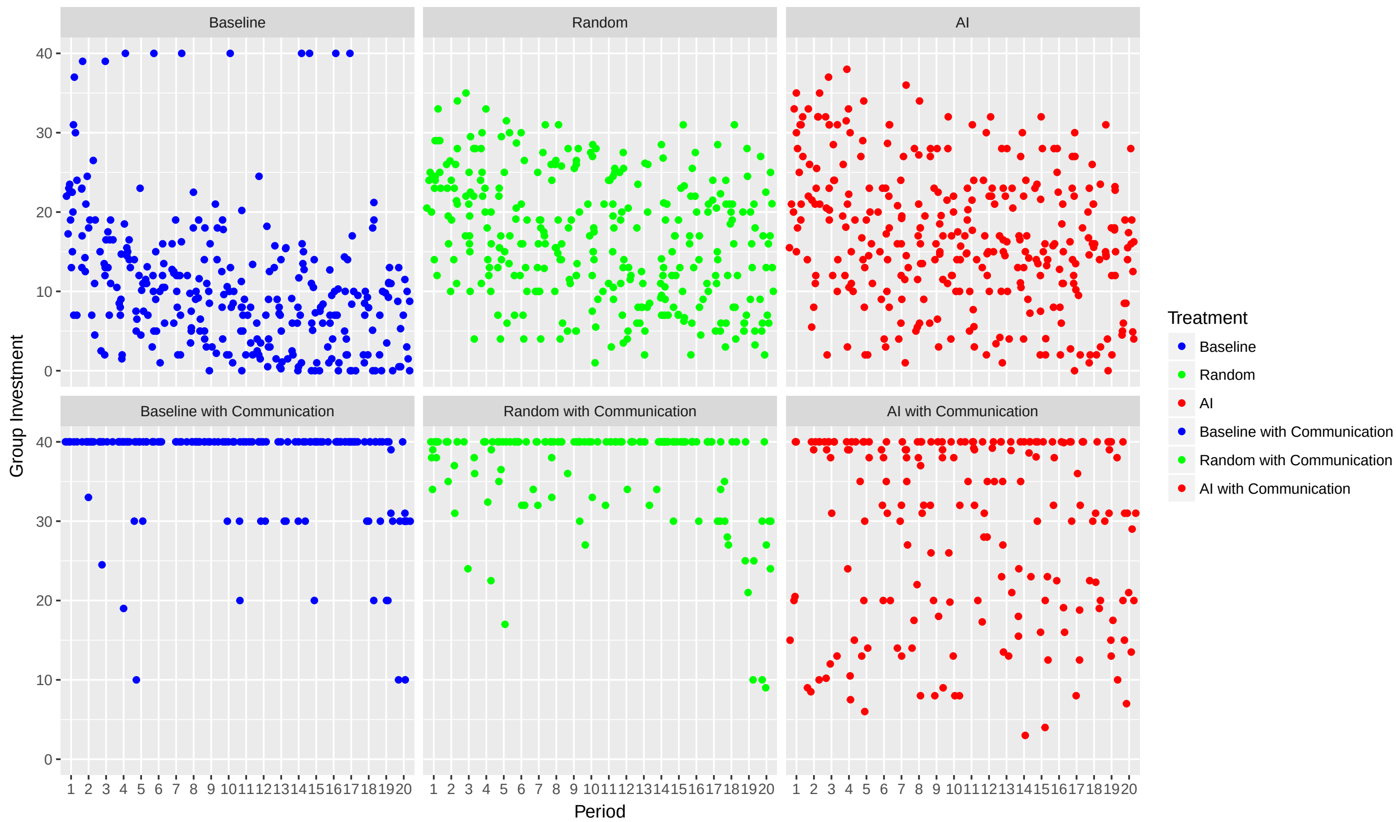

Figure 3: **Total Group Investments in Each Period, Separated by Treatment**

In light of our results to date, one might say that we went searching for "religion" and instead found "magic." In designing our initial "god game," we underestimated our subjects' propensity to magically/mechanically search for winning *numbers* rather than religiously/morally strive to behave in a manner that might win an AI-god's approval.

We want to see if relatively small changes to the experiment can create an environment in which players are more likely to view the AI as a quasi-supernatural being rather than an impersonal force.

We could, of course, program our AI to clearly and consistently reward cooperation, but that would eliminate the uncertainty and collective inference of non-verifiable cause and effect that is key to real-world supernaturalism.

pernatural beings. In contrast to religion, *magic* consists of beliefs, practices, and institutions that concern impersonal *supernatural forces* or the impersonal manipulation of *supernatural beings* (Iannaccone and Berman, 2006).

We there plan to modify our current design by incorporating different *types* of AI's modeled on different religions' view of their god(s), leaving players to infer which type of AI/god they happen to be facing. Players might again get an AI that behaves in a random (unpredictable, capricious) manner, but with other AI's the probability of a positive adjustment would increase in the group's most recent contributions. A third type of AI will take account of the group's entire history of contributions, but give greater weight to the more recent past. Yet a forth will judge the group in relation to the contributions of other groups, etc. We will carefully designed the different AI decision algorithms to mirror different types of theologies that arise in different religions, such as the loving (but also rewarding and punishing) God of Judaism, Christianity, and Islam, versus the self-centered and capricious gods of the Greco-Roman pantheon, or impersonal god of Deism, etc.

We are particularly interested in seeing if players tend to impute particular (Judeo-Christian) motives and propensities to their AI-god even when their actual (but unknown) AI is programmed for randomness. Even in the current experiment (which provided no information about the AI) many students claimed (in the post-game survey) that they thought the AI was rewarding cooperation.

Cultural Evolution: We also plan to run future experiments that incorporate different forms of group selection – through conversion (opportunities to switch groups), indoctrination (advising new players), proselytizing (advising members of other groups), and competition (publicizing the earnings of other groups), and the like. As we noted in section 4, supernaturalism is more likely to prove functional where groups compete and beliefs change. While encouraged by the findings of this paper, we are anxious to experiment with environments that offer more opportunities for adaptation.

Common Enemies: It's no coincidence that we're ending this study without drawing any conclusions about the large jump in collective investment caused by explicitly random shocks. The result surprised us and everyone we've consulted. It's all the more surprising that the jump appears in the first stage of the first period, even before any group received its first adjustment and hence before the adjustment stage could induce any response or "resetting" of player's expectations. Nor is something as simple and transparent as the virtual flip of 5¢ coin likely to confuse players or overwhelm their capacity to reason. Absent additional evidence, we can only speculate. But the theory that strikes us as the place to start is what we're tentatively calling the "common enemy" effect. Though a random number generator is a singularly bloodless opponent, each player knew from the start that this outsider was meddling with the welfare of the entire group, punishing every member 50% of the time, and immune to persuasion. Few human tendencies 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 – initial seating patterns, randomly distributed colors, irrelevant preferences, not to mention hair color, eye color, small age differences, and countless more substantive differences. In seeking to create a neutral alternative to our AI treatment, we may have stumbled on an understudied source of group solidarity. And in some cases at least, when our players got the chance to communicate they (in contrast to their AI group counterparts) explicitly admonished each other to "pool our investments and get as much of [the school's] money as possible" and others advised "ignore those -5's and just do what's best for the group." In any case, the impact of seemingly irrelevant risk is yet another topic for further study, and in a world full of risk and uncertainty, one that might influence outcomes in any number of economic experiments and real-world contexts.

In short, we view our initial results as novel, striking, and surprising – more than enough to warrant further work. Hence, the plural "god games". We hope others will play along.

References

BECK, J. and FORSTMEIER, W. (2007). Superstition and belief as inevitable by-products of an adaptive learning strategy. *Human Nature*, **18** (1), 35–46.

BENJAMIN, D. J., CHOI, J. J. and FISHER, G. (2016). Religious identity and economic behavior. *Review of Economics and Statistics*, **98** (4), 617–637.

BERGER, P. L. (1969). *The social reality of religion*. Faber London.

BLEAK, J. L. and FREDERICK, C. M. (1998). Supersititious behavior in sport: Levels of effectiveness and determinants of use in three collegiate sports. *Journal of Sport Behavior*, **21** (1), 1.

BROOKS, B. A., HOFF, K. and PANDEY, P. (2016). Can the culture of honor lead to inefficient conventions? experimental evidence from india.

BUTERA, L. and LIST, J. A. (2017). *An Economic Approach to Alleviate the Crises of Confidence in Science: With an Application to the Public Goods Game*. Tech. rep., National Bureau of Economic Research.

CHAUDHARY, L. and RUBIN, J. (2016). Religious identity and the provision of public goods: Evidence from the indian princely states. *Journal of Comparative Economics*, **44** (3), 461–483.

CHAUDHURI, A. (2011). Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Experimental Economics*, **14** (1), 47–83.

— (2016). Recent advances in experimental studies of social dilemma games.

FISCHBACHER, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental economics*, **10** (2), 171–178.

FOSTER, K. R. and KOKKO, H. (2009). The evolution of superstitious and superstition-like behaviour. *Proceedings of the Royal Society of London B: Biological Sciences*, **276** (1654), 31–37.

GERSHMAN, B. (2016). Witchcraft beliefs and the erosion of social capital: Evidence from sub-saharan africa and beyond. *Journal of Development Economics*, **120**, 182–208.

GILOVICH, T., VALLONE, R. and TVERSKY, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive psychology*, **17** (3), 295–314.

HAYEK, F. A. (1989). The fatal conceit. edited by ww bartley, iii.

HOFFMANN, R. (2013). The experimental economics of religion. *Journal of Economic surveys*, **27** (5), 813–845.

HOLT, C. and LAURY, S. (2002). Risk aversion and incentive effects. *American Economic Review*, **92** (5), 1644–1655.

IANNACCONE, L. R. (1992). Sacrifice and stigma: Reducing free-riding in cults, communes, and other collectives. *Journal of political economy*, **100** (2), 271–291.

— (2006). The market for martyrs. *Interdisciplinary Journal of Research on Religion*, **2**.

— and BERMAN, E. (2006). Religious extremism: The good, the bad, and the deadly. *Public choice*, **128** (1), 109–129.

ISAAC, R. M., WALKER, J. M. and THOMAS, S. H. (1984). Divergent evidence on free riding: An experimental examination of possible explanations. *Public choice*, **43** (2), 113–149.

KNUTH, D. E. (1997). The art of computer programming, 3rd edn. seminumerical algorithms, vol. 2.

KURAN, T. (2012). *The long divergence: How Islamic law held back the Middle East*. Princeton University Press.

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.

— (2010). *Bourgeois dignity: Why economics can't explain the modern world*. University of Chicago Press.

— (2016). *Bourgeois equality: How ideas, not capital or institutions, enriched the world*. University of Chicago Press.

Morse, W. and Skinner, B. (1957). A second type of superstition in the pigeon. *The American Journal of Psychology*, **70** (2), 308–311.

Norenzayan, A. (2013). *Big gods: How religion transformed cooperation and conflict*. Princeton University Press.

Nunn, N. and de la Sierra, R. S. (2017). *Why Being Wrong can be Right: Magical Warfare Technologies and the Persistence of False Beliefs*. Tech. rep., National Bureau of Economic Research.

Palmer, C. T. (1989). The ritual taboos of fishermen: An alternative explanation. *MAST. Maritime anthropological studies*, **2** (1), 59–68.

Park, S. K. and Miller, K. W. (1988). Random number generators: Good ones are hard to find. *Commun. ACM*, **31** (10), 1192–1201.

R Development Core Team (2012). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.

Reberg, D., Mann, B. and Innis, N. K. (1977). Superstitious behavior for food and water in the rat. *Physiology & Behavior*, **19** (6), 803–806.

Rubin, J. (2017). *Rulers, Religion, and Riches: Why the West got rich and the Middle East did not*. Cambridge University Press.

RUFFLE, B. J. and SOSIS, R. (2006). Cooperation and the in-group-out-group bias: A field test on israeli kibbutz members and city residents. *Journal of Economic Behavior & Organization*, **60** (2), 147–163.

SKINNER, B. F. (1948). " superstition" in the pigeon. *Journal of Experimental Psychology: General*, **38**, 168–272.

STARK, R. (1999). A theory of revelations. *Journal for the Scientific Study of Religion*, pp. 287–308.

— (2005). The victory of reason. *How Christianity Led to Freedom, Capitalism, and Western Success, New York*.

— and BAINBRIDGE, W. S. (1996). *A theory of religion*. Rutgers Univ Pr.

STEADMAN, L. B. and PALMER, C. T. (2015). *Supernatural and natural selection: religion and evolutionary success*. Routledge.

TIMBERLAKE, W. and LUCAS, G. A. (1985). The basis of superstitious behavior: chance contingency, stimulus substitution, or appetitive behavior? *Journal of the experimental analysis of behavior*, **44** (3), 279–299.

WADE, N. (2009). *The faith instinct: How religion evolved and why it endures*. Penguin.

WARNER, C. M., KILINÇ, R., HALE, C. W., COHEN, A. B. and JOHNSON, K. A. (2015). Religion and public goods provision: Experimental and interview evidence from catholicism and islam in europe. *Comparative Politics*, **47** (2), 189–209.

WIKIPEDIA (). Clustering illusion. *The Free Encyclopedia*.

— (2017). Linear congruential generator — wikipedia, the free encyclopedia. [Online; accessed 3-November-2017].

WILSON, R. K. and SELL, J. (1997). "liar, liar..." cheap talk and reputation in repeated public goods settings. *Journal of Conflict Resolution*, **41** (5), 695–717.

WRIGHT, R. (2010). *The Evolution of God: The origins of our beliefs*. Hachette UK.

Appendices

A Experiment Instructions

In this section, we, first, provide the instructions for the core treatments and then we offer the communication treatments' instructions. The core treatments' instructions mainly differ in the description of the adjustment algorithm.[31] Hence, **Stage 2**'s narration is the main variation across treatments.

A.1 Instruction of Core Treatments

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.

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.

A.1.1 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.

[31]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.

A.1.2 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.)

The Instruction of Stage 2 for the 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.

The Instruction of Stage 2 for the AI Treatment

Stage 2: 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.

A.1.3 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.

The Instruction of Stage 2 for the Baseline Treatment

Your overall earnings in each period will be your earnings from the joint account plus your earnings from your personal account.

The Instruction of Stage 2 for the Random Treatment

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 Instruction of Stage 2 for the AI Treatment

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.

A.1.4 Examples for the AI Adjustment Treatment

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 0.4×20 = 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.

	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	AI chooses adjustment = +5
Overall earnings for period:	17	20	22	13	= Earnings from both accounts plus AI's adjustment

Table A1: **Table 1**

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.

	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	AI chooses adjustment = -5
Overall earnings for period:	9.1	6.1	12.6	3.6	= Earnings from both accounts plus AI's adjustment

Table A2: **Table 2**

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.

	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	AI chooses adjustment = +5
Overall earnings for period:	21	21	21	21	= Earnings from both accounts plus AI's adjustment

Table A3: **Table 3**

A.2 Instruction of the Communication Treatments

Prior to the Your Earnings stage, we add the below description of group chat to the instruction of the core treatments.

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.

A.3 Risk preference elicitation instructions

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		
A	B	Your Choice
$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

A.4 Questionnaire

The following questions were asked of all subjects (except the questions in bold, which only were asked only in the AI treatments. Question 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

B Adjustment Algorithm

In the second stage of each period, the computer adjusts each subject's earnings by adding either a high or low value, $a_1 = 5$ or $a_2 = -5$. Within each group j, all players receive the same adjustment, but adjustments vary across groups. We designed the adjustment algorithm so that simple parameter changes would yield different types of AI's.

Details: Each group of receives an independently calculated adjustment. Group j receives the high adjustment A_{jt} in period t with probability $p_{t_{jt}} = Prob(A_{jt} = a_1)$ defined as follows:

$$p_{t_{jt}} = P(G_{jt}, G^*, \bar{G}_{-jt}) \equiv \frac{1}{2}\left[b_0 + b_1(\frac{G_{jt}}{G^*}) + b_2(\frac{G_{jt}}{G_{jt} + \bar{G}_{-jt}})\right] \tag{1}$$

G_{jt} is the total invested in group j during round t, $\bar{G}_{-jt}$ is the average group investment in groups other than j during round t, G^* is the maximum possible group investment ($= n * E$ for groups of size n and endowment E), and $b_0 + b_1 + b_2 = 1$. Hence the probability of group j getting a high adjustment is the weighted average of a random component, group j's investment relative to the maximum possible investment, and group j's investment relative to the average investment of the other groups.

Group j then receives the high adjustment in period t if a random number calculated by the ztree function Random( ) is less than the probability threshold determined by the P function above. For just two adjustment levels, the probability of the low adjustment is $1 - P$. Each group receives its own separately calculated adjustment.

The Random( ) function in ztree is based on a C++ pseudo-random number generator, which is one of the class of functions known as "linear congruential generators" (Wikipedia, 2017). Computing the values of such functions is straightforward, but as Knuth (1997), Park and Miller (1988), and others have demonstrated, the behavior of the function depends critically on the choice of underlying parameters, determining the degree of randomness associated with any given functional form is mathematically challenging, and identifying optimal parameters for a given functional form is quite difficult.

To extend the adjustment algorithm to $n \geq 2$ adjustments we choose a sequence of adjustments $(a_1 > a_2 > ... > a_n)$ and associated probabilities $(p_1, p_2, ..., p_n)$ such that $p_1 = P(G_{jt}, G^*, G_{-jt})$ and $p_k = \beta^k P$ for $k = 1...n$ with β determined by the fact that the probabilities must sum to 1.

Changing the b weights yields different types of AI's. For example, the adjustment process approaches pure randomness (with equal probability of high or low adjustment) as b_0 approaches 1, and the probability of a high adjustment approaches the size of j's group investment relative to the maximum possible investment as b_1 approaches 1.

To implement an AI that takes account of past and current behavior, we simply redefine $p_{1_{jt}}$ as functions of P in periods t and earlier.

Setting $b_1 = 1$ and $b_1 = b_2 = 0$ yields random treatments like those reported in this paper. To explore responses to "effectively" random AI's, as in this paper, it suffices to set b_0 near 1.

To see whether behavior changes when the probability of a higher adjustment is more closely linked to a group's relative and absolute rates of investment, we are currently running additional AI treatments with $b_0 = b_1 = b_2 = 1/3$ and reporting the results in the next draft of this paper.

C Complex AI with Communication

In this treatment we keep all the aspects of the AI with communication treatment; however, we modify the adjustment algorithm so that there are 5 different values for the monetary prizes $\theta \in \{-5, -2, 0, 2, 5\}$. Below is the instructions for this treatment:

> **Stage 2**: 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, +2, 0, -2, 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 and 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.

We ran 3 sessions of this treatment (henceforth complex AI with communication treatment). We had 16 subjects per session for a total of 48 subjects. The average total payment was 24.25$ with the minimum payment of 20.25$ and maximum payment of 28.25$. The mean, standard error, and median of the group investment levels in this treatment are 32.56, 0.618, and 40, respectively.

While the mean level of the group investment in the complex AI with communication treatment is not significantly different from AI with communication treatment (two-sided MWW rank sum test, $p - value = 0.73$), The random with communication treatment significantly promotes higher level of investment than the complex AI with communication treatment ($p - value = 0.03$).[32]

[32]Considering the median group investments yield similar results. That is, the median investment levels in complex AI with communication is not significantly different from the AI with communication treatment ($p - value = 0.57$)and it is significantly lower than the random with communication treatment ($p - value = 0.03$).

D Additional Figures and Robustness Checks

	(1) $Invetment_t$	(2) $Invetment_t$
Random	1.378**	2.091**
	(0.607)	(0.857)
AI	1.487**	1.435
	(0.641)	(1.102)
Risk		1.291
		(1.287)
Random × Risk		-2.292
		(1.751)
AI × Risk		-0.714
		(2.425)
Period		-0.140***
		(0.018)
Random × Period		0.029
		(0.028)
AI × Period		0.035
		(0.026)
Female		-0.680
		(0.416)
Intercept	2.748***	4.037***
	(0.494)	(0.808)
Observations	3840	3840
R^2	0.04	0.10

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

Table D1: **GLS analysis of Investment Decision**.

	(1) Absolute Change	(2) No Change	(3) No Change No corner solution
Random	-0.136	0.016	-0.002
	(0.223)	(0.046)	(0.045)
Adjustment$_{t-1}$	-0.850***	0.162***	0.170***
	(0.128)	(0.033)	(0.033)
Random × Adjustment$_{t-1}$	0.338**	-0.082**	-0.091**
	(0.170)	(0.036)	(0.040)
Intercept	2.188***	0.302***	0.193***
	(0.139)	(0.034)	(0.035)
Observations	2432	2432	2065
R^2	0.02	0.02	0.03

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

Table D2: **GLS analysis of Absolute and No-Change in Investment Decision**.